



Competition between Visual Events Modulates the Influence of Saliency during Free-Viewing of Naturalistic Videos

Davide Nardo^{1,2*}, Paola Console¹, Carlo Reverberi^{3,4} and Emiliano Macaluso^{1,5}

¹ Neuroimaging Laboratory, Santa Lucia Foundation, Rome, Italy, ² Institute of Cognitive Neuroscience, University College London, London, UK, ³ Department of Psychology, University of Milano-Bicocca, Milan, Italy, ⁴ NeuroMi—Milan Center for Neuroscience, University of Milano-Bicocca, Milan, Italy, ⁵ Impact Team, Lyon Neuroscience Research Center, Lyon, France

In daily life the brain is exposed to a large amount of external signals that compete for processing resources. The attentional system can select relevant information based on many possible combinations of goal-directed and stimulus-driven control signals. Here, we investigate the behavioral and physiological effects of competition between distinctive visual events during free-viewing of naturalistic videos. Nineteen healthy subjects underwent functional magnetic resonance imaging (fMRI) while viewing short video-clips of everyday life situations, without any explicit goal-directed task. Each video contained either a single semantically-relevant event on the left or right side (Lat-trials), or multiple distinctive events in both hemifields (Multi-trials). For each video, we computed a saliency index to quantify the lateralization bias due to stimulus-driven signals, and a gaze index (based on eye-tracking data) to quantify the efficacy of the stimuli in capturing attention to either side. Behaviorally, our results showed that stimulus-driven saliency influenced spatial orienting only in presence of multiple competing events (Multi-trials). fMRI results showed that the processing of competing events engaged the ventral attention network, including the right temporoparietal junction (R TPJ) and the right inferior frontal cortex. Saliency was found to modulate activity in the visual cortex, but only in the presence of competing events; while the orienting efficacy of Multi-trials affected activity in both the visual cortex and posterior parietal cortex (PPC). We conclude that in presence of multiple competing events, the ventral attention system detects semantically-relevant events, while regions of the dorsal system make use of saliency signals to select relevant locations and guide spatial orienting.

Keywords: spatial attention, eye movements, real world, saliency, competition, fMRI, MVPA

INTRODUCTION

In any everyday life situation the sensory system receives more information from the outside world than the brain can fully process, with many objects and events that compete for the limited processing resources. Spatial attention control allows focussing on signals located in one region of space, selecting relevant information for in-depth processing (Desimone and Duncan, 1995). Both sensory characteristics of the external stimuli, and internal signals related to current goals, intentions, and expectations contribute to spatial orienting control (Posner, 1980; Yantis, 2000).

OPEN ACCESS

Edited by:

Carol Seger,
Colorado State University, USA

Reviewed by:

Radek Ptak,
Geneva University Hospital,
Switzerland
Joy Geng,
University of California Davis, USA

*Correspondence:

Davide Nardo
davidenardo@gmail.com

Received: 30 March 2016

Accepted: 13 June 2016

Published: 28 June 2016

Citation:

Nardo D, Console P, Reverberi C and Macaluso E (2016) Competition between Visual Events Modulates the Influence of Saliency during Free-Viewing of Naturalistic Videos. *Front. Hum. Neurosci.* 10:320. doi: 10.3389/fnhum.2016.00320

Here, using short video-clips of real-world scenes we investigated how the competition between visual events affects overt spatial orienting in the absence of any explicit goal-directed task, and how such competition modulates the impact of stimulus salience on orienting behavior and brain activity.

Behaviorally, the issue of competition in spatial attention has been often addressed using visual search tasks. Subjects are required to search for a given target (e.g., a letter) within a visual array containing many distractors that compete for processing (i.e., non-target letters). When the target is characterized by a very distinctive feature (e.g., a red target-letter among green distractor-letters), the target seems to “pop-out” of the visual array, the detection is fast and the search times are largely independent of the number of distractors (efficient search). In this situation, the sensory characteristics of the visual input guide the selection process and attention is attracted automatically towards the salient pop-out target (stimulus-driven control, Theeuwes, 1992; Yantis, 1993). By contrast, when a target is defined by a combination of features, some of which are shared with distractors, the search is slow and dependent on the number of distractors (inefficient search). In this case, to solve the competition between the target and the distractors, spatial attention is shifted sequentially between the display items until the target is found, or all the items have been classified as distractors (Treisman and Gelade, 1980; Treisman, 1986; but for criticisms, see Wolfe et al., 1989; Cave and Wolfe, 1990). The selection process is governed primarily by internal knowledge about the target-defining features and attention is controlled in a voluntary manner (goal-directed control, Bacon and Egeth, 1997; see also Duncan and Humphreys, 1989, 1992).

Non-invasive imaging studies provide us with extensive evidence about the neurophysiological bases of visuospatial control in humans (for reviews, see Corbetta and Shulman, 2002; Vossel et al., 2014). For example, Kastner et al. (1999) investigated how voluntary control of spatial attention influences the processing of competing visual stimuli. Using functional magnetic resonance imaging (fMRI), the authors showed that the simultaneous presentation of several irrelevant stimuli in one visual quadrant lead to the suppression of the corresponding functional responses in the visual cortex, compared to the sequential presentation of the same stimuli (“competitive interactions”). However, when the participants directed voluntary attention to the stimulated quadrant such interactions were reduced (cf. “Biased Competition Theory” of selective attention, Desimone and Duncan, 1995). Besides these effects of competition in the visual cortex, the same study also highlighted the involvement of the dorsal frontoparietal cortex that is assumed to control goal-directed, voluntary attention (Kastner and Pinsk, 2004; Pinsk et al., 2004; Foley et al., 2012). On the other hand, regions in the right inferior frontoparietal cortex (i.e., the temporoparietal junction, R TPJ; and the inferior frontal gyrus, R IFG) have been associated with stimulus-driven control. These regions activate when participants are presented with isolated targets at an unattended location (e.g., invalid trials in spatial cueing tasks; Arrington et al., 2000; Corbetta et al., 2000). Related to the issue of attentional selection between competing stimuli, Serences et al. (2005) showed activation of the R TPJ in

response to peripheral distractors while participants performed a central detection task, but only when the distractors contained some goal-related, task-relevant target-defining feature (i.e., distractors of the same color as the target).

These and many other imaging studies provide us with important insights about brain regions involved in voluntary and stimulus-driven attention, but the vast majority of these studies involved simple stimuli and very specific, goal-directed tasks. Typically, participants are instructed to detect or discriminate a given target item (e.g., a letter or some other simple shape) and—for hundreds of trials—they will process the incoming visual input with the sole scope of detecting/discriminating that specific stimulus. This approach allows controlling for many experimental factors, such as the number, size and position of the stimuli. However, it should be considered that attention control in real-life has to face a rather different set of constraints (e.g., see Wolfe et al., 2011; Peelen and Kastner, 2014). In the real world the attentional system has to deal with highly complex, variable and dynamic sensory input. Further, while both stimulus-driven and goal-related signals contribute to attention control (see below), we seldom explore the environment with the sole scope of detecting one specific object or event: that is, following some specific task-instructions. Instead, we have a vast amount of prior internal knowledge about the spatial layout of objects in the real world, which provides us with additional information for the allocation of spatial attention (Castelhano and Heaven, 2011; Spotorno et al., 2014; Wu et al., 2014).

One approach to investigate attention control in more realistic, life-like conditions is to utilize pictures or videos of real-world scenes (e.g., Carmi and Itti, 2006; Summerfield et al., 2006; Einhäuser et al., 2007; Nardo et al., 2011, 2014; Malcolm and Shomstein, 2015; Ogawa and Macaluso, 2015; for a review, see also Peelen and Kastner, 2014). Many studies have now demonstrated that both the sensory characteristics of the visual input, as well as internal information (such as task-related goals, prior knowledge and expectations) contribute to the processing of such complex stimulus material. Stimulus-driven signaling can be investigated using computational models that seek to extract salient locations in the image (e.g., “saliency maps” proposed by Itti et al., 1998; Itti and Koch, 2001). Saliency maps characterize the stimulus-driven contribution to spatial orienting and have been shown to successfully predict patterns of eye movements during free-viewing of complex scenes (Carmi and Itti, 2006; Elazary and Itti, 2008) and to modulate brain activity associated with spatial attention (e.g., Bogler et al., 2011; Nardo et al., 2011; Santangelo and Macaluso, 2013). At the same time, higher-level factors strongly influence spatial orienting in natural scenes. When participants are asked to search for a target-object defined by specific task-instructions, prior knowledge about where that object can be usually found makes the search more efficient (Evans et al., 2011; Greene and Wolfe, 2011; Wolfe et al., 2011), and the discrimination between targets and distractors is largely dependent on semantic knowledge (Peelen et al., 2009; Draschkow et al., 2014; Seidl-Rathkopf et al., 2015). These internal factors can influence scanpaths (Humphrey and Underwood, 2009; Malcolm and Henderson, 2010) and reduce the impact of stimulus-driven signals during overt exploration

of real-world scenes (Einhäuser et al., 2008; Stoll et al., 2015). Nonetheless, it should be noted that most of the previous studies that investigated spatial attention with naturalistic stimuli made use of explicit goal-directed tasks (e.g., search for a specific target-object within real scenes, Neider and Zelinsky, 2006; Castelhana and Heaven, 2010; Malcolm and Henderson, 2010; see also Santangelo et al., 2015), which is likely to influence the neural systems involved in control operations (see section above about regions associated with goal-directed attention).

Here we investigated the role of stimulus-driven signals for spatial orienting in natural scenes, specifically in the absence of any explicit task. In a previous fMRI study that involved free-viewing of a virtual visual environment (i.e., without any goal-directed task), we reported activation of the dorsal frontoparietal network when stimulus-driven signals successfully attracted subjects' gaze/attention (Nardo et al., 2011; "efficacy" of salience for spatial orienting). The same study also showed that semantically-relevant, gaze/attention-grabbing events (moving avatars, in that study) were processed in the ventral attention system, indicating a possible dissociation between the dorsal and the ventral frontoparietal systems when orienting attention within complex visual environments, in the absence of any explicit goal-directed task. In the current study, we extend this line of investigation by specifically enquiring how competing visual events are processed in the attention control networks, and how such competition interacts with stimulus-driven (salience) signaling.

In the current fMRI study, participants were presented with short (1.5 s) video-clips that included either multiple semantically-relevant events on both sides of space (Multi-trials: strong competition), or a single main event lateralized on the left or right hemifield (Lat-trials: weak competition; for a few examples, see **Figure 1A**). Participants were asked to simply watch the videos and were allowed to move their eyes, that is, a naturalistic free-viewing situation without any explicit goal-directed task. For each stimulus, we quantified the strength of the stimulus-driven signals using saliency maps (Itti et al., 1998). We created a saliency lateralization-index (Sal_idx) as the difference between the saliency in the two hemifields. In order to index to what extent each stimulus was able to drive spatial attention to either side of space we used eye-movements recorded during the viewing of the videos. We computed a gaze lateralization index (Gaze_idx) as the difference in time the participants spent looking towards the two hemifields.

Behaviorally, we assessed the contribution of stimulus-driven signals to spatial orienting testing for the relationship between the saliency index and the gaze index, separately for videos with multiple vs. single-lateralized events. We predict that in Lat-trials participants would systematically orient towards the side of the single semantically-relevant event (cf. Einhäuser et al., 2008; Nuthmann and Henderson, 2010; Stoll et al., 2015). If so, the "video-specific" level of lateralization of saliency will have little effect on the time spent looking towards the relevant hemifield, without any correlation between the saliency and gaze indexes for these trials. By contrast, we expect that stimulus-driven saliency would contribute to spatial orienting in Multi-trials, when the competition between co-occurring visual events will

reduce any spatial orienting bias associated with the distinctive events. Now, videos with high levels of saliency lateralization should lead to greater gaze-orienting towards the most salient hemifield, compared with Multi-trials entailing low levels of saliency lateralization. The latter would result in a significant correlation between the saliency and the gaze indexes in Multi-trials.

For the analyses of imaging data, we first tested for the overall effect of competition between distinctive events by directly comparing Multi-trials vs. Lat-trials. We used both standard univariate methods and multivariate pattern analysis (MVPA). The latter has several advantages over univariate methods (Norman et al., 2006) and is gaining increasing interest in cognitive neuroscience (for reviews, see Mahmoudi et al., 2012; Serences and Saproo, 2012; Haynes, 2015). Recently, we have shown that MVPA can unveil the role of specific nodes of the attentional system (i.e., processing of invalid trials in the left TPJ), which would be missed using traditional subtraction methods (Silvetti et al., 2015; for recent applications of MVPA in studies with naturalistic stimuli, see also Bogler et al., 2011; Epstein and Morgan, 2012; Preston et al., 2013; Johnson and Johnson, 2014; Linsley and MacEvoy, 2014; Watson et al., 2014).

We formulated two main hypotheses. On the one hand, if the high levels of competition in Multi-trials is handled via sequential orienting towards the multiple visual events, we would anticipate the engagement of the ventral attention system, which we previously associated with the detection of distinctive events in free-viewing conditions (Nardo et al., 2011). Further, if stimulus-driven signaling takes effect to solve competition in Multi-trials (cf. also section of behavioral predictions, above), we may expect the engagement of the dorsal attention system that have been found activated when visual saliency is used to guide spatial orienting and attentional selection (Bogler et al., 2011; Nardo et al., 2011; see also Santangelo and Macaluso, 2013). The further characterization of the response of the frontoparietal attention systems in these naturalistic conditions will contribute to bridge the gap between traditional experiments in the laboratory and brain functioning in the real world (Hasson et al., 2004; Peelen and Kastner, 2014).

MATERIALS AND METHODS

Subjects

This study is part of a larger research project about spatial attention control in people who suffered from stroke to the right hemisphere (neglect patients). In this framework, we recruited a cohort of healthy controls who will be later compared with neurological patients. Here, we report the data of 19 healthy aged volunteers (i.e., all the healthy subjects who participated in the stroke project; age range: 44–81 years, mean age: 62 ± 12.1 ; 10 males and 9 females). Their inclusion criteria were: right-handedness, no history of psychiatric or neurological disease or drug abuse, and normal or corrected-to-normal (contact lenses) visual acuity. After having received instructions, all participants gave their written informed consent. The study was approved by the independent Ethics Committee of the Santa Lucia

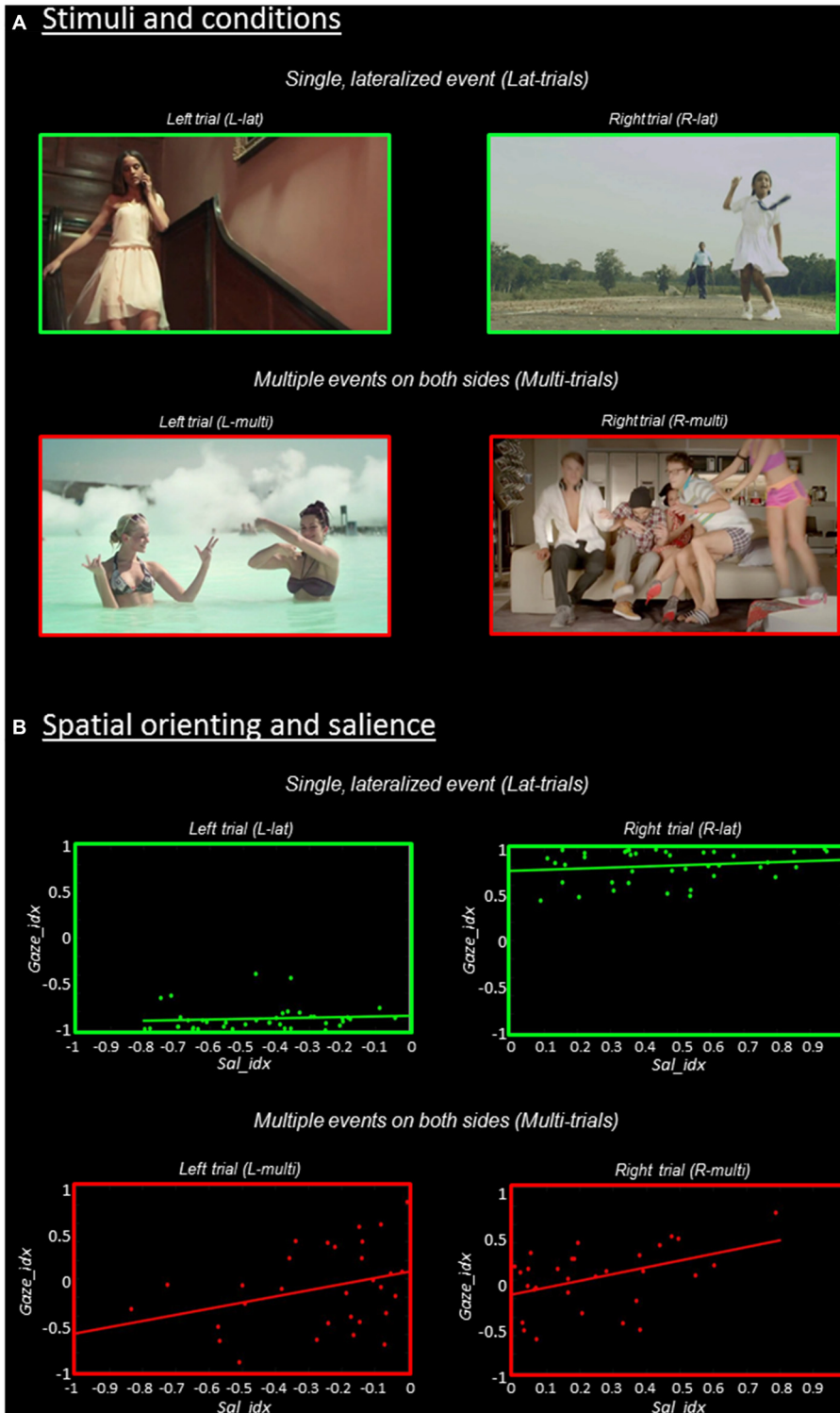


FIGURE 1 | Stimuli and behavioral results. (A) Stimuli and experimental conditions. Single frames extracted from the videos that contained a single, lateralized visual event are shown in the upper panel (“Lat-trials”, green borders); while frames extracted from the videos including multiple events on both sides are shown in the lower panel (“Multi-trials”, red borders). The further categorization in “Left” vs. “Right” trials was done based on saliency maps (cf. “Materials and Methods” Section). **(B)** Relationship between spatial orienting (Gaze_idx) and stimulus salience (Sal_idx), separately for the four experimental conditions. The Gaze_idx correlated with Sal_idx only in trials containing multiple competing events (“Multi-trials”, plots in red), while in Lat-trials the side of the main visual event largely determined spatial orienting (cf. plots in green). For both Gaze_idx and Sal_idx positive values indicate a rightward bias, while negative values indicate a leftward bias.

Foundation (Scientific Institute for Research, Hospitalization and Health Care).

Stimuli and Saliency Index

The set of stimuli consisted of 140 videos showing everyday life situations. We obtained a collection of non-Italian TV commercial clips, partly purchased from an Advertising Archive¹ and partly downloaded from YouTube (see **Figure 1A** for a few screen shots). Using a video editing software (Final Cut Pro, Apple Inc., Cupertino, CA, USA), we selected 1.5 s video-segments that included a single continuous scene with either one lateralized distinctive event (Lat-trials) or multiple events in both hemifields (Multi-trials). Most of the distinctive, semantically-relevant events consisted in one or more persons in the foreground, who either performed an action (walking, dancing, manipulating objects, etc.) or changed posture. In approximately 10% of videos, the event consisted in a moving vehicle (plane, car, motorbike, etc.; equally distributed across conditions). We discarded any segments that included writings, which allowed us to left-right flip the videos as an additional experimental control (cf. the experimental factor of “left/right”, below). Stimuli were presented in a randomized order, balancing the video flipping across subjects. The inter-trial intervals ranged from 3 to 7.4 s (mean 5.2). During the inter-trial interval a fixation point was presented in the center of the screen.

For each stimulus, we indexed the strength of stimulus-driven signals using saliency maps (Itti et al., 1998; Itti and Koch, 2001). For each frame, we obtained a saliency map based on multi-scale contrasts of intensity, color, orientation, motion and flicker, as computed using the MT_TOOLS toolbox.² Next, for each video we generated a single value (Sal_idx) that represents the level of lateralization of the saliency signals. The Sal_idx was computed by averaging the saliency values separately in the left and the right side of each map (excluding a central region of 2°, cf. also “gaze index” below) and then averaging these values over the entire stimulus duration. Finally, the level of saliency lateralization was computed as a single normalized coefficient $\text{Sal_idx} = (L - R) / (L + R)$, with negative/positive values indicating greater saliency on the left/right side of the video (see also Nardo et al., 2014).

The Sal_idx was used for both behavioral analyses—that is, to evaluate the contribution of stimulus-driven signaling for gaze-orienting (cf. Gaze_idx, below)—and for correlational analyses using fMRI BOLD time-series. In addition, the Sal_idx allowed us to categorize each stimulus as either “Left” or “Right”, which was included as a factor in the modeling of the fMRI data (L- vs. R-trials), together with the main factor concerning the level of competition (Multi- vs. Lat-trials). The full set of videos included 80 lateralized trials and 60 multiple trials, equally split into left- and right-trials.

In addition, we examined the saliency maps to gain some further insight about the spatial organization of the scenes in the two main conditions (Multi- and Lat-videos). Despite saliency maps are based on local discontinuities in low-level features (e.g., color, orientation, motion, etc.), previous work showed that such

discontinuities tend to correspond to “interesting objects” in the scene (Elazary and Itti, 2008). Therefore, here we predicted that the Multi-trials should include more saliency clusters as compared with Lat-trials. For each video, we computed the number of saliency clusters by adding the saliency maps associated with each video-frame and counting the number of saliency cluster per video (i.e., continuous patches of pixels with saliency values larger than zeros). We compared the number of saliency clusters in Multi- vs. Lat-trials using a two-sample *t*-test. With these 2D maps we also further quantified the overt orienting behavior of the participants, by asking the additional question of whether the likelihood to fixate a salient location/cluster changed as a function of Multi- vs. Lat-conditions. For each video, we computed the percentage of fixations falling inside the saliency clusters and compared these values between Multi- vs. Lat-trials using a two-sample *t*-test.

Eye-Movements and Gaze Index

Subjects viewed the 140 videos twice, first in a behavioral session outside the MR scanner and then 1 week later during the fMRI experiment. Behavioral data were collected as part of the research project that included comparisons with brain-damaged patients (cf. also “Subjects” section, above). The behavioral session was carried out in a quiet, dimly lit room. Subjects seated in front of a laptop equipped with a portable eye-tracking system operating at 120 Hz (RED-m Eye Tracking System 3.2; SensoMotoric Instruments³), at a distance of approximately 60 cm from the screen. Subjects were simply asked to freely-view the stimuli. During the fMRI sessions, gaze-position was recorded with a long-range MR-compatible eye-tracking system operating at 60 Hz (Applied Science Laboratories, Bedford, MA, USA; Model 504). The videos were back-projected onto a screen at the back of the MR bore, visible to the subjects via an MR-compatible mirror. The horizontal visual angle was approximately 24° pre-scanning, and ca. 19° in-scanner. The eye-tracking data were processed with custom scripts using MT_TOOLS⁴. Fixations were defined as gaze-position remaining within an area of 1.5° × 1.5° for a minimum duration of 100 ms.

In order to obtain the most accurate indexes of orienting efficacy, we imposed several constraints to select the eye-tracking data that were included for the computation of the video-specific Gaze indexes. For each subject and each trial/video, we considered only eye-traces where the pre-stimulus gaze-position was within ±2° of the center of the screen and at least 50% of data-points during the presentation of video could be categorized as fixations (i.e., excluding trials including many blinks and/or other artifacts). We counted how many trials satisfied these criteria and selected participants who had at least six trials for each of the four experimental conditions. The number of participants who fitted these criteria was 11 for data acquired during scanning vs. 16 in the pre-scanning session. Therefore, for the computation of the Gaze_idx we used the eye-tracking data of 16 participants in the pre-scanning session (cf. also Nardo et al.,

¹<http://www.coloribus.com>

²<http://www.slneuroimaginglab.com/mt-tools>

³<http://www.smivision.com/en/gaze-and-eye-tracking-systems/products/redm.html>

⁴<http://www.slneuroimaginglab.com/mt-tools>

2011; where we validated the use of eye-movement data collected outside the scanner to compute indexes of spatial orienting). Note that the pre- vs. during-scanning Gaze_idx were highly correlated ($r = 0.92$; $p < 0.001$).

The Gaze_idx represents the efficacy of each video to generate a systematic shift of overt attention towards either side. For each video, the index was computed as the proportion of time that participants spent looking towards the two hemifields, that is, $(Ltime - Rtime)/(Ltime + Rtime)$ (see Nardo et al., 2014). Since any small deviation of horizontal gaze-position around the center of the screen (even below the spatial precision of gaze measurement) would strongly affect this index, the gaze data-points falling within $\pm 2^\circ$ of the center of the screen were not considered for the computation of the L/R-time values.

In addition, for each video we computed the number of saccades that the 16 participants made during free-viewing of the videos in the pre-scanning session. For each video, the number of saccades was defined as the number of fixations (see above) minus one. Under the hypothesis that the presence of multiple distinctive events would lead to sequential exploration of the different events, we compared Multi- vs. Lat-trials expecting a larger number of saccades in the former.

fMRI Data Acquisition and Preprocessing

A Siemens Allegra (Siemens Medical Systems, Erlangen, Germany) 3 T scanner equipped for echo-planar imaging (EPI) was used to acquire fMRI. A quadrature volume head coil was used for radio frequency transmission and reception. Head movement was minimized by mild restraint and cushioning. Thirty-two slices of fMRI were acquired using blood-oxygen-level dependent (BOLD) imaging (3×3 mm, 2.5 mm thick, 50% distance factor, repetition time (TR) = 2.08 s, echo time (TE) = 30 ms), covering the entirety of the cortex. We also acquired a Magnetization-Prepared Rapid Gradient-Echo (MPRAGE) sequence as an anatomical reference (TR = 2.5 s, TE = 2.74 ms, voxel size $1 \times 1 \times 1$ mm, matrix resolution $256 \times 256 \times 176$, axial acquisition). Each subject underwent two fMRI runs, each including 230 volumes. Data preprocessing was performed with SPM12b (Wellcome Department of Cognitive Neurology). The first four volumes of each run were discarded to allow T1 saturation effects. Images were first manually realigned along the anterior commissure-posterior commissure (AC-PC) axis, then realigned, unwarped, and slice-timed with the middle slice as a reference.

Standard Univariate Analysis

Standard univariate fMRI analysis was performed with SPM12b, and included first- and second-level analyses. Stimuli were modeled as delta functions (duration = 1.5 s, corresponding to the duration of the videos), convolved with the standard hemodynamic response function (HRF). The first-level model included four conditions, obtained by combining the factors “competition” (Multi- vs. Lat-trials) and “side” (L- vs. R-trials). All models included the head-motion realignment parameters as additional covariates of no interest. The time series were high-pass filtered at 128 s and pre-whitened by means of autoregressive

model AR(1). Linear contrasts were used to average the parameter estimates across the two fMRI-runs, separately for the four conditions of interest (Lmulti, Rmulti, Llat, Rlat). To allow for group-level analyses, the resulting four contrast images per subject were normalized to the Montreal Neurological Institute (MNI) space. The normalization parameters were estimated from the (co-registered and segmented) individual T1 volume, re-sampled ($3 \times 3 \times 3$ mm) and smoothed with a Gaussian kernel (Full width at half maximum, FWHM = $8 \times 8 \times 8$ mm), including the SPM12b brain-mask as an explicit mask.

For second-level group analysis, a flexible factorial design modeled the four conditions of interest (Lmulti, Rmulti, Llat, Rlat), plus the main effect of subjects. Sphericity correction was applied to account for any non-independent error term for repeated measures and any difference in error variance across conditions (Friston et al., 2002).

We report the main contrast of interest that directly compared the conditions with multiple events (“Multi”, high competition) vs. conditions with a single distinctive event (“Lat”, low competition), averaging across the factor of side. The level of significance was set to p -Family-wise error (FWE)-corr. = 0.05, cluster-level corrected for multiple comparison (cluster size estimated at p -unc. = 0.001). We also report some effects at a lower p -unc. < 0.001, with the aim of evaluating the results of the univariate analysis against the results of the multivariate MVPA approach (see below, and Table 1).

Multi-Voxel Pattern Analysis

Multivariate analyses were performed with the Decoding Toolbox⁵. At the subject-level, with unnormalized and unsmoothed data, we performed MVPA, using a Linear Support Vector Machine (LSVM) classifier with a 6-fold cross-validation procedure. We used the stimulus/video conditions as classes and the parameter estimates (beta-images) derived from individual fMRI first-level analyses as features. For the MVPA analysis the first-level models were re-constructed now including three separate regressors for each of the four condition (Lmulti, Rmulti, Llat, Rlat). Specifically, for each fMRI-run the trials belonging to each condition were pseudo-randomly split into three sub-groups. The pseudo-randomization took into account the possible influence of the previous stimulus by ensuring, for each trial, that the previous trial belonged to one of the four conditions with equal probability. Moreover, we balanced the overall temporal distribution of the trials in the three sub-groups by balancing the number of trial-onsets from the first and the second half of the run. With this, for each fMRI-run, we obtained three regressors modeling the two Lat-conditions (including 6 or 7 trials) and three regressors modeling the Multi-conditions (5 trials per regressor). Accordingly, the new first-level models included 12 regressors of interest per run, which—over the two runs—enabled us to obtain six parameter estimates (beta-images) for each of the four conditions of interest.

These 24 beta-images were divided into six image-sets, with each set containing one beta-image per condition (i.e., Llat, Rlat, Lmulti, Rmulti). To avoid over-fitting and biased

⁵<https://sites.google.com/site/tdtdecodingtoolbox>

TABLE 1 | Results of the univariate and multivariate analyses (MVPA; cross-validation).

Region	Univariate analysis (Multi > Lat)				Multivariate analysis (Multi vs. Lat)				% acc
	Cluster		Peak		Cluster		Peak		
	<i>p</i>	<i>k</i>	<i>t</i>	<i>x, y, z</i>	<i>p</i>	<i>k</i>	<i>t</i>	<i>x, y, z</i>	
R calcarine	<0.001	1287	6.85	15 -88 5	<0.001	3637	7.01	12 -79 11	65.4
R lingual			7.46	18 -70 -7			5.99	27 -67 -4	58.1
L calcarine			7.14	-9 -88 2			5.96	-9 -85 -4	60.7
L lingual			7.54	-9 -79 -10			6.19	-12 -79 -1	62.0
L/R PCN	0.008	114	5.12	3 -55 50			6.96	6 -61 56	61.3
R LOC	0.408	30	4.40	45 -64 17			4.61	48 -70 2	62.1
L LOC	0.731	17	4.24	-51 -76 5			7.09	-48 -73 2	59.9
R SOC	-	-	-	-			5.85	30 -67 29	62.7
L SOC	-	-	-	-			5.62	-21 -73 29	60.1
R TPJ	0.621	21	4.11	54 -52 23			4.14	48 -58 20	61.9
R STS	-	-	-	-	0.001	97	5.65	48 -28 -13	51.4
R IFG/MFG	-	-	-	-	<0.001	194	5.19	36 5 26	54.0

Multi: videos with multiple events; Lat: videos with a single lateralized event. Cluster: FWE-corrected *p*-value and cluster size (*k* = number of voxels). Peak: *t*-statistics and *x, y, z*-coordinates in MNI space. % acc: average accuracy of the MVPA. Regions: R/L, right/left hemisphere; PCN, precuneus; LOC, lateral occipital cortex; SOC, superior occipital cortex; TPJ, temporoparietal junction; STS, superior temporal sulcus; IFG/MFG, inferior/middle frontal gyri. In italics: areas fully significant in MVPA, but activated only at uncorrected *p*-values (*p* < 0.001) in the univariate analysis.

accuracy estimation, we applied a leave-one-out cross-validation procedure. A classifier was trained to discriminate between Multi (Lmulti, Rmulti) and Lat (Llat, Rlat) conditions using five of the six image-sets available. The classifier was then tested on the last remaining set. Thus, for each run of cross-validation, the classifier was tested over four beta-images, one for each of the four experimental conditions (Llat, Rlat, Lmulti and Rmulti).

The MVPA was performed with a searchlight approach (sphere with a radius of four voxels). The searchlight considered only gray-matter voxels. The latter were identified using the individual T1 volume that was segmented, smoothed (FWHM = 4 × 4 × 4 mm), re-sampled to the EPI resolution (3 × 3 × 3.75 mm) and thresholded to identify the voxels belonging to the gray matter (probability threshold = 10%). For each subject, the MVPA produced an accuracy map including the decoding accuracy above chance (i.e., accuracy minus 50%).

For statistical inference at the group-level, these accuracy maps were normalized to the MNI space. The normalization parameters were estimated from the (co-registered and segmented) individual T1 volume, re-sampled (3 × 3 × 3 mm) and smoothed with a Gaussian kernel (FWHM = 4 × 4 × 4 mm), including the SPM12b brain-mask as an explicit mask. Statistical inference was obtained with a one-sample *t*-test that assessed where the classification accuracy was larger than chance using between-subjects variance. The level of significance was set to *p*-FWE-corr. = 0.05, cluster-level corrected for multiple comparison (cluster size estimated at *p*-unc. = 0.001).

Parametric Modulations Analyses

We used the Sal_idx and the Gaze_idx to identify any brain region where activity was modulated according to stimulus-driven saliency and orienting efficacy, as a function of the level of competition in the sensory input (Multi- vs. Lat-trials). For this, the two indexes were now included as video/trial-specific parametric modulators. This approach utilizes inter-stimulus variability to identify any region where the BOLD signal varied

according to the lateralization of the stimulus-driven signals (co-variation with Sal_idx) or with the efficacy of the stimuli in attracting gaze/attention towards one hemifield (co-variation with Gaze_idx). For each run of each subject, the indexes were scaled to a 0–1 range before entering the fMRI analyses. The latter step was done so as to avoid comparing parameter estimates derived from predictors that included different amounts of variance (cf. gaze_idx in Lat-trials, with values always around +0.8 or –0.8; see **Figure 1B**, in green).

Separately for the two indexes, we re-estimated the first-level models now adding one regressor for each of the four conditions, cf. univariate model above. The additional regressors modeled the level of lateralized saliency/gaze-efficacy for each single trial (parametric modulators). For each condition, linear contrasts averaged the parameter estimates associated with the saliency (or gaze) index across the two fMRI runs. At the group level, two separate analyses considered the Sal_idx and the Gaze_idx. Flexible factorial designs modeled four conditions of interest corresponding to the effect of Sal_idx (or Gaze_idx) separately in Lmulti, Rmulti, Llat, Rlat trials, plus the main effect of subjects. These models allowed us to test for regions where activity co-varied with saliency (or gaze) differentially in videos with high vs. low competition (i.e., main effect “Multi > Lat”). The statistical thresholds were set to *p*-FWE-corr. = 0.05, cluster-level corrected for multiple comparison (cluster size estimated at *p*-unc. = 0.001).

RESULTS

Overt Orienting Behavior

Spatial orienting behavioral data are shown in **Figure 1B**, where the ratio of the time spent in the two hemifields (Gaze_idx) is plotted against the corresponding saliency ratio (Sal_idx), separately for Lat- vs. Multi-trials and Left vs. Right-trials. Overall, the gaze data highlight that spatial orienting behavior in Lat- vs. Multi-trials was very different. The Lat-videos lead

to a systematic shift of gaze towards the corresponding visual hemifield, with Gaze_idx values around ± 0.8 , irrespective of saliency (see plots in green). By contrast, the Multi-videos lead to Gaze_idx values around zeros, indicating that the participants spent approximately the same amount of time fixating in the two hemifields (see plots in red). But notably, now saliency appeared to affect the spatial orienting behavior.

We formally assessed the relationship between gaze and saliency by correlating the two indexes, separately in the four experimental conditions. It should be noted that because these tests considered “left” and “right” trials separately, any “categorical” effect associated with the side of the main visual event on gaze orienting (cf. Lat-trials) did not contribute to the statistics reported here, which instead refer specifically to the effect of saliency on orienting within-hemifield. The results showed that stimulus saliency and the efficacy of the stimuli in orienting overt attention correlated significantly for the Rmulti-trials ($r = 0.44$, $p = 0.014$) and showed a statistical trend for the Lmulti-trials ($r = 0.33$, $p = 0.074$). For both trial-types, the largest level of saliency lateralization, the longer the time participants spent looking towards the most salient visual field (see **Figure 1B**, bottom panels in red). By contrast, when the videos included a single distinctive event, gaze was strongly lateralized on the side of the event and saliency did not affect orienting ($r = 0.10$, $p = 0.538$ for Llat; $r = 0.17$, $p = 0.302$ for Rlat; see top panels in green).

Considering the presence of a rather wide age-range in our sample, we also tested for possible correlations between the saliency-gaze relationship and age, now using subject-by-subject variance. This did not reveal any significant correlation, but a statistical trend was found for the R-lat videos, where age correlated negatively with the relationship between saliency and gaze ($r = -0.47$; $p = 0.069$).

In addition to these analyses regarding the time spent within each hemifield, we sought to further characterize spatial orienting behavior by comparing the overall number of saccades in Multi- vs. Lat-trials and by assessing whether participants fixated salient locations/clusters with a different likelihood in Multi- vs. Lat-conditions. On average subjects made 1.9 saccades per second (± 0.1 , s.e.m.) when the stimulus contained a single lateralized event, while they made 2.4 saccades per second (± 0.1 , s.e.m.) during Multi-trials. A paired t -test formally confirmed the difference between the two stimulus conditions ($t_{(15)} = 6.88$, $p < 0.001$). The percentage of fixations falling inside the saliency clusters was larger in Lat- as compared to Multi-trials (37% vs. 32%, $T_{(138)} = 2.0$; $p < 0.043$), despite the number of these clusters was larger in Multi- as compared to Lat-videos (9.3 vs. 7.3, $T_{(138)} = 4.0$; $p < 0.001$). These additional results suggest that in Lat-trials the participants remained focused onto the most relevant event/object (fewer saccades, and more fixations falling within the saliency clusters), while in Multi-trials they shifted attention sequentially between the multiple distinctive events that characterized these high-competition videos (see also “Discussion” Section).

In sum, the gaze-data highlighted a qualitatively different orienting behavior in Lat- vs. Multi-trials, despite similar levels of saliency lateralization. This supports the notion that

mechanisms other than pure bottom-up saliency contributed to spatial orienting during free viewing of these complex stimuli. Specifically, the participants tended to orient towards events displaying meaningful actions. In Lat-trials the presentation of a single, lateralized distinctive event governed orienting behavior (**Figure 1B**, in green). In Multi-trials there was high competition between multiple co-occurring events and now stimulus-driven saliency was found to contribute to overt spatial orienting (see **Figure 1B**, in red).

Overall Effect of Competition Between Visual Events

Before addressing the main issue concerning the competition between multiple events (i.e., “Multi vs. Lat”), for completeness we compared all four video-conditions vs. rest. This revealed the activation of a large portion of the cortex, including the entire occipital visual cortex, lateral and medial frontal regions, the insulae, medial temporal regions (comprising the hippocampus), the intraparietal sulcus (IPS) and the frontal eye fields (FEF), as well as the TPJ and the IFG. Because of the lack of specificity of this contrast we will not discuss these effects any further, aside briefly pointing out here that the activated areas included regions belonging to both the dorsal (IPS and FEF) and the ventral (TPJ and IFG) frontoparietal attention networks.

We tested for the influence of competition between distinctive events by directly comparing conditions with multiple vs. single-lateralized visual events. The univariate whole-brain analysis tested for “Multi > Lat”, averaging across left and right trials (see below for additional tests considering left and right trials separately). The results showed a significant cluster of activation comprising striate and extrastriate occipital visual cortex, extending into the ventral occipitotemporal cortex (see **Figure 2A**). A second significant cluster of activation was located in the precuneus (see **Table 1**).

The MVPA decoding revealed significant effects in the same regions, but with a larger cluster that now extended dorsally into the superior occipital cortex (SOC), and laterally including the lateral occipital cortex (LOC) and the R TPJ (see **Figure 2B**). Two additional significant clusters were found in the right superior temporal sulcus (R STS) and in the right inferior premotor cortex (R IFG/ middle frontal gyrus, RMFG; see **Figure 2B**, panel on the right). However, it should be noted that—albeit significant—the decoding accuracy for these two regions was just above chance level (see **Table 1**, rightmost column).

Accordingly, the MVPA results confirmed the results of the standard univariate analyses, but also highlighted several additional areas, in particular suggesting competition-related effects in regions belonging to the ventral attention system (i.e., R TPJ and R IFG/MFG). For completeness, we asked whether these differences between univariate and multivariate results merely reflected higher sensitivity of MVPA (Norman et al., 2006), or rather revealed some form of inter-digitated population coding within the ventral attention network (see also Silvetti et al., 2015). For this, we lowered the threshold of the whole-brain univariate analysis to p -unc. < 0.001 , and looked for peaks around the regions observed in the MVPA. This revealed that R TPJ and

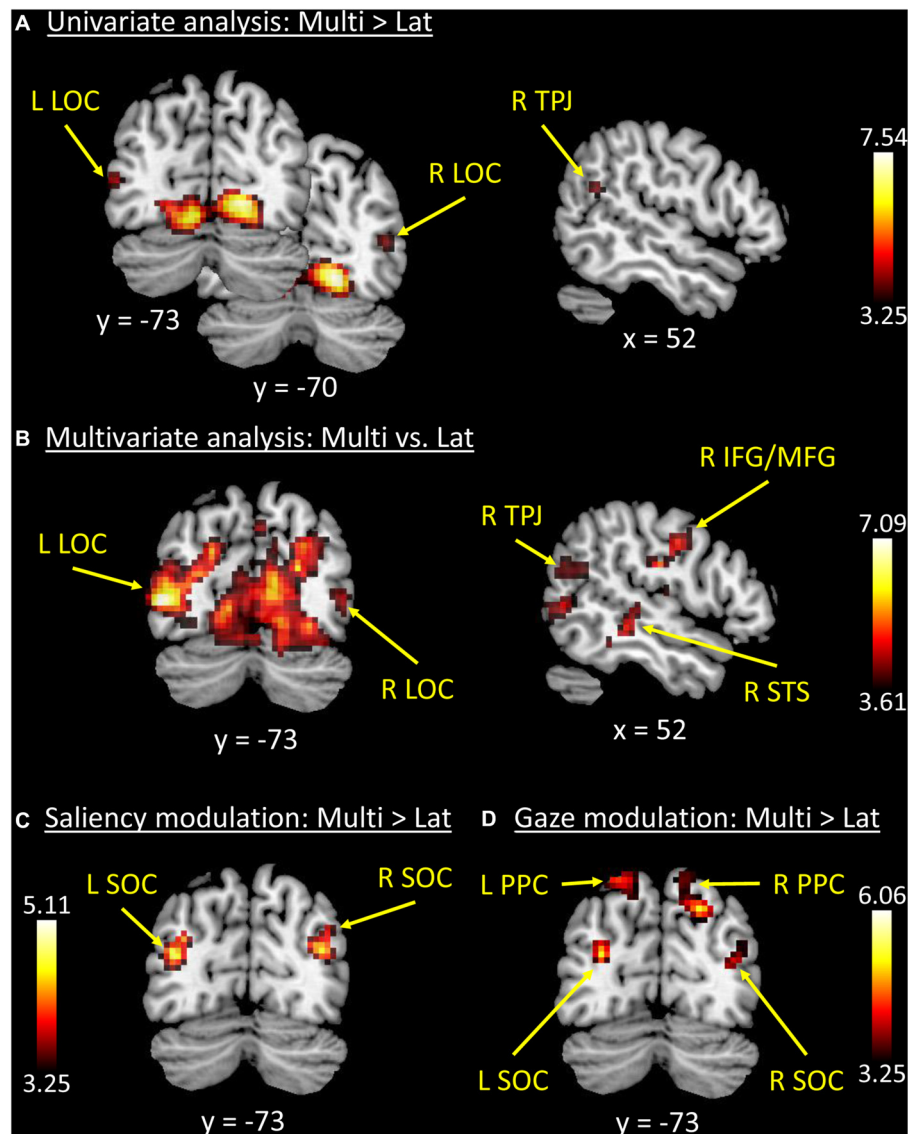


FIGURE 2 | Results of the whole-brain functional MR imaging (fMRI) analyses. (A) Results of the univariate analysis comparing “Multi > Lat-trials”. These showed fully significant effects in occipital visual areas, plus a cluster in the right temporoparietal junction (R TPJ) that did not survive correction for multiple comparisons. **(B)** Results of the multivariate analysis (classification of “Multi vs. Lat” trials) that confirmed the effects in visual cortex, now also showing fully significant effects in the R TPJ and right inferior/middle frontal gyri (R IFG/MFG). **(C,D)** Trial-by-trial modulatory effects associated with the level of stimulus saliency (Sal_idx, panel C) and orienting efficacy (Gaze_idx, panel D). Superior occipital regions plus the posterior parietal cortex (PPC) showed increasing BOLD responses with increasing levels of lateralization of saliency and/or gaze, but more so on Multi-trials compared with Lat-trials (i.e., direct comparison between parametric modulators: “Multi > Lat”). See **Tables 1, 2** for peak-coordinates and detailed statistics. Legend: L, left; R, right; Lat, lateralized trials; Multi, multiple trials; TPJ, temporoparietal junction; LOC, lateral occipital cortex; SOC, superior occipital cortex; IFG/MFG, inferior/middle frontal gyri; STS, superior temporal sulcus; PPC, posterior parietal cortex; x/y, coordinates in MNI space.

bilateral LOC showed an effect for the contrast “Multi > Lat” (see **Table 1**, stats reported in *italics*; and **Figure 2A**). These additional tests indicate higher sensitivity of MVPA compared to standard univariate methods.

The lateralization of the TPJ findings in the right hemisphere prompted the question of whether this would hold for both left and right-trials, which were pooled in all the analyses above. Accordingly, we re-tested the effects of “Multi vs. Lat”, now

considering left- and right-trials separately. For the univariate analysis, the two corresponding simple main effects showed analogous activations of the R TPJ (L-multi > L-lat: $xyz = 63 -49 20$, $T = 3.28$; R-multi > R-lat: $xyz = 57 -40 14$, $T = 3.38$; p -unc. < 0.001), without any effect in the L TPJ. The multivariate test that considered the left-trials only showed again an effect lateralized in the R TPJ (L-multi vs. L-lat: $xyz = 48 -58 20$, $T = 3.72$; p -unc. < 0.001), without any effect in the left

hemisphere. The multivariate test with right-trials only revealed a robust effect in the R TPJ ($xyz = 51 -55 14$, $T = 3.95$, p -unc. < 0.001), but now also a weak effect in the left TPJ ($xyz = -60 -58 17$, $T = 2.68$, p -unc. < 0.01). Overall, these additional tests seem to suggest some dominance of the right TPJ for the processing of the videos including multiple events.

Modulation According to Stimulus Saliency and Orienting Efficacy

The analyses of the behavioral data highlighted a different contribution of stimulus-driven saliency, depending on the level of competition present in the stimulus: saliency was found to influence orienting behavior only when videos contained multiple competing events (see **Figure 1B**). We investigated the possible underlying neural basis of this condition-specific finding with two additional fMRI analyses. First, we tested for regions showing larger co-variation between the saliency lateralization index (Sal_idx) and the BOLD signal in Multi- compared with Lat-trials. This revealed a significant cluster of activation in the left SOC (see **Figure 2C** and **Table 2**). The corresponding region in the right hemisphere showed a statistical trend in the same direction (p -FWE-corr. = 0.079, corrected at whole-brain level). Thus, in these occipital regions the more saliency was lateralized to one hemifield, the larger the BOLD response while viewing the video, but—critically—this effect was significantly larger in videos that included multiple competing events compared with videos with a single lateralized event.

These results suggest that saliency processing in these regions may occur only when stimulus-driven signals contribute to guide spatial orienting, that is, when videos include multiple events. We further explored this condition-specific link between brain activity and spatial orienting by considering the gaze lateralization index (Gaze_idx). It should be noted that in the relevant Multi-trials the Gaze_idx correlated with the Sal_idx (cf. behavioral results above, and **Figure 1B**, plots in red). Thus, testing for any BOLD co-variation with gaze may reveal similar

results as using saliency. Nonetheless, the gaze index should identify more specifically the Multi-trials where gaze/attention was *effectively* captured towards one side, a factor that we have previously shown to be critical for the activation of the dorsal attention system (Nardo et al., 2011). Indeed, the results of the co-variation analysis with Gaze_idx confirmed the modulation of activity in the SOC, but now also revealed a modulation of activity in the posterior parietal cortex (PPC, extending medially into the precuneus; see **Figure 2D** and **Table 2**).

DISCUSSION

The aim of this study was to investigate the role of stimulus-driven saliency and competition between distinctive visual events during free-viewing of naturalistic stimuli, in the absence of any explicit goal-directed task. We expected that high-levels of competition would result in an increased demand of selective processing, which we hypothesized would increase the contribution of stimulus-driven saliency to spatial orienting. Behaviorally, we found that indeed saliency determined within-hemifield gaze orienting only in presence of multiple competing events. Functional imaging results showed that the processing of competing visual events engaged visual areas and the precuneus, but also the R STS and key nodes of the ventral attention network (R TPJ and R IFG/MFG, using multivariate analyses). Stimulus-driven saliency was found to modulate activity in the SOC, selectively when the videos included multiple competing events. Moreover, the activation of the SOC as well as the PPC (in the dorsal attention network) co-varied with an index of spatial orienting efficacy. These results demonstrate that both dorsal and ventral attention systems contribute to the processing of competing events in naturalistic conditions, with a segregation between event detection and spatial selection in ventral and dorsal regions, respectively (cf. also Nardo et al., 2011).

The present study was designed to investigate spatial orienting as it takes place in life-like conditions similar to those we encounter in our everyday experience. Participants were asked to free-view short video-clips of dynamic real-world scenes, without any explicit task. These experimental settings are rather different from those used in most previous studies of visuospatial attention, where instead subjects are asked to actively engage in a specific task, such as searching for a target among many distractors. Thus, we make a distinction between “general attention control” and “goal-directed control”. The former would include all possible stimulus-related (external) influences, as well as many types of endogenous (internal) signals. Internal signals comprise both implicit information (such as that related to the recognition of semantically-relevant events in naturalistic conditions) and information that may be associated with an explicit task (e.g., task-relevant features defining a specific search template). A feature of the present study is that it did not involve any signal of the latter type: that is, there was no requirement of task-related, goal-directed control.

Our videos were characterized by the presence of either a single distinctive visual event lateralized in one hemifield or multiple events on both sides of space. The events included moving objects—in most cases people performing meaningful

TABLE 2 | Results of the analyses testing for activation associated with saliency (Sal_idx) and gaze (Gaze_idx) indexes.

	Region	Cluster		Peak	
		<i>p</i>	<i>k</i>	<i>t</i>	<i>x, y, z</i>
Sal_idx	L SOC	0.001	177	5.11	-39 -76 20
	R SOC	<i>0.079</i>	66	4.47	39 -73 20
Gaze_idx	L SOC	0.042	72	4.05	-24 -82 32
	R SOC	<i>0.132</i>	50	4.74	45 -79 20
	L PPC	<0.001	596	4.57	-24 -70 59
	R PPC			4.48	18 -67 59
	R PCN			4.37	9 -58 65
	L PCN			4.89	-15 -61 62

Cluster: FWE-corrected *p*-value and cluster size (k = number of voxels). Peak: *t*-statistics and *x, y, z*-coordinates in MNI space. Regions: R/L, right/left hemisphere; SOC, superior occipital cortex; PPC, posterior parietal cortex; PCN, precuneus. In italics: the right LOC showed the relevant effects only at uncorrected *p*-values (p -unc < 0.001) and is reported because of the fully significant findings in the corresponding region of the left hemisphere.

actions—that were semantically relevant, yet without being explicitly task/goal-related. This allowed us to ask the question of how the neural substrate traditionally involved in the selection of task-relevant stimuli (i.e., the frontoparietal attention networks, see Corbetta and Shulman, 2002) process such meaningful events, representative of real-world situations. Our approach partly relies on previous evidence that complex naturalistic stimuli can yield consistent patterns of brain activity, even in the absence of any explicit task (Hasson et al., 2004).

The standard univariate contrast that compared directly conditions with high vs. low competition (Multi- vs. Lat-trials) showed significant activation only within the occipital and occipitotemporal cortex, plus the precuneus. However, using a multivariate approach we showed that the level of competition was also represented in the R TPJ and the right inferior premotor cortex, that is, the two main nodes of the ventral attention network. Additional assessment of the imaging data revealed that the lack of significant effects in R TPJ and R IFG in the univariate analysis merely reflected the lower statistical power of the univariate compared to multivariate approach (see **Table 1** showing that both regions activated at $p\text{-unc.} < 0.001$; cf. also Norman et al., 2006; Davis et al., 2014).

The R TPJ has been associated with stimulus-driven attentional reorienting, when a behaviorally relevant stimulus is presented outside the current focus of attention (Corbetta et al., 2008). Other studies have emphasized non-spatial aspects of attention control in R TPJ, including breaches/updates of expectations and responses to low frequency stimuli (Bledowski et al., 2004; Mulert et al., 2004; Geng and Vossel, 2013; for a proposal concerning R TPJ de-activation related to the filtering of spatial and non-spatial distractors, see also Shulman et al., 2009). Albeit enclosing substantial differences, these formulations emphasize the role of the ventral attention system in mediating the interaction between the current internal goals/expectations and the processing of the stimuli in the external environment (see Geng and Vossel, 2013; Macaluso and Doricchi, 2013; cf. also the notion of the R TPJ acting as a “circuit-breaker” in Corbetta and Shulman, 2002). By contrast, the current study did not involve any explicit task (see also above) and each trial included the presentation of a unique stimulus/video, which should minimize any task-based expectation and any need of updating thereof.

In a previous fMRI study that also involved passive viewing of complex and dynamic visual stimuli (i.e., a virtual visual environment), we found that the R TPJ activated on the presentation of unique, task-irrelevant events (moving avatars in that study; Nardo et al., 2011). The meaningful events engaged the R TPJ particularly when participants oriented gaze/attention towards these stimuli. This prompted us to suggest that the R TPJ activation reflected the transient engagement of a detection system, independent of task-relevance. Here, the high competition videos included multiple distinctive events, thus an explanation for the current findings would be that the Multi-trials lead to multiple, sequential detections with greater activation of the R TPJ as compared to the Lat-trials (single detection of just one distinctive event). In line with this, the eye-tracking data showed that participants made more saccades, and corresponding detections and shifts of spatial attention, in Multi-

compared with Lat-trials (see also below, for a discussion of the possible confounding effect due to different number of saccades between conditions).

The current results are consistent with a previous fMRI study comparing bilateral vs. unilateral target-conditions using very simple visual stimuli (i.e., dots; see Beume et al., 2015). This study showed that, as compared to unilaterally-presented dots, bilaterally-presented dots were associated with increased activity in the right ventral attention network, including the R TPJ, the R IFG and perisylvian regions located in the temporal lobe (superior/middle temporal gyri). The authors suggested that the processing of bilateral events requires more attentional resources compared with unilateral processing, and that attentional regions were recruited to convey input from visual association areas to higher-order spatial representations that would integrate spatial signals arising in the two hemifields. Related to our proposal above, it should be noted that in this previous study the stimuli were presented for 400 ms and therefore, in bilateral trials, participants may have shifted attention between the two hemifields. However, the participants were asked to maintain central fixation during the detection task (covert orienting of attention) and no evidence with regard to this possibility was available.

Comparing conditions including bilateral vs. unilateral events is highly relevant for the understanding of spatial orienting deficits typically associated with right-hemisphere damage (unilateral spatial Neglect). Spatial Neglect is a complex neurological syndrome that include an orienting bias towards the ipsilesional (right) side of space, consistently reported in patients with lesions including the R TPJ (Corbetta and Shulman, 2002; Chechlacz et al., 2010; Ptak and Schnider, 2011). The deficit is exacerbated when patients are confronted with multiple stimuli presented across the two hemifields, that is, in conditions entailing high levels of competition (Riddoch and Humphreys, 1983; Bartolomeo and Chokron, 2001; Corben et al., 2001; Geng and Behrmann, 2006; Coulthard et al., 2008). Concurrent presentation of two stimuli in the two hemifields can lead to visual extinction of the stimulus presented on the contralesional side, again a deficit that has been associated with lesions of the R TPJ and neighboring regions in the posterior/superior temporal cortex (Friedrich et al., 1998; Karnath et al., 2003; Ticini et al., 2010; de Haan et al., 2012; Chechlacz et al., 2013; for related work using TMS in healthy subjects, see also Meister et al., 2006).

Our current results confirm the involvement of the right ventral attention network in the processing of competing events, with a lateralization to the right hemisphere that appeared to be largely independent of the stimulus-lateralization (cf. additional analyses that considered separately left and right trials). The eye-tracking data suggested that with the current stimulus material, which involve the presentation of meaningful events for a relatively long duration—as would happen in any everyday life situation—the high levels of competition in Multi-trials lead to the sequential detection of the distinctive events and shifting of spatial attention. Adding to previous findings using simple stimuli and explicit goal-directed tasks, here we demonstrate that the ventral attention system detects distinctive visual events within complex and dynamic visual scenes, even in the absence

of any explicit goal-directed task. Moreover, unlike the few previous studies that showed activation of the R TPJ/R IFG for fully irrelevant stimuli (e.g., Asplund et al., 2010; Nardo et al., 2011), in the current study these effects can hardly be explained by low stimulus probability or some general effect of surprise. The finding that the R TPJ and R IFG engaged in the absence of any explicit goal-directed task poses a challenge for models of attention control that emphasize the role of task-set and/or expectations in the ventral attention system (Downar et al., 2001; Kincade et al., 2005; Geng and Mangun, 2011).

The main hypothesis of the current study was that stimulus-driven saliency would contribute to spatial orienting primarily in conditions entailing high levels of competition, when the presence of multiple visual events implies uncertain/conflicting information about the most relevant location in space. We predicted that in this situation the additional information provided by sensory salience would be most effective for guiding spatial orienting. Indeed, the behavioral results showed a correlation between saliency and gaze data only in Multi-trials (see **Figure 1B**, plots in red). By contrast, in the Lat-trials the subjects oriented their gaze systematically towards the distinctive events and there was no effect of salience (despite the good amount of *Sal_idx* variability also in these trials; cf. x-axis in the top-panels of **Figure 1B**). These behavioral data indicate that in Lat-trials the distinctive events governed spatial orienting over and above any bottom-up influence, whereas the latter did contribute to spatial orienting in presence of multiple, competing events.

The analyses of imaging data sought to identify where in the brain this condition-specific influence of saliency on spatial orienting was implemented. The results showed that both stimulus salience (*Sal_idx*) and stimulus efficacy (*Gaze_idx*) modulated activity in the SOC, selectively in high competition Multi-trials (see **Figures 2C,D**). In addition, the index related to the stimulus orienting efficacy was found to modulate responses of the PPC, in the dorsal attention network (Corbetta and Shulman, 2002).

These results are in agreement with previous findings that demonstrated coding of salience in PPC and—more specifically—that response of the dorsal attention network reflects the efficacy of these stimulus-driven signals for spatial orienting and selection (Bogler et al., 2011; Nardo et al., 2011, 2014; Santangelo and Macaluso, 2013). Here, the engagement of these dorsal regions provide us with a possible physiological substrate for the behavioral finding that salience affected orienting behavior selectively in high competition trials. Beyond the classical view that the dorsal attention system is associated with goal-directed attentional and oculomotor control (Corbetta and Shulman, 2002; Müri, 2006; Corbetta et al., 2008), recent evidence emphasizes the integration of multiple control signals for spatial selection in the PPC (see also Macaluso and Doricchi, 2013). The PPC is sensitive to bottom-up attentional influences associated with stimulus salience (Geng and Mangun, 2009), contains attention maps indexing space to support the selection of multiple objects at the same time (Somers and Sheremata, 2013), and is implicated in feature binding, but only when spatial information is available to resolve ambiguities about

the relationships between object features (Shafritz et al., 2002). These findings are consistent with the notion that PPC contains topographical representations of the visual space that code for the relative relevance of different locations in the environment (“saliency maps”: Gottlieb et al., 1998; Gottlieb, 2007; Bogler et al., 2011; Nardo et al., 2011), and that integrate bottom-up and top-down signals for the control of spatial orienting (see also, “priority maps”: Bisley and Goldberg, 2010; Arcizet et al., 2011; Ptak, 2012).

Our results implicate both dorsal and ventral attention systems during spatial orienting in naturalistic viewing conditions that included multiple competing events. The notion that spatial orienting relies on the interplay between the two frontoparietal control systems is well acknowledged in the literature (e.g., Corbetta and Shulman, 2002; Macaluso, 2010; Vossel et al., 2014). This interplay has been studied primarily in tasks entailing explicit goal-directed attention (such as spatial cueing tasks), leading to the proposal that the ventral system acts as a “circuit breaker” that re-sets top-down control that the dorsal system exerts on visual areas (Corbetta and Shulman, 2002). Subsequent investigations further specified the constraints that can lead to the engagement of the ventral network, including stimulus saliency/novelty (Downar et al., 2000, 2001, 2002; Bledowski et al., 2004; Mavritsaki et al., 2010), task-relatedness (Kincade et al., 2005; Corbetta et al., 2008; Geng and Mangun, 2011) and contextual updating (Geng and Vossel, 2013). Our current data do not provide us with any information about the dynamics of the interaction between the two attention networks (Vossel et al., 2012), but indicate that both systems engage even in the absence of any explicit goal-directed task. We propose that here the ventral system detects the occurrence of multiple competing events, signaling to the dorsal (posterior parietal) system that salience needs to be taken into account in order to resolve the competition and select the spatial location with the highest priority (cf. McMains and Kastner, 2011; for related proposal concerning selective processing in the visual cortex). Future studies may use TMS and/or analyses of effective connectivity to shed light on the temporal sequence of the activation of ventral and dorsal frontoparietal regions during free-viewing of naturalistic stimuli.

While the use of complex stimulus material and free-viewing provide us with experimental settings that approximate attention control in the real world, it should be acknowledged that this approach entails also several limitations. First, here the definition of what constitutes a distinctive event, and thus whether a specific video was classified as a Multi- or Lat-trial, had to be relatively arbitrary. We sought to select a pool of videos where there was a clear difference between the distinctive foreground event/s and the scene background. On average, the eye-tracking data confirmed that there was indeed a measurable, objective difference between the videos assigned to the different conditions (see **Figure 1B**). This was further supported by the different number of saliency clusters in Multi- vs. Lat-videos, as well as by the different percentage of fixations falling inside the saliency clusters in the two conditions (see “Results” Section). Taken together, these data indicate that indeed there was a systematic difference between the videos selected for the two conditions.

Nonetheless, it is possible that for some specific video the condition-assignment (or inclusion in the final pool of videos) could have been done in a different way.

Second, the different conditions lead to different overt motor behavior (e.g., number of saccades), which is a limitation for the interpretation of the imaging results. Indeed, the large activation of visual areas in both univariate and multivariate analyses (including the striate cortex) is most likely due to differences in the visual input arriving to the occipital visual cortex, as a consequence of the different oculomotor behavior. Nonetheless, in previous studies we have explicitly addressed the impact of overt eye-movements on the pattern of activity associated with event-detection and salience-processing by scanning participants also during covert viewing of the stimuli (Nardo et al., 2011, 2014). By directly comparing overt vs. covert viewing conditions we found that while activity in early visual areas was indeed strongly dependent on eye-movements, the pattern of activation in R TPJ, R IFG and PPC, as well as extra-striate visual cortex, were largely unaffected by overt/covert viewing condition. Thus, we expect that also the current findings about the effect of competition between events and the interaction between this and salience processing does not merely reflect a motor-related confound. This view is also supported by the current finding that in PPC there was a positive co-variation between BOLD response and Gaze_idx: that is, activity in PPC increased for those Multi-videos that led to some gaze-lateralization. High gaze-lateralization means that the participants spent longer time looking towards a specific hemifield. In general, this is associated with a decrease (rather than an increase) of the number of saccades. For instance, the comparison between Lat- vs. Multi-trials, which are associated with high vs. low Gaze_idx respectively (cf. **Figure 1B**), showed that the number of saccades was indeed smaller in the former (1.9 vs. 2.4 saccades/s, see “Results” Section). Accordingly, it seems unlikely that the positive co-variation between the BOLD signal in PPC and Gaze_idx can be attributed to some increase in eye movements.

Third, besides the differences in visual complexity and number of saccades, Lat- and Multi-trials may differ in the number of “interesting” or “relevant” objects. While this issue could in principle be addressed using manual object-labeling (e.g., see “labelme” project; Russell et al., 2008), it should be pointed out that the stimuli presented in the current study included a total of 5250 frames, which makes any such manual labeling procedure unfeasible. Instead, here we opted for a careful selection of the videos and the computational analysis of the stimuli (saliency maps), which appeared suitable to achieve the

current aim of studying the interplay between event-competition and bottom-up saliency. Moreover, it should be noted that the computational approach has the advantage that it can be easily applied to any new set of stimuli, thus facilitating future research based on the current results.

In conclusion, this study was aimed at investigating how competition between distinctive visual events and stimulus-driven salience affect spatial orienting, when viewing naturalistic stimuli without any goal-directed task. We found that salience contributed to spatial orienting only when the level of competition was high (Multi-trials). While in situations of low competition (Lat-trials) spatial orienting was driven by the main visual event, the presence of multiple competing events implicated an additional role of stimulus-driven signals for the selection of the most relevant spatial locations. The imaging analyses highlighted the engagement of the right ventral attention system (R TPJ and R IFG) for the processing of highly competing stimuli. We relate the increased activation of the ventral system with the detection of multiple events in the high competition trials and sequential shifting of spatial attention between these events. Selectively for the high competition trials, we found that salience modulated activity in the SOC and that the efficacy of the stimuli for spatial orienting modulated the same occipital areas, as well as the PPC. We link these effects in dorsal regions with the representation of bottom-up signals that effectively contribute to spatial orienting (Gottlieb, 2007; Bogler et al., 2011; Nardo et al., 2011). We conclude that dorsal and ventral frontoparietal attention networks play specific roles during the processing of competing events and spatial orienting in life-like conditions. Our results contribute to the emerging field of neuroimaging studies that try to characterize mechanisms of attention control relevant for brain functioning in the real world (Peelen and Kastner, 2014; see also Hasson et al., 2004).

AUTHOR CONTRIBUTIONS

DN, CR and EM designed the study. DN created the stimuli, ran the experiments and analyzed behavioral data. DN, CR and PC analyzed fMRI data. DN and EM wrote the manuscript. DN, PC, CR and EM edited the manuscript.

ACKNOWLEDGMENTS

This research was supported by the Italian Ministry of Health, grant N. GR-2010-2311171 to EM.

REFERENCES

- Arcizet, F., Mirpour, K., and Bisley, J. W. (2011). A pure salience response in posterior parietal cortex. *Cereb. Cortex* 21, 2498–2506. doi: 10.1093/cercor/bhr035
- Arrington, C. M., Carr, T. H., Mayer, A. R., and Rao, S. M. (2000). Neural mechanisms of visual attention: object-based selection of a region in space. *J. Cogn. Neurosci.* 12, 106–117. doi: 10.1162/089892900563975
- Asplund, C. L., Todd, J. J., Snyder, A. P., and Marois, R. (2010). A central role for the lateral prefrontal cortex in goal-directed and stimulus-driven attention. *Nat. Neurosci.* 13, 507–512. doi: 10.1038/nn.2509
- Bacon, W. J., and Egeth, H. E. (1997). Goal-directed guidance of attention: evidence from conjunctive visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 948–961. doi: 10.1037/0096-1523.23.4.948
- Bartolomeo, P., and Chokron, S. (2001). Unilateral neglect: the effect of competing stimuli on estimated line length. *Brain Cogn.* 46, 34–38. doi: 10.1016/s0278-2626(01)80028-2
- Beume, L. A., Kaller, C. P., Hoeren, M., Klöppel, S., Kuemmerer, D., Glauche, V., et al. (2015). Processing of bilateral versus unilateral conditions: evidence for the functional contribution of the ventral attention network. *Cortex* 66, 91–102. doi: 10.1016/j.cortex.2015.02.018

- Bisley, J. W., and Goldberg, M. E. (2010). Attention, intention and priority in the parietal lobe. *Annu. Rev. Neurosci.* 33, 1–21. doi: 10.1146/annurev-neuro-060909-152823
- Bledowski, C., Prvulovic, D., Goebel, R., Zanella, F. E., and Linden, D. E. (2004). Attentional systems in target and distractor processing: a combined ERP and fMRI study. *Neuroimage* 22, 530–540. doi: 10.1016/j.neuroimage.2003.12.034
- Bogler, C., Bode, S., and Haynes, J. D. (2011). Decoding successive computational stages of saliency processing. *Curr. Biol.* 21, 1667–1671. doi: 10.1016/j.cub.2011.08.039
- Carmi, R., and Itti, L. (2006). Visual causes versus correlates of attentional selection in dynamic scenes. *Vision Res.* 46, 4333–4345. doi: 10.1016/j.visres.2006.08.019
- Castelhano, M. S., and Heaven, C. (2010). The relative contribution of scene context and target features to visual search in scenes. *Atten. Percept. Psychophys.* 72, 1283–1297. doi: 10.3758/APP.72.5.1283
- Castelhano, M. S., and Heaven, C. (2011). Scene context influences without scene gist: eye movements guided by spatial associations in visual search. *Psychon. Bull. Rev.* 18, 890–896. doi: 10.3758/s13423-011-0107-8
- Cave, K. R., and Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search. *Cogn. Psychol.* 22, 225–271. doi: 10.1016/0010-0285(90)90017-x
- Chechlacz, M., Rotshtein, P., Bickerton, W. L., Hansen, P. C., Deb, S., and Humphreys, G. W. (2010). Separating neural correlates of allocentric and egocentric neglect: distinct cortical sites and common white matter disconnections. *Cogn. Neuropsychol.* 27, 277–303. doi: 10.1080/02643294.2010.519699
- Chechlacz, M., Rotshtein, P., Hansen, P. C., Deb, S., Ridloch, M. J., and Humphreys, G. W. (2013). The central role of the temporoparietal junction and the superior longitudinal fasciculus in supporting multi-item competition: evidence from lesion-symptom mapping of extinction. *Cortex* 49, 487–506. doi: 10.1016/j.cortex.2011.11.008
- Corben, L. A., Mattingley, J. B., and Bradshaw, J. L. (2001). A kinematic analysis of distractor interference effects during visually guided action in spatial neglect. *J. Int. Neuropsychol. Soc.* 7, 334–343. doi: 10.1017/s1355617701733073
- Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., and Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nat. Neurosci.* 3, 292–297. doi: 10.1038/73009
- Corbetta, M., Patel, G., and Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58, 306–324. doi: 10.1016/j.neuron.2008.04.017
- Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi: 10.1038/nrn755
- Coulthard, E. J., Nachev, P., and Husain, M. (2008). Control over conflict during movement preparation: role of posterior parietal cortex. *Neuron* 58, 144–157. doi: 10.1016/j.neuron.2008.02.009
- Davis, T., LaRocque, K. F., Mumford, J. A., Norman, K. A., Wagner, A. D., and Poldrack, R. A. (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. doi: 10.1016/j.neuroimage.2014.04.037
- de Haan, B., Karnath, H. O., and Driver, J. (2012). Mechanisms and anatomy of unilateral extinction after brain injury. *Neuropsychologia* 50, 1045–1053. doi: 10.1016/j.neuropsychologia.2012.02.015
- Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222. doi: 10.1146/annurev-neuro.18.1.193
- Downar, J., Crawley, A. P., Mikulis, D. J., and Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nat. Neurosci.* 3, 277–283. doi: 10.1038/72991
- Downar, J., Crawley, A. P., Mikulis, D. J., and Davis, K. D. (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. doi: 10.1006/nimg.2001.0946
- Downar, J., Crawley, A. P., Mikulis, D. J., and Davis, K. D. (2002). A cortical network sensitive to stimulus salience in a neutral behavioral context across multiple sensory modalities. *J. Neurophysiol.* 87, 615–620. doi: 10.1152/jn.00636.2001
- Draschkow, D., Wolfe, J. M., and Vö, M. L. (2014). Seek and you shall remember: scene semantics interact with visual search to build better memories. *J. Vis.* 14:10. doi: 10.1167/14.8.10
- Duncan, J., and Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychol. Rev.* 96, 433–458. doi: 10.1037/0033-295x.96.3.433
- Duncan, J., and Humphreys, G. (1992). Beyond the search surface: visual search and attentional engagement. *J. Exp. Psychol. Hum. Percept. Perform.* 18, 578–588; discussion 589–593. doi: 10.1037/0096-1523.18.2.578
- Einhäuser, W., Mundhenk, T. N., Baldi, P., Koch, C., and Itti, L. (2007). A bottom-up model of spatial attention predicts human error patterns in rapid scene recognition. *J. Vis.* 7, 6.1–6.13. doi: 10.1167/7.10.6
- Einhäuser, W., Spain, M., and Perona, P. (2008). Objects predict fixations better than early saliency. *J. Vis.* 8, 18.1–18.26. doi: 10.1167/8.14.18
- Elazary, L., and Itti, L. (2008). Interesting objects are visually salient. *J. Vis.* 8, 3.1–3.15. doi: 10.1167/8.3.3
- Epstein, R. A., and Morgan, L. K. (2012). Neural responses to visual scenes reveals inconsistencies between fMRI adaptation and multivoxel pattern analysis. *Neuropsychologia* 50, 530–543. doi: 10.1016/j.neuropsychologia.2011.09.042
- Evans, K. K., Horowitz, T. S., and Wolfe, J. M. (2011). When categories collide: accumulation of information about multiple categories in rapid scene perception. *Psychol. Sci.* 22, 739–746. doi: 10.1177/0956797611407930
- Foley, N. C., Grossberg, S., and Mingolla, E. (2012). Neural dynamics of object-based multifocal visual spatial attention and priming: object cueing, useful-field-of-view and crowding. *Cogn. Psychol.* 65, 77–117. doi: 10.1016/j.cogpsych.2012.02.001
- Friedrich, F. J., Egly, R., Rafal, R. D., and Beck, D. (1998). Spatial attention deficits in humans: a comparison of superior parietal and temporal-parietal junction lesions. *Neuropsychology* 12, 193–207. doi: 10.1037/0894-4105.12.2.193
- Friston, K. J., Glaser, D. E., Henson, R. N., Kiebel, S., Phillips, C., and Ashburner, J. (2002). Classical and bayesian inference in neuroimaging: applications. *Neuroimage* 16, 484–512. doi: 10.1006/nimg.2002.1091
- Geng, J. J., and Behrmann, M. (2006). Competition between simultaneous stimuli modulated by location probability in hemispatial neglect. *Neuropsychologia* 44, 1050–1060. doi: 10.1016/j.neuropsychologia.2005.10.026
- Geng, J. J., and Mangun, G. R. (2009). Anterior intraparietal sulcus is sensitive to bottom-up attention driven by stimulus salience. *J. Cogn. Neurosci.* 21, 1584–1601. doi: 10.1162/jocn.2009.21103
- Geng, J. J., and Mangun, G. R. (2011). Right temporoparietal junction activation by a salient contextual cue facilitates target discrimination. *Neuroimage* 54, 594–601. doi: 10.1016/j.neuroimage.2010.08.025
- Geng, J. J., and Vessel, S. (2013). Re-evaluating the role of TPJ in attentional control: contextual updating? *Neurosci. Biobehav. Rev.* 37, 2608–2620. doi: 10.1016/j.neubiorev.2013.08.010
- Gottlieb, J. (2007). From thought to action: the parietal cortex as a bridge between perception, action and cognition. *Neuron* 53, 9–16. doi: 10.1016/j.neuron.2006.12.009
- Gottlieb, J. P., Kusunoki, M., and Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature* 391, 481–484. doi: 10.1038/35135
- Greene, M. R., and Wolfe, J. M. (2011). Global image properties do not guide visual search. *J. Vis.* 11:18. doi: 10.1167/11.6.18
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., and Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science* 303, 1634–1640. doi: 10.1126/science.1089506
- Haynes, J.-D. (2015). A primer on pattern-based approaches to fMRI: principles, pitfalls and perspectives. *Neuron* 87, 257–270. doi: 10.1016/j.neuron.2015.05.025
- Humphrey, K., and Underwood, G. (2009). Domain knowledge moderates the influence of visual saliency in scene recognition. *Br. J. Psychol.* 100, 377–398. doi: 10.1348/000712608X344780
- Itti, L., and Koch, C. (2001). Computational modelling of visual attention. *Nat. Rev. Neurosci.* 2, 194–203. doi: 10.1038/35058500
- Itti, L., Koch, C., and Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* 20, 1254–1259. doi: 10.1109/34.730558
- Johnson, M. R., and Johnson, M. K. (2014). Decoding individual natural scene representations during perception and imagery. *Front. Hum. Neurosci.* 8:59. doi: 10.3389/fnhum.2014.00059
- Karnath, H. O., Himmelbach, M., and Küker, W. (2003). The cortical substrate of visual extinction. *Neuroreport* 14, 437–442. doi: 10.1097/00001756-200303030-00028

- Kastner, S., and Pinsk, M. A. (2004). Visual attention as a multilevel selection process. *Cogn. Affect. Behav. Neurosci.* 4, 483–500. doi: 10.3758/cabn.4.4.483
- Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., and Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22, 751–761. doi: 10.1016/s0896-6273(00)80734-5
- Kincade, J. M., Abrams, R. A., Astafiev, S. V., Shulman, G. L., and Corbetta, M. (2005). An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. *J. Neurosci.* 25, 4593–4604. doi: 10.1523/JNEUROSCI.0236-05.2005
- Linsley, D., and MacEvoy, S. P. (2014). Evidence for participation by object-selective visual cortex in scene category judgments. *J. Vis.* 14:19. doi: 10.1167/14.9.19
- Macaluso, E. (2010). Orienting of spatial attention and the interplay between the senses. *Cortex* 46, 282–297. doi: 10.1016/j.cortex.2009.05.010
- Macaluso, E., and Doricchi, F. (2013). Attention and predictions: control of spatial attention beyond the endogenous-exogenous dichotomy. *Front. Hum. Neurosci.* 7:685. doi: 10.3389/fnhum.2013.00685
- Mahmoudi, A., Takerkart, S., Regragui, F., Boussaoud, D., and Brovelli, A. (2012). Multivoxel pattern analysis for fMRI data: a review. *Comput. Math. Methods Med.* 2012:961257. doi: 10.1155/2012/961257
- Malcolm, G. L., and Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real-world scene search. *J. Vis.* 10, 4.1–4.11. doi: 10.1167/10.2.4
- Malcolm, G. L., and Shomstein, S. (2015). Object-based attention in real-world scenes. *J. Exp. Psychol. Gen.* 144, 257–263. doi: 10.1037/xge0000060
- Mavritsaki, E., Allen, H. A., and Humphreys, G. W. (2010). Decomposing the neural mechanisms of visual search through model-based analysis of fMRI: top-down excitation, active ignoring and the use of saliency by the right TPJ. *Neuroimage* 52, 934–946. doi: 10.1016/j.neuroimage.2010.03.044
- McMains, S., and Kastner, S. (2011). Interactions of top-down and bottom-up mechanisms in human visual cortex. *J. Neurosci.* 31, 587–597. doi: 10.1523/JNEUROSCI.3766-10.2011
- Meister, I. G., Wienemann, M., Buelte, D., Grünwald, C., Sparing, R., Dambeck, N., et al. (2006). Hemiextinction induced by transcranial magnetic stimulation over the right temporoparietal junction. *Neuroscience* 142, 119–123. doi: 10.1016/j.neuroscience.2006.06.023
- Mulert, C., Jäger, L., Schmitt, R., Bussfeld, P., Pogarell, O., Möller, H. J., et al. (2004). Integration of fMRI and simultaneous EEG: towards a comprehensive understanding of localization and time-course of brain activity in target detection. *Neuroimage* 22, 83–94. doi: 10.1016/j.neuroimage.2003.10.051
- Müri, R. M. (2006). MRI and fMRI analysis of oculomotor function. *Prog. Brain Res.* 151, 503–526. doi: 10.1016/s0079-6123(05)51016-1
- Nardo, D., Santangelo, V., and Macaluso, E. (2011). Stimulus-driven orienting of visuospatial attention in complex dynamic environments. *Neuron* 69, 1015–1028. doi: 10.1016/j.neuron.2011.02.020
- Nardo, D., Santangelo, V., and Macaluso, E. (2014). Spatial orienting in complex audiovisual environments. *Hum. Brain Mapp.* 35, 1597–1614. doi: 10.1002/hbm.22276
- Neider, M. B., and Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision Res.* 46, 614–621. doi: 10.1016/j.visres.2005.08.025
- Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* 10, 424–430. doi: 10.1016/j.tics.2006.07.005
- Nuthmann, A., and Henderson, J. M. (2010). Object-based attentional selection in scene viewing. *J. Vis.* 10:20. doi: 10.1167/10.8.20
- Ogawa, A., and Macaluso, E. (2015). Orienting of visuospatial attention in complex 3D space: search and detection. *Hum. Brain Mapp.* 36, 2231–2247. doi: 10.1002/hbm.22767
- Peelen, M. V., Fei-Fei, L., and Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature* 460, 94–97. doi: 10.1038/nature08103
- Peelen, M. V., and Kastner, S. (2014). Attention in the real world: toward understanding its neural basis. *Trends Cogn. Sci.* 18, 242–250. doi: 10.1016/j.tics.2014.02.004
- Pinsk, M. A., Doniger, G. M., and Kastner, S. (2004). Push-pull mechanism of selective attention in human extrastriate cortex. *J. Neurophysiol.* 92, 622–629. doi: 10.1152/jn.00974.2003
- Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25. doi: 10.1080/00335558008248231
- Preston, T. J., Guo, F., Das, K., Giesbrecht, B., and Eckstein, M. P. (2013). The attention network of Neural representations of contextual guidance in visual search of real-world scenes. *J. Neurosci.* 33, 7846–7855. doi: 10.1523/jneurosci.5840-12.2013
- Ptak, R. (2012). The frontoparietal attention network of the human brain: action, saliency and a priority map of the environment. *Neuroscientist* 18, 502–515. doi: 10.1177/1073858411409051
- Ptak, R., and Schneider, A. (2011). The attention network of the human brain: relating structural damage associated with spatial neglect to functional imaging correlates of spatial attention. *Neuropsychologia* 49, 3063–3070. doi: 10.1016/j.neuropsychologia.2011.07.008
- Riddoch, M. J., and Humphreys, G. W. (1983). The effect of cueing on unilateral neglect. *Neuropsychologia* 21, 589–599. doi: 10.1016/0028-3932(83)90056-8
- Russell, B. C., Torralba, A., Murphy, K. P., and Freeman, W. T. (2008). LabelMe: a database and web-based tool for image annotation. *Int. J. Comput. Vis.* 77, 157–173. doi: 10.1007/s11263-007-0090-8
- Santangelo, V., Di Francesco, S. A., Mastroberardino, S., and Macaluso, E. (2015). Parietal cortex integrates contextual and saliency signals during the encoding of natural scenes in working memory. *Hum. Brain Mapp.* 36, 5003–5017. doi: 10.1002/hbm.22984
- Santangelo, V., and Macaluso, E. (2013). Visual salience improves spatial working memory via enhanced parieto-temporal functional connectivity. *J. Neurosci.* 33, 4110–4117. doi: 10.1523/JNEUROSCI.4138-12.2013
- Seidl-Rathkopf, K. N., Turk-Browne, N. B., and Kastner, S. (2015). Automatic guidance of attention during real-world visual search. *Atten. Percept. Psychophys.* 77, 1881–1895. doi: 10.3758/s13414-015-0903-8
- Serences, J. T., and Saproo, S. (2012). Computational advances towards linking BOLD and behavior. *Neuropsychologia* 50, 435–446. doi: 10.1016/j.neuropsychologia.2011.07.013
- Serences, J. T., Shomstein, S., Leber, A. B., Goyal, X., Egeth, H. E., and Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychol. Sci.* 16, 114–122. doi: 10.1111/j.0956-7976.2005.00791.x
- Shafritz, K. M., Gore, J. C., and Marois, R. (2002). The role of the parietal cortex in visual feature binding. *Proc. Natl. Acad. Sci. U S A* 99, 10917–10922. doi: 10.1073/pnas.152694799
- Shulman, G. L., Astafiev, S. V., Franke, D., Pope, D. L., Snyder, A. Z., McAvoy, M. P., et al. (2009). Interaction of stimulus-driven reorienting and expectation in ventral and dorsal frontoparietal and basal ganglia-cortical networks. *J. Neurosci.* 29, 4392–4407. doi: 10.1523/JNEUROSCI.5609-08.2009
- Silvetti, M., Lasaponara, S., Lecce, F., Dragone, A., Macaluso, E., and Doricchi, F. (2015). The response of the left ventral attentional system to invalid targets and its implication for the spatial neglect syndrome: a multivariate fMRI investigation. *Cereb. Cortex* doi: 10.1093/cercor/bhv208 [Epub ahead of print].
- Somers, D. C., and Sheremata, S. L. (2013). Attention maps in the brain. *Wiley Interdiscip. Rev. Cogn. Sci.* 4, 327–340. doi: 10.1002/wcs.1230
- Spotorno, S., Malcolm, G. L., and Tatler, B. W. (2014). How context information and target information guide the eyes from the first epoch of search in real-world scenes. *J. Vis.* 14:7. doi: 10.1167/14.2.7
- Stoll, J., Thrun, M., Nuthmann, A., and Einhäuser, W. (2015). Overt attention in natural scenes: objects dominate features. *Vision Res.* 107, 36–48. doi: 10.1016/j.visres.2014.11.006
- Summerfield, J. J., Lepsiens, J., Gitelman, D. R., Mesulam, M. M., and Nobre, A. C. (2006). Orienting attention based on long-term memory experience. *Neuron* 49, 905–916. doi: 10.1016/j.neuron.2006.01.021
- Theeuwes, J. (1992). Perceptual selectivity for color and form. *Percept. Psychophys.* 51, 599–606. doi: 10.3758/bf03211656
- Ticini, L. F., de Haan, B., Klose, U., Nägele, T., and Karnath, H. O. (2010). The role of temporoparietal cortex in subcortical visual extinction. *J. Cogn. Neurosci.* 22, 2141–2150. doi: 10.1162/jocn.2009.21315
- Treisman, A. M. (1986). Features and objects in visual processing. *Sci. Am.* 255, 114–125. doi: 10.1038/scientificamerican1186-114b
- Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136. doi: 10.1016/0010-0285(80)90005-5
- Vossel, S., Geng, J. J., and Fink, G. R. (2014). Dorsal and ventral attention systems: distinct neural circuits but collaborative roles. *Neuroscientist* 20, 150–159. doi: 10.1177/1073858413494269

- Vossel, S., Weidner, R., Driver, J., Friston, K. J., and Fink, G. R. (2012). Deconstructing the architecture of dorsal and ventral attention systems with dynamic causal modeling. *J. Neurosci.* 32, 10637–10648. doi: 10.1523/JNEUROSCI.0414-12.2012
- Watson, D. M., Hartley, T., and Andrews, T. J. (2014). Patterns of response to visual scenes are linked to the low-level properties of the image. *Neuroimage* 99, 402–410. doi: 10.1016/j.neuroimage.2014.05.045
- Wolfe, J. M., Cave, K. R., and Franzel, S. L. (1989). Guided search: an alternative to the feature integration model for visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 15, 419–433. doi: 10.1037/0096-1523.15.3.419
- Wolfe, J. M., Võ, M. L., Evans, K. K., and Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends Cogn. Sci.* 15, 77–84. doi: 10.1016/j.tics.2010.12.001
- Wu, C. C., Wang, H. C., and Pomplun, M. (2014). The roles of scene gist and spatial dependency among objects in the semantic guidance of attention in real-world scenes. *Vision Res.* 105, 10–20. doi: 10.1016/j.visres.2014.08.019
- Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. *J. Exp. Psychol. Hum. Percept. Perform.* 19, 676–681. doi: 10.1037/0096-1523.19.3.676
- Yantis, S. (2000). “Goal-directed and stimulus-driven determinants of attentional control,” in *Attention and Performance XVIII*, eds S. Monsell and J. Driver (Cambridge, MA: MIT Press), 73–103.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Nardo, Console, Reverberi and Macaluso. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.