Behavioral/Cognitive

Visual Salience Improves Spatial Working Memory via Enhanced Parieto-Temporal Functional Connectivity

Valerio Santangelo[1,2] **and Emiliano Macaluso**[2]

[1]Department of Human and Educational Sciences, University of Perugia, 06123 Perugia, Italy and [2]Neuroimaging Laboratory, Santa Lucia Foundation, 00179 Rome, Italy

In everyday life, the brain is bombarded with a multitude of concurrent and competing stimuli. Only some of these enter consciousness and memory. Attention selects relevant signals for in-depth processing depending on current goals, but also on the intrinsic properties of stimuli. We combined behavior, computational modeling, and functional imaging to investigate mechanisms supporting access to memory based on intrinsic sensory properties. During fMRI scanning, human subjects were presented with pictures of naturalistic scenes that entailed high levels of competition between possible target objects. Following a retention interval of 8 s, participants judged the location (same/different) of a target object extracted from the initial scene. We found that memory performance at retrieval increased with increasing object salience at encoding, indicating a "prior entry" for salient information. fMRI analyses revealed encoding-related activation in the posterior parietal cortex, selectively for salient objects that were later remembered. Moreover, parietal cortex showed increased functional coupling with the medial-temporal lobe, for remembered objects only. These findings reveal a parietotemporal circuit that integrates available sensory cues (based on attention-grabbing saliency signals) and current memory requirements (storing objects' locations) to encode object-related spatial information in working memory.

Introduction

Working memory (WM) plays a key role in maintaining information that is no longer present in the environment but is still relevant for the immediate future. WM can store only a limited amount of information (Luck and Vogel, 1997), implying a link between encoding in WM and attention selection processes (Chun and Turk-Browne, 2007). Previous behavioral studies showed that attended items have a greater chance to be remembered compared to nonattended items (Schmidt et al., 2002; Botta et al., 2010), and that explicit memory is impaired when attentional resources are divided between multiple tasks during encoding (Craik et al., 1996). The link between attention and WM is also supported by neuroimaging studies that have revealed common frontoparietal activation for attention and WM (Naghavi and Nyberg, 2005). A few studies have directly assessed the interplay between attention and WM by manipulating both factors within the same paradigm (Gazzaley et al., 2005; Gazzaley and Nobre, 2012). These demonstrated that goal-directed/voluntary attention can enhance processing of the to-be-remembered stimuli during WM encoding (Murray et al., 2011).

These previous studies have not yet considered two issues that characterize the operation of attention and WM in the real world.

First, these studies have assessed only the role of goal-directed attention, without considering the possible contribution of stimulus-driven attention. Extensive investigation of visuospatial attention orienting has highlighted that goal-directed and stimulus-driven attention can operate via separate mechanisms (Klein and Shore, 2000; Corbetta and Shulman, 2002). A few behavioral studies have provided initial evidence that also stimulus-driven attention can facilitate the encoding in WM (Schmidt et al., 2002; Fine and Minnery, 2009; Botta et al., 2010), but the underlying neurophysiological bases are entirely unexplored. The second issue concerns the level of stimulus complexity used in previous studies. In everyday situations, the selection of relevant information typically involve many objects and events competing for access to memory. However, previous imaging studies used oversimplified paradigms in which the to-be-remembered stimuli consisted of simple and repeated stimuli, such as arrows (Murray et al., 2011) or colored/moving dots (Zanto and Gazzaley, 2009). Even when more complex stimuli were used, e.g., faces or buildings, these were presented in isolation without any context and competing information (Gazzaley et al., 2005; Majerus et al., 2007).

In the current fMRI study, we presented unrepeated pictures of naturalistic scenes that included multiple objects placed within realistic contexts. We used a biologically plausible computational model to characterize stimulus-driven visuospatial attention (Itti et al., 1998; Elazary and Itti, 2008) ("saliency maps"). For each picture, we indexed the saliency of a to-be-remembered target object, and we used these indexes to investigate the contribution of stimulus-driven signals for successful encoding in WM. We expected that objects at salient locations during encoding would be more likely successfully remembered at retrieval, compared to
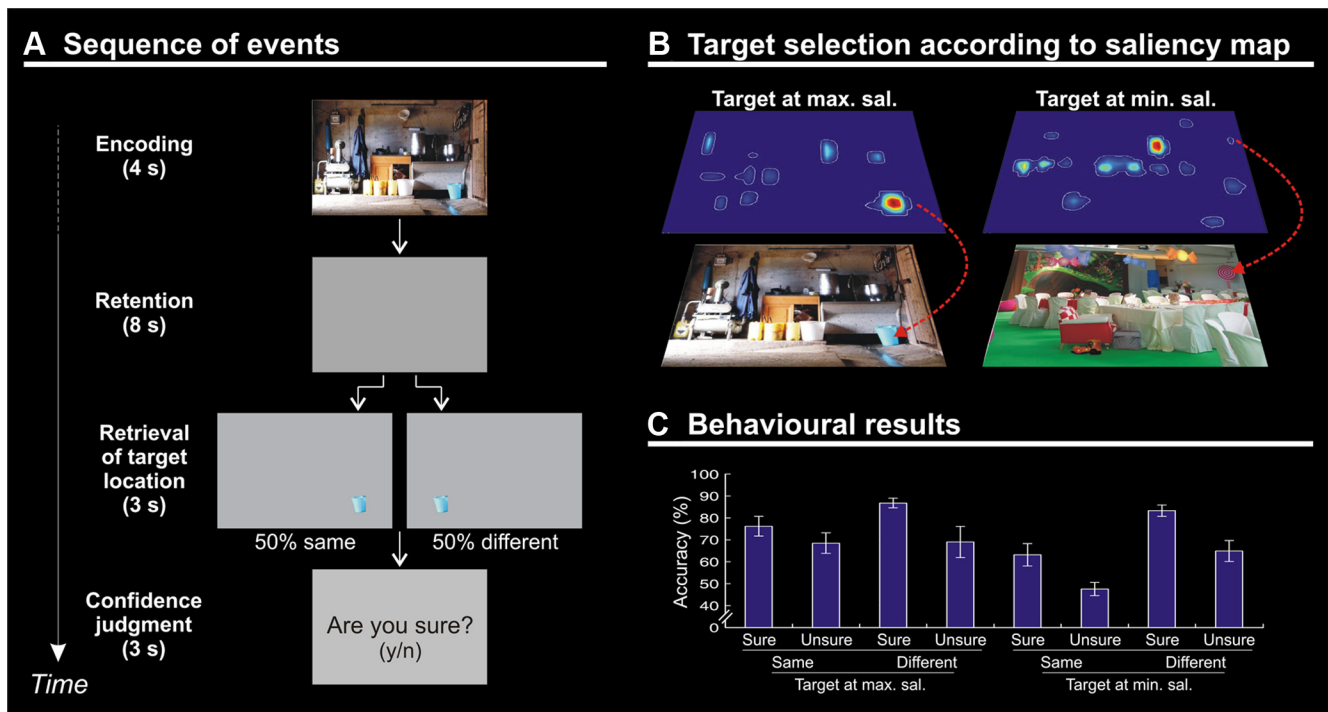
Figure 1. *A*, Diagram showing the sequence of events during one trial. The trial began with a picture presented for 4 s (encoding). A blank display was then shown for 8 s (delay). At retrieval, a single target object was presented either in the same or different location compared with the object's location in the original picture. Participants had to judge the same versus different location of the target (3 s) and then express a confidence judgment requiring a yes/no response (3 s). *B*, Examples of pictures including target objects selected at the point of maximal (left) or minimal salience (right). *C*, Mean accuracy (±SE) plotted according to the min/max saliency of target object at encoding, the same/different location at retrieval, and the sure/unsure subjective confidence rating (Table 1).

low-saliency objects. At the neurophysiological level, we expected an involvement of the parietal cortex that has been previously associated both with salience processing (Constantinidis and Steinmetz, 2005; Nardo et al., 2011) and WM (Todd and Marois, 2004; Xu and Chun, 2006).

Materials and Methods

Participants. Sixteen right-handed healthy volunteers took part in the fMRI experiment. Two participants were excluded from data analysis because of technical problems in the eye-movement recording, leaving 14 participants (six males; mean age, 24.9 years; range, 20–32 years). All participants gave written consent to the study, which was approved by the independent Ethics Committee of the Santa Lucia Foundation.

Stimuli and task. The task consisted in an encoding phase (4 s), a maintenance phase (8 s delay), a retrieval phase (discrimination of target-object location; 3 s), and, finally, a memory confidence judgment (3 s; Fig. 1 *A*). In the encoding phase, subjects were presented with pictures depicting scenes of everyday life. The picture set consisted in one hundred images collected on the World Wide Web. Pictures were characterized by the presence of multiple elements and objects (obviously) existing in relation to one another, resulting in a defined visual space (Henderson, 2003). The pictures included internal (e.g., a kitchen, a bathroom, etc.) and external scenes (e.g., a garden, a street, etc.), but no single-object photo or living things such as people or animals. Subjects were required to memorize as many details as possible for later retrieval. Pictures were displayed at 18 × 12° of visual angle (resolution in pixels, 680 × 448). The identity of the target object was unknown until retrieval (i.e., a delayed match-to-sample task). Following the 8 s delay (blank screen), subjects were shown a single target object cut out from the original picture and presented on a gray background (retrieval phase). In half of the trials, the target object was presented at the same location as in the original image, while in the other half of the trials the target object was presented at the mirror location in the opposite hemifield (Fig. 1 *A*; note that the object location was flipped, but the stimulus vertical axis re-

mained unchanged). Subjects had to report whether the target object was in the "same" versus "different" location with respect to the position at encoding by pressing one of two buttons with their right hand. Participants were then required to express a confidence judgment concerning their response. A display with the question "Are you sure? (y/n)" was presented, and participants pressed again one of the two response buttons. This was followed by a variable intertrial interval (1–3 s, uniformly distributed).

Participants underwent two fMRI runs (lasting ∼16.5 min each), including 50 trials each. The order of trials within and across runs was randomized with two constraints: (1) each run included half high-saliency targets and half low-saliency targets (see below, Computation of saliency indexes), and (2) for each target category (low or high salience), at retrieval, approximately half of the targets were presented in the same location (12 or 13 trials), and the other half in the different location (i.e., the remaining 12 or 13 trials).

Computation of saliency indexes. We used saliency maps to characterize the strength of the bottom-up external signals and the spatial distribution of these signals with respect to the designated target object (saliency indexes). Each picture was analyzed with the Saliency Toolbox 2.2 (http://www.saliencytoolbox.net/), which computes saliency maps using local discontinuities in line orientation, intensity contrast, and color opponency (Itti et al., 1998). Saliency maps were resampled at a resolution of 680 × 448 pixels. For half of the pictures, we designated as the target object the object located at the point of maximal salience (i.e., high-saliency targets), while for the other half we selected the target object at the location of minimal salience (i.e., low-saliency targets; Fig. 1 *B*). As an additional constraint, we made sure that the size of target objects did not significantly differ between high- and low- salience conditions ($p = 0.704$). The two-dimensional saliency maps were also used to extract picture/target-specific saliency indexes (Fig. 2 *A*). For this, each map was first scaled by its sum. This ensured that all maps had a total value equal to 1, while retaining the information about the relative spatial distribution of the saliency peaks. Next, we extracted object specific saliency
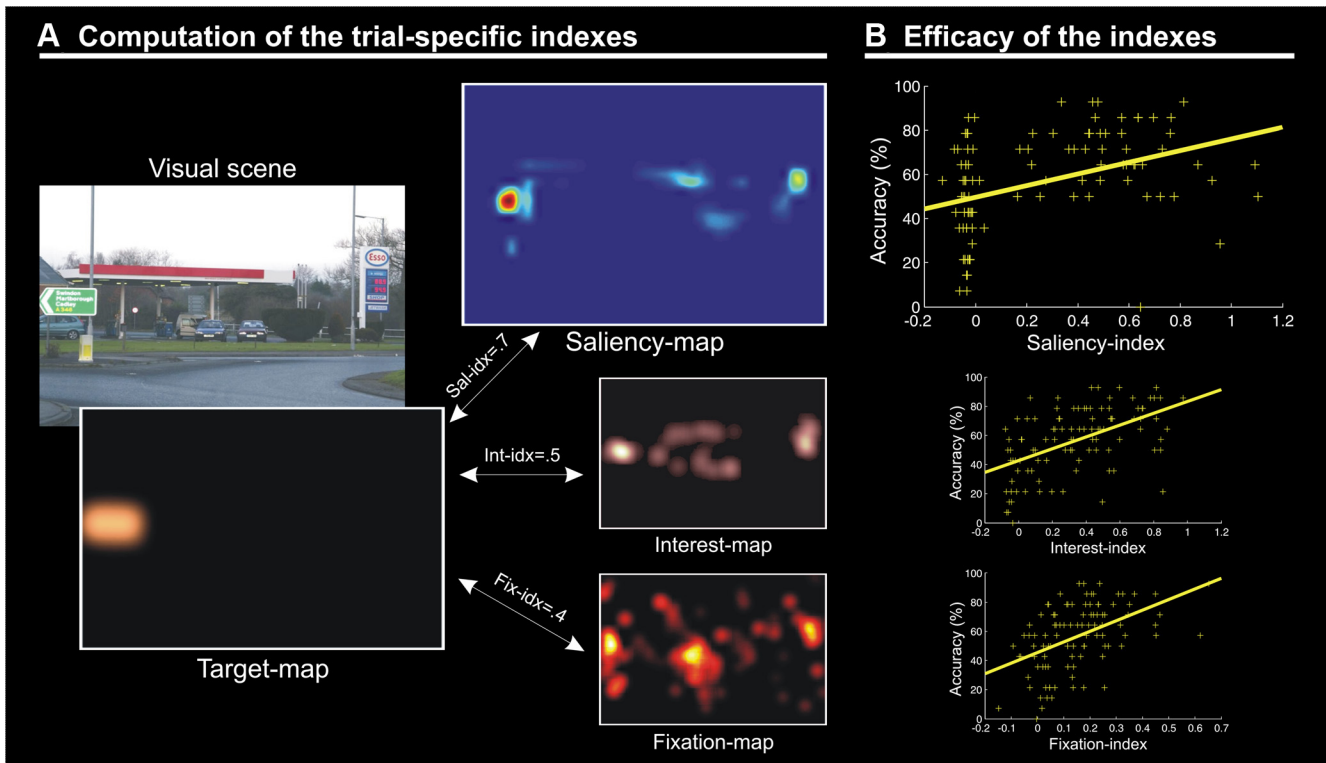
Figure 2. ***A***, Schematic illustration showing the computation of picture/target-specific saliency indexes from the two-dimensional saliency maps. For each picture, we constructed a target map including the location/shape of the target object within original image. This was then correlated with the saliency map associated with the same image, thus producing a numerical index for each picture. With the same procedure we generated also interest and fixation indexes related to endogenous/knowledge-based selection mechanisms. ***B***, Correlations between memory performance and saliency, interest, and fixation indexes. For each picture, the WM accuracy (i.e., the percentage of subjects who responded correctly to the same/different position discrimination and answered "sure" to the confidence judgment question) is plotted separately against the three indexes. This shows that retrieval performance increased when the target object was salient (top; $r = 0.39$, $p <$ 0.001), when the target object was judged to be interesting/relevant (middle; $r = 0.52$, $p <$ 0.001), and when subjects fixated the target object during encoding (bottom; $r = 0.47$, $p <$ 0.001).

indexes by computing the correlation (Fisher transformed to Z) between the saliency map and the "target map". The target map consisted of a 680×448 matrix with values set equal to 1 at the coordinates of the target object and zeros everywhere else, smoothed with a Gaussian filter (FWHM, 1°) and scaled by its sum (Fig. 2A). The resulting index will be high when the designated target object corresponds to a salient location in the image.

Together with the main saliency index, for each picture we also computed interest and fixation indexes (Fig. 2). The interest indexes were derived from "interest maps" that are thought to capture the contribution of internal knowledge/representations for the processing of complex images (Einhäuser et al., 2008; Masciocchi et al., 2009). A separate group of 18 volunteers (nine males; mean age, 23.0 years; range, 21–25 years) was asked to view all 100 pictures on a computer display (sequentially, in a random order) and to report the three most relevant/interesting objects within each picture. Participants clicked with a mouse on the most relevant/interesting object and then provided a verbal label for that object by typing on a computer keyboard the object's name. With the same procedure, they then indicated the second and the third most relevant/interesting object. The interest maps were generated by assigning the values 3, 2, or 1 (first, second, or third most relevant/interesting object) to the coordinates corresponding to the positions of the mouse click. These interest maps have the highest values in regions of space comprising objects that were considered most relevant/interesting by the majority of the participants.

Finally, for each picture we also computed a "fixation index." Fixation indexes were derived from "fixation maps" based on fixations' location and duration. The gaze position was recorded during fMRI using an ASL eye-tracking system, adapted for use in the scanner (Applied Science Laboratories; Model 504; sampling rate 60 Hz). Eye-position traces were examined in a 4 s window, starting from the picture onset (i.e., the duration of the encoding phase). We identified fixations as time windows

(minimum duration, 100 ms) where the velocity of the eye traces did not exceed 50°/s. For each picture, the fixation map was generated by assigning the value of the fixations' durations to the coordinates of the corresponding gaze positions (median of the x/y position within the fixation window; Fig. 2A). The fixation maps have high values at locations that were fixated more frequently and for longer times (for review, see Henderson, 2003). Interest and fixation indexes were used primarily to confirm the specificity of the effect of saliency indexes on WM performance (see below, Behavioral data).

Magnetic resonance imaging. A Siemens Allegra (Siemens Medical Systems) operating at 3T and equipped for echoplanar imaging (EPI) was used to acquire the functional magnetic resonance images. 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 functional MR images were acquired using blood oxygenation level-dependent imaging (3×3 mm; 2.5 mm thick; 50% distance factor; repetition time, 2.08 s; time echo, 30 ms), covering the entirety of the cortex.

fMRI data analysis. We used SPM8 (Wellcome Department of Cognitive Neurology, London, UK) implemented in MATLAB 7.4 (Math-Works) for data preprocessing and statistical analyses. Each participant underwent two fMRI runs, each comprising 490 volumes. After having discarded the first four volumes of each run, all images were corrected for head movements. Slice-acquisition delays were corrected using the middle slice as reference. All images were normalized to the standard SPM8 EPI template, resampled to 2 mm isotropic voxel size, and spatially smoothed using an isotropic Gaussian kernel of 8 mm FWHM. Time series at each voxel for each participant were high-pass filtered at 128 s and prewhitened by means of autoregressive model AR(1).

Statistical inference was based on a random effects approach (Penny and Holmes, 2004), which comprised two steps: first-level analyses estimating contrasts of interest for each subject followed by second-level
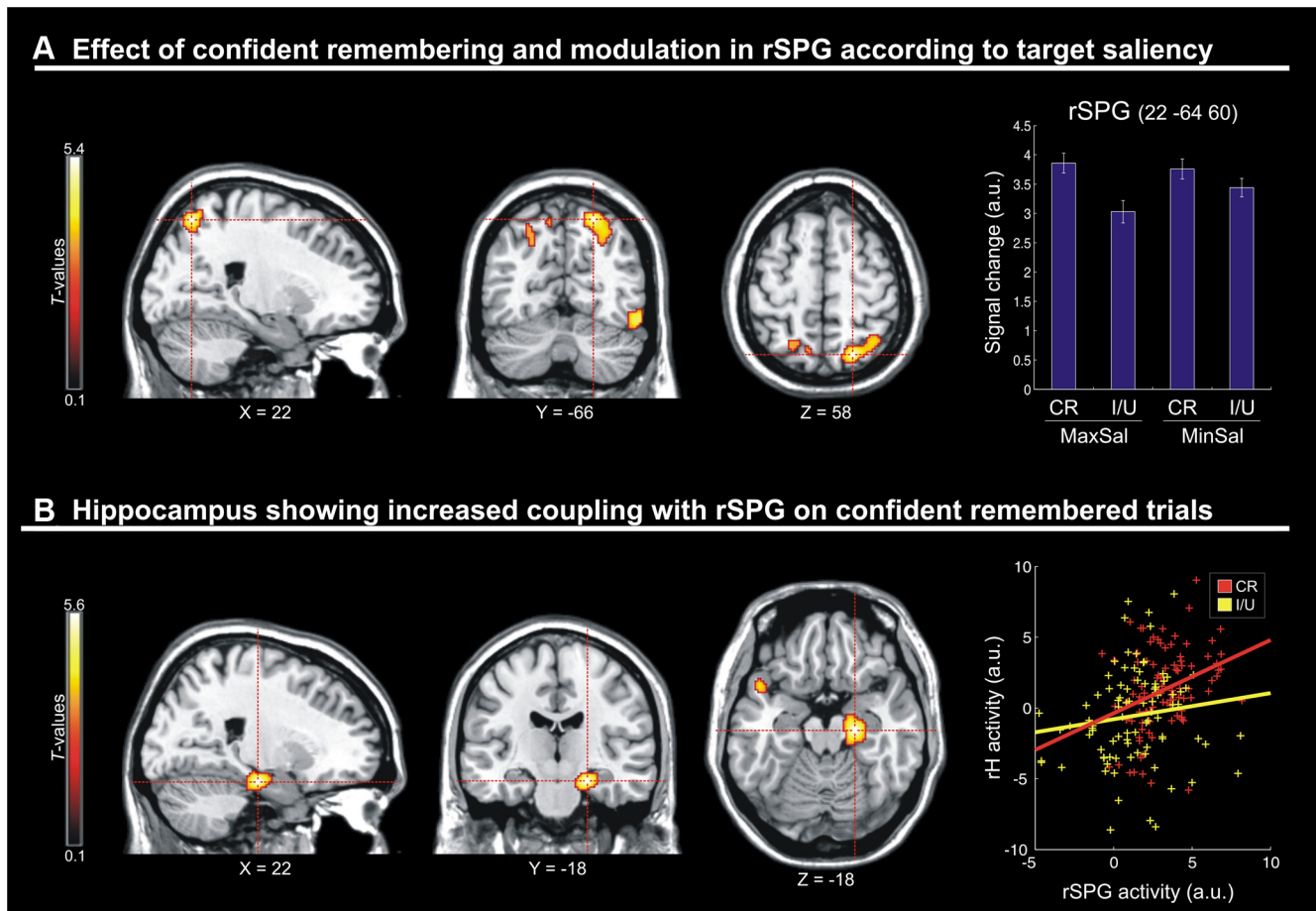
Figure 3. *A*, Sections through the peak of activation in the rSPG that showed a significant main effect of confident remembering at encoding [confident remembered (CR) vs incorrect and/or unsure (I/U) targets]. Activity in this regions was further modulated by the saliency of the target object, with a greater effect of confident memory when, at encoding, the target object was at a point of maximal compared to minimal salience (compare CR bar minus I/U bar for MaxSal and MinSal conditions in the signal plot). *B*, Sections through the right hippocampus that showed increased functional coupling with the rSPG selectively on confident remembered trials. The rightmost panel shows a scatter plot with the signals in the two areas, as a function of memory performance. Data were extracted from the single-subject BOLD time series, five volumes after the onset of each fMRI run (to account for the hemodynamic delay). The plot includes the mean signals for 99 confident remembered pictures and 89 incorrect/unsure pictures, for which there were at least three subjects contributing to the computation of the mean. For display purposes, all activation maps are displayed at a threshold of $p < 0.001$ (uncorrected).

analyses for statistical inference at the group level (with nonsphericity correction) (Friston et al., 2002). For each subject, the first-level multiple regression model included four conditions at encoding, given by the crossing of "target saliency" (high vs low) and "confident memory" [correct responses followed by confident judgment ("yes, I'm sure") vs incorrect responses and/or unsure judgments], plus four conditions at retrieval (confident memory same vs different target location). It should be noted that with the current two-alternative forced-choice task (same vs different target location), subjects may provide quite a few "correct" responses even when failing to actually recollect the target-object location (50% chance level). This was one of the reasons why the main fMRI analysis investigated confident memory by comparing correct/sure trials with all other trials (for an additional control analysis using "sure" trials only, see Results).

The four conditions at encoding were modeled as miniblocks, time locked at the onset of the pictures with a duration of 4 s. The retrieval conditions were modeled as delta functions time locked at the onset of the test stimulus (duration, 0 s; to minimize any influence of the subsequent parts of the trial, i.e., motor responses and confidence judgment). All predictors were convolved with the SPM8 hemodynamic response function, and the parameters of head movements were included as covariates of no interest.

The group-level analysis concerned the encoding phase of the trial, when any stimulus-driven effect related to the salience of the target object is expected to take place (Gazzaley and Nobre, 2012). A within-subject

ANOVA modeled the four relevant events types [i.e., saliency (high/low) × memory (successful/unsuccessful)]. First, we highlighted activity associated with confident memory (correct/sure vs incorrect and/or unsure trials). The statistical threshold was set to $p < 0.05$, FWE corrected at the voxel level, considering the whole brain as the volume of interest. Next, we tested whether the area showing an effect of successful memory was further modulated depending on target saliency, i.e., whether there was a larger effect of correct/sure versus incorrect/unsure trials for high-compared to low-saliency targets. For this, we considered a sphere (radius, 8 mm; matching the FWHM of smoothing filter) centered on the parietal voxel showing a significant effect of confident memory (Fig. 3A) as the volume of interest (small volume correction) (Worsley et al., 1996).

Together with these standard intraregional analyses, we used analyses of inter-regional connectivity [psychophysiological interactions (PPIs)] (Friston, 2004) to explore the functional connectivity of the parietal region [the right superior parietal gyrus (rSPG)] that showed a significant effect of confident memory and that was further modulated according to saliency. We used the "Generalized Form of Context-Dependent Psychophysiological Interactions" SPM toolbox (McLaren et al., 2012). At the subject level, each PPI analysis included eight regressors corresponding to the psychological variables of interest (i.e., four event types at encoding and four at retrieval, as in the main analysis), the time course of the parietal seed area (the physiological variable), and the critical cross-products (i.e., the psychophysiological interaction term) between the

Table 1. Retrieval performance in the different experimental conditions

	Maximal saliency				Minimal saliency			
	Same		Different		Same		Different	
	Sure	Unsure	Sure	Unsure	Sure	Unsure	Sure	Unsure
Accuracy (%)	76.2 ± 4.5	68.5 ± 4.7	86.8 ± 2.2	69.0 ± 7.1	63.2 ± 5.1	47.6 ± 3.0	83.3 ± 2.6	64.9 ± 4.8
Total	287	41	284	57	216	107	218	117
Range	12–25	0–13	15–24	1–7	6–16	0–16	7–24	0–14

Accuracy (percentage correct ± SE), total number of correct trials across subjects ($n = 14$), and range of the number of correct trials within subjects, are shown as a function of the target saliency (maximum/minimum) at encoding, the target position at retrieval (same/different), and the subjective confidence rating (sure/unsure).

four psychological variables at encoding and time course of the seed areas. The head motion realignment parameters were included as covariates of no interest. The parameter estimates of the four PPI regressors entered a two-by-two within-subject ANOVA for statistical inference at the group level. We tested for changes of parietal connectivity with the rest of the brain, as a function of memory (i.e., areas showing a increased coupling with parietal cortex at encoding, when the trial would subsequently be remembered with high confidence at retrieval), and for the interaction between target saliency and memory. The statistical threshold was set to $p < 0.05$, FWE corrected at the voxel level, considering the whole brain as the volume of interest.

Results

Behavioral data

On each trial, subjects reported the same/different location of the target object and then provided a confidence judgment (sure/unsure). On the majority of trials, participants were confident about their response (72% sure vs 25% unsure, plus 3% no response trials; Table 1).

A three-way within-subjects ANOVA with the factors of target saliency at encoding (high vs low), target location (same vs different), and confidence judgment (sure vs unsure) at retrieval was performed on the accuracy data (Fig. 1C). This analysis revealed a main effect of target saliency ($F_{(1,13)} = 11.1$, $p = 0.005$): Subjects were more accurate in judging the target location at retrieval when, at encoding, the target object was at the point of maximal saliency compared with the point of minimal saliency (75 vs 65%). The ANOVA also showed a main effect of target location ($F_{(1,13)} = 6.0$, $p = 0.030$), with greater accuracy when at retrieval the target object was presented in a different location compared with the object's location at encoding (76 vs 64%). The effect of location was driven primarily by targets at points of minimal saliency, as formally evidenced by the significant interaction between saliency and location ($F_{(1,13)} = 6.5$, $p = 0.024$; different minus same, 19% for minimum vs 6% for maximum saliency). Finally, this analysis revealed a main effect of confidence ($F_{(1,13)} = 36.2$, $p < 0.001$), indicating that when participants were 'sure' about their response, they were indeed more accurate in remembering the target location than when they were unsure (77 vs 63%). This analysis did not reveal any other significant effect (all F values < 1.5; all p values > 0.248).

We confirmed the effect of saliency on memory performance using object-specific saliency indexes on a trial-by-trial basis (Fig. 2A). For each picture, we computed the correlation between target-specific saliency indexes and the corresponding WM performance. Figure 2B (top) shows the positive correlation between target salience and the proportion of subjects that confidently remembered the targets ($r = 0.39$, $p < 0.001$): the higher the value of the saliency index, the greater the number of participants who confidently remembered that target.

We further confirmed the specificity of the effect of stimulus-driven signals on WM performance by performing a multiple regression analysis that included additional indexes related to

voluntary, knowledge-based selection (Einhäuser et al., 2008) (interest and fixation indexes; see Materials and Methods; Fig. 2B). The omnibus F test was significant ($F_{(3,96)} = 20.6$, $p < 0.001$), demonstrating an overall effect of the three predictors on WM performance. We assessed the contribution of the saliency index revealing that the intrinsic object's low-level proprieties predicted retrieval performance ($t = 2.3$, $p = 0.023$), even after accounting for the other two indexes. This demonstrates that stimulus-driven signals provide a distinctive contribution to successful memory performance, over and above any contribution of knowledge-based factors. Analogous t tests now considering the distinctive contribution of the interest and fixation indexes were also significant (interest index, $t = 3.6$, $p < 0.001$; fixation index, $t = 3.3$, $p = 0.001$). Thus, the more an object was labeled as interesting/relevant by an independent group of participants (interest index), or the more and object was fixated (fixation index), the higher the chance for that target object to be confidently remembered in the current discrimination task. These findings extend the previous literature using knowledge-based indexes (Einhäuser et al., 2008; Masciocchi et al., 2009), demonstrating here, for the first time, the impact of these factors on WM performance.

Finally, we examined the possibility that factors related to the global configuration of the scene, rather than aspects specifically associated with the target object (i.e., its saliency level), also contributed to the retrieval performance. We used an "overall saliency distribution" index and an "overall object distribution" index to recategorize the 100 images, and we tested whether these global indexes also predicted retrieval accuracy. The overall saliency distribution was computed using the saliency maps and counting the number of pixels with values larger than zero. This index will be larger for pictures including a widespread saliency distribution than for pictures with a more restricted pattern of saliency (i.e., just a few peaks). The overall object distribution was computed using the interest maps and again counting the number of pixels with values larger than zero. This index will tend to increase when the scene includes many objects that subjects' considered relevant/interesting. Two-sample t tests assessed whether pictures with high versus low saliency distributions and pictures with high versus low object distributions lead to different retrieval performance. Both tests were not significant (saliency distribution, $t = -0.13$, $p = 0.900$; object distribution, $t = 1.14$, $p = 0.255$), indicating that these factors associated with the global configuration of the scene did not affect memory performance (for additional details about this point, see Discussion).

fMRI data

First, the fMRI analyses tested for areas showing an effect of successful confident memory during the encoding phase, comparing correct/sure trials minus incorrect and/or unsure trials. This revealed a significant cluster of activation in the rSPG ($x, y, z = 22, -66, 58$; $t = 5.38$; $p = 0.044$, corrected; Fig. 3A). Activity in this

region at encoding predicted whether the target object would be confidently remembered at retrieval, highlighting the importance of stimulus processing at encoding for successful WM performance (Pessoa et al., 2002). At a lower uncorrected level of significance, the same comparison showed activation also in other regions of the frontoparietal network, including the left posterior parietal cortex (PPC) and the left middle frontal gyrus (Fig. 3A).

Next, we turned to our main hypothesis testing whether target saliency would further modulate activity associated with the successful encoding in WM. Indeed, in the rSPG we found a significant interaction between WM performance and target saliency ($x, y, z = 22, -64, 60; t = 2.99; p = 0.044$, small volume corrected; see Materials and Methods). The effect of successful memory (correct/sure vs incorrect and/or unsure) was larger in the context of high-saliency targets compared to low-saliency targets (Fig. 3A, signal plot). Accordingly, stimulus-driven attentional signals affected rSPG activation associated specifically with successful encoding in WM.

To ensure that our main analysis targeted processes associated with reliable recollection, rather than correct response by chance or familiarity-based processes (Diana et al., 2007), all trials followed by unsure confidence ratings were combined with the incorrect responses. However, we also performed an additional fMRI analysis considering only trials followed by sure confidence judgments. This control analysis crossed again the factors of target saliency (maximum vs minimum) and memory performance (successful vs unsuccessful), but now with the unsuccessful conditions including only incorrect trials followed by sure judgments. In the rSPG, we found a pattern of activity similar to that in the main analysis, albeit with lower statistical values (salience by memory interaction, $t = 2.04; p = 0.016$, uncorrected). The lower statistics is not surprising given that this control analysis included, overall, ~25% fewer trials than the main analysis, and almost 60% fewer trials in the "forgotten" minimum saliency condition (Table 1). Despite this limitation, this control analysis confirmed the pattern of interaction in the rSPG, now considering only trials with equal levels of confidence rating (i.e., sure trials only).

Finally, we investigated the possible contribution of intraregional network connectivity, using psychophysiological interactions (Friston, 2004). We seeded the connectivity analysis in the rSPG and asked whether the functional coupling of this region with any other region of the brain changed as a function of WM performance. This analysis revealed that the functional coupling between rSPG and the medial temporal lobe (MTL) increased selectively in correct/sure trials ($x, y, z = 22, -18, -18; t = 5.63; p = 0.048$, corrected; Fig. 3B), without any further modulation according to saliency. The medial temporal cluster showing increased connectivity with rSPG was centered in the right hippocampus and extended ventrally to the adjacent parahippocampal cortex.

Discussion

We used a delayed match-to-sample task to examine the contribution of stimulus-driven signals for the successful encoding of object-related information in WM. Participants judged the location of target objects (retrieval phase) that were extracted from complex and unrepeated visual scenes (encoding phase). Retrieval performance was more accurate when, at encoding, the to-be-judged target was located at a salient location. The functional imaging analyses associated the effect of stimulus-driven salience on WM performance with encoding-related activity in

the rSPG. This region showed an increased functional coupling with the medial temporal cortex selectively in confidently remembered trials.

The enhancement of WM performance for objects at salient locations demonstrates that objects' intrinsic properties determine encoding prioritization, consistent with the "prior entry" hypothesis (Spence and Parise, 2010) and the capture of attention by salient stimuli (Desimone and Duncan, 1995; Kastner and Ungerleider, 2000). Our findings agree with a previous behavioral study that also reported enhanced WM performance for objects at salient locations (Fine and Minnery, 2009). Here, we further specify the distinctive contribution of bottom-up signals for successful WM in a task where the identity of the target object was discovered only at retrieval. This rules out any contribution of strategic/voluntary attention during encoding (Fine and Minnery, 2009). Moreover, additional analyses of the behavioral data revealed that object salience predicted WM performance, even after accounting for the effect of other knowledge-based factors that here were also found to affect WM performance (i.e., interest and fixation indexes). In contrast, the overall saliency and overall object distributions did not predict performance (see Results).

Nonetheless, we acknowledge that naturalistic scenes can be categorized using a virtually unlimited number of parameters, and that other global configuration factors may affect retrieval performance (Jiang et al., 2000; Blalock and Clegg 2010). The role of global scene configuration could in fact explain why subjects were more accurate in different- versus same-location trials. The presentation of a target object in a different position compared to the original picture can yield to violations of the global scene configuration (e.g., a glass presented in a position where there was no table in the original scene). This would make the position judgment easier in different-location trials. Thus, in the context of the current localization task, the overall scene configuration may have provided subjects with a spatial frame to encode the objects' position. With respect to global scene's proprieties, it is also important to notice that the computation of the saliency maps includes several steps involving reciprocal influences between pixels/locations that are far apart in the image (e.g., cross-scale subtractions, differences of Gaussian). Accordingly, the number, spatial distribution and sensory characteristics (i.e., salience) of the various objects in the scene will influence the saliency level of the target object, and vice versa. These interdependencies between the target object and other global properties of the scene constitute a limitation for the interpretation of the current results, but they are also at the core of the competitive mechanisms that we seek to characterize using computational models of bottom-up salience (Itti et al., 1998).

The fMRI data associated the interplay between salience and successful encoding in WM with activation of the PPC. The PPC has been implicated in many aspects of cognitive control, also prompting the suggestion of a general role of the parietal cortex in "task effort." However, mere cognitive effort seems unlikely to account for our findings. First, cognitive effort is typically associated with concurrent activation of parietal and frontal regions (Crittenden and Duncan, 2012), while here the critical salience-by-memory interaction affected parietal cortex only. Moreover, the current interaction pattern would imply that salient target objects require greater effort than low-saliency objects, which is contrary to previous behavioral data showing that subjects are faster to discriminate salient compared to nonsalient stimuli (Geng and Mangun, 2009).

Instead we associate the interaction between salience and memory performance with trial-to-trial variations of the re-

sources dedicated to the processing of salient objects. Previous studies showed that brain activity during encoding can predict subsequent memory performance (Pessoa et al., 2002). This relationship is thought to arise because the amount of resources that subjects dedicate to the processing of the stimuli at encoding "fluctuates" from trial to trial, and this leads to different levels of performance at retrieval (Anticevic et al., 2010; White et al., 2012). Here we go a step further showing that these fluctuations specifically affected processing/encoding of salient signals. This further selectivity is consistent with the view that PPC represents salient signals, and therefore, any trial-to-trial fluctuation will determine memory performance only when the retrieval task probes stimuli/objects at salient locations.

The finding of stimulus-driven attentional effects in dorsal parietal cortex during successful WM encoding challenges traditional WM models. These either emphasize a "storage account," with parietal cortex acting as a capacity-limited buffer (Todd and Marois, 2004; Xu and Chun, 2006; Mitchell and Cusack, 2008), or advocate for an active attentional involvement primarily during WM maintenance (Berryhill et al., 2011; see also Santangelo and Macaluso, 2013). A few previous studies reported WM encoding-related effects in PPC, but always involving manipulation of endogenous/goal-directed attention (Majerus et al., 2007; Gillebert et al., 2012). In contrast, here we demonstrate for the first time the specific contribution of bottom-up salience rather than explicit task requirements [e.g., recall vs recognition (Berryhill et al., 2011); judgments of stimulus identity vs order (Majerus et al., 2007)].

The interaction between saliency and successful memory was localized in the posterior part of the superior parietal lobule that includes many different subregions (Husain and Nachev, 2007). Based on anatomical criteria, we localized the relevant cluster in the SPG. The superior parietal cortex plays a key role in the representation of space (Gross and Graziano, 1995; Buneo and Andersen, 2006), and parietal areas have been associated previously with the representation of visual salience, both in monkey electrophysiology (Constantinidis and Steinmetz, 2005) and in human fMRI (Geng and Mangun, 2009; Bogler et al., 2011; Nardo et al., 2011). Furthermore, the parietal cortex has been involved in the integration of multiple sources of (contextual) information when localizing objects in space (Shafritz et al., 2002; Walter and Dassonville, 2008). Our results are consistent with previous proposals that PPC prioritizes salient signals for spatial perception and localization (Gottlieb, 2007; Berman and Colby, 2009; Nardo et al., 2011), but here extend this to the memory domain for the first time. We propose that attention control in PPC assigns spatial priorities to objects at salient locations and contributes to the representation of the spatial layout of these objects. This, in turn, facilitates the encoding of spatial information in WM.

Our findings also revealed changes of functional connectivity between PPC and the MTL depending on memory performance. During encoding, we found greater functional coupling between rSPG and the hippocampus for scenes that included objects that would be successfully retrieved compared with incorrect or unsure retrieval. The MTL has been long associated with "item–context binding," with some distinction between the perirhinal cortex that would primarily process item information, the parahippocampal cortex supporting contextual representations, and the hippocampus that would combine item and context information (Epstein and Kanwisher, 1998; Davachi, 2006; Diana et al., 2007). Here we report increased connectivity between parietal cortex and the hippocampus, consistent with the proposal that

the current spatial judgment task involved linking object and global contextual information. Related to this, a previous study by Howard et al. (2011) showed that the hippocampus responds to changes in the spatial relationship between objects and context, while the parahippocampal cortex was found to engage for scene novelty. Moreover, this previous study manipulated the object/context relationship within relatively short time frames [hundreds of milliseconds in the study by Howard et al. (2011) vs a few seconds in the present study], highlighting the rapid object-context binding operated by the hippocampus.

We interpret the functional coupling between PPC and MTL in relation to Cowan's (2011) model of WM. This model links attention and WM, suggesting that attention-related processes in PPC activate a limited number of preexisting (long-term) representations in MTL (Messinger et al., 2001; Squire et al., 2004). Accordingly, the limited capacity of WM would be intrinsically related to corresponding limitations in focusing (and maintaining) attention on just a few item representations at the time. We propose that PPC processes low-level sensory signals to assign priorities to a subset of objects in the scene (selected according to bottom-up saliency). The spatial layout of the selected objects is coded within the global context of the scene. This would enable the storage of object-related spatial information in WM, in a process requiring dynamic PPC/MTL coupling. These stimulus-driven selection processes during encoding then enable correct spatial judgments at retrieval.

We conclude that PPC plays a central role in stimulus-driven spatial selection for successful WM encoding. Using ecologically valid material, and combining behavioral, computational, and imaging methods, we identified a parietotemporal circuit that integrates available sensory cues (attention-grabbing saliency signals) and current WM requirements (storing objects' locations). These results contribute to bridge the gap between attention and memory, two cognitive systems that jointly govern spatial orienting behavior in complex, everyday life situations.

References

Anticevic A, Repovs G, Shulman GL, Barch DM (2010) When less is more: TPJ and default network deactivation during encoding predicts working memory performance. Neuroimage 49:2638–2648. CrossRef Medline

Berman R, Colby C (2009) Attention and active vision. Vision Res 49:1233–1248. CrossRef Medline

Berryhill ME, Chein J, Olson IR (2011) At the intersection of attention and memory: the mechanistic role of the posterior parietal lobe in working memory. Neuropsychologia 49:1306–1315. CrossRef Medline

Blalock LD, Clegg BA (2010) Encoding and representation of simultaneous and sequential arrays in visuospatial working memory. Q J Exp Psychol 63:856–962. CrossRef

Bogler C, Bode S, Haynes JD (2011) Decoding successive computational stages of saliency processing. Curr Biol 21:1667–1671. CrossRef Medline

Botta F, Santangelo V, Raffone A, Olivetti Belardinelli M, Lupianez J (2010) Exogenous and endogenous spatial attention effects on visuo-spatial working memory. Q J Exp Psychol 63:1590–1602. CrossRef Medline

Buneo CA, Andersen RA (2006) The posterior parietal cortex: Sensorimotor interface for planning and online control of visually guided movements. Neuropsychologia 44:2594–2606. CrossRef Medline

Chun MM, Turk-Browne NB (2007) Interactions between attention and memory. Curr Opin Neurobiol 17:177–184. CrossRef Medline

Constantinidis C, Steinmetz MA (2005) Posterior parietal cortex automatically encodes the location of salient stimuli. J Neurosci 25:233–238. CrossRef Medline

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

Cowan N (2011) The focus of attention as observed in visual working memory tasks: making sense of competing claims. Neuropsychologia 49:1401–1406. CrossRef Medline

Craik FI, Govoni R, Naveh-Benjamin M, Anderson ND (1996) The effects

of divided attention on encoding and retrieval processes in human memory. J Exp Psychol Gen 125:159–180. CrossRef Medline

Crittenden BM, Duncan J (2012) Task difficulty manipulation reveals multiple demand activity but no frontal lobe hierarchy. Cereb Cortex. Advance online publication. Retrieved Nov 6, 2012. doi:10.1093/cercor/bhs333. CrossRef

Davachi L (2006) Item, context and relational episodic encoding in humans. Curr Opin Neurobiol 16:693–700. CrossRef Medline

Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222. CrossRef Medline

Diana RA, Yonelinas AP, Ranganath C (2007) Imaging recollection and familiarity in the medial temporal lobe: A three-component model. Trends Cogn Sci 11:379–386. CrossRef Medline

Einhäuser W, Spain M, Perona P (2008) Objects predict fixation better than early saliency. J Vis 8(14):18 1–26. CrossRef

Elazary L, Itti L (2008) Interesting objects are visually salient. J Vis 8(3):3 1–15. CrossRef

Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. Nature 392:598–601. CrossRef Medline

Fine MS, Minnery BS (2009) Visual salience affects performance in a working memory task. J Neurosci 29:8016–8021. CrossRef Medline

Friston KJ (2004) Functional and effective connectivity in neuroimaging: a synthesis. Hum Brain Mapp 2:56–78.

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

Gazzaley A, Nobre AC (2012) Top-down modulation: bridging selective attention and working memory. Trends Cogn Sci 16:129–135. CrossRef Medline

Gazzaley A, Cooney JW, McEvoy K, Knight RT, D'Esposito M (2005) Top-down enhancement and suppression of the magnitude and speed of neural activity. J Cogn Neurosci 17:507–517. CrossRef Medline

Geng JJ, Mangun GR (2009) Anterior intraparietal sulcus is sensitive to bottom-up attention driven by stimulus salience. J Cogn Neurosci 21:1584–1601. Medline

Gillebert CR, Dyrholm M, Vangkilde S, Kyllingsbæk S, Peeters R, Vandenberghe R (2012) Attentional priorities and access to short-term memory: parietal interactions. Neuroimage 62:1551–1562. CrossRef Medline

Gottlieb J (2007) From thought to action: the parietal cortex as a bridge between perception, action, and cognition. Neuron 53:9–16. CrossRef Medline

Gross CG, Graziano MS (1995) Multiple representations of space in the brain. Neuroscientist 1:43–50. CrossRef

Henderson JM (2003) Human gaze control during real-world scene perception. Trends Cogn Sci 7:498–504. CrossRef Medline

Howard LR, Kumaran D, Ólafsdóttir HF, Spiers HJ (2011) Double dissociation between hippocampal and parahippocampal responses to object-background context and scene novelty. J Neurosci 31:5253–5261. CrossRef Medline

Husain M, Nachev P (2007) Space and the parietal cortex. Trends Cogn Sci 11:30–36. CrossRef Medline

Itti L, Koch C, Niebur E (1998) A model of saliency-based visual attention for rapid scene analysis. IEEE Trans Pattern Anal Mach Intell 20:1254–1259. CrossRef

Jiang Y, Olson IR, Chun MM (2000) Organization of visual short-term memory. J Exp Psychol Learn Mem Perf 26:683–702. CrossRef

Kastner S, Ungerleider LG (2000) Mechanisms of visual attention in the human cortex. Annu Rev Neurosci 23:315–341. CrossRef Medline

Klein RM, Shore D (2000) Control of cognitive processes. In: Attention and performance XVIII. (Monsell S, Driver J, eds), pp 195–208. Cambridge, MA: MIT.

Luck SJ, Vogel EK (1997) The capacity of visual working memory for features and conjunctions. Nature 390:279–281. CrossRef Medline

Majerus S, Bastin C, Poncelet M, Van der Linden M, Salmon E, Collette F, Maquet P (2007) Short-term memory and the left intraparietal sulcus: focus of attention? Further evidence from a face short-term memory paradigm. Neuroimage 35:353–367. CrossRef Medline

Masciocchi CM, Nihalas S, Parkhurst D, Niebur E (2009) Everyone knows what is interesting: salient locations which should be fixated. J Vis 9(11):25 1–22. CrossRef

McLaren DG, Ries ML, Xu G, Johnson SC (2012) A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. Neuroimage 61:1277–1286. CrossRef Medline

Messinger A, Squire LR, Zola SM, Albright TD (2001) Neuronal representations of stimulus associations develop in the temporal lobe during learning. Proc Natl Acad Sci U S A 98:12239–12244. CrossRef Medline

Mitchell DJ, Cusack R (2008) Flexible, capacity-limited activity of posterior parietal cortex in perceptual as well as visual short-term memory tasks. Cereb Cortex 18:1788–1798. CrossRef Medline

Murray AM, Nobre AC, Stokes MG (2011) Markers of preparatory attention predict visual short-term memory performance. Neuropsychologia 49:1458–1465. CrossRef Medline

Naghavi HR, Nyberg L (2005) Common fronto-parietal activity in attention, memory, and consciousness: share demands on integrations? Conscious Cogn 14:390–425. CrossRef Medline

Nardo D, Santangelo V, Macaluso E (2011) Stimulus-driven orienting of visuo-spatial attention in complex dynamic environments. Neuron 69:1015–1028. CrossRef Medline

Penny WD, Holmes AP (2004) Random effects analysis. In: Human brain function II, Ed 2 (Frackowiak RSJ, Friston KJ, Frith CD, Dolan R, Price CJ, Zeki S, Ashburner J, Penny WD, eds), pp 843–851. San Diego: Elsevier.

Pessoa L, Gutierrez E, Bandettini P, Ungerleider L (2002) Neural correlates of visual working memory: fMRI amplitude predicts task performance. Neuron 35:975–987. CrossRef Medline

Santangelo V, Macaluso E (2013) The contribution of working memory to divided attention. Hum Brain Mapp 34:158–175. CrossRef Medline

Schmidt BK, Vogel EK, Woodman GF, Luck SJ (2002) Voluntary and automatic attentional control of visual working memory. Percept Psychophys 64:754–763. CrossRef Medline

Shafritz KM, Gore JC, Marois R (2002) The role of the parietal cortex in visual feature binding. Proc Natl Acad Sci U S A 99:10917–10922. CrossRef Medline

Spence C, Parise C (2010) Prior-entry: a review. Conscious Cogn 19:364–379. CrossRef Medline

Squire LR, Stark CE, Clark RE (2004) The medial temporal lobe. Annu Rev Neurosci 27:279–306. CrossRef Medline

Todd JJ, Marois R (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. Nature 428:751–754. CrossRef Medline

Walter E, Dassonville P (2008) Visuospatial contextual processing in the parietal cortex: an fMRI investigation of the induced Roelofs effect. Neuroimage 42:1686–1697. CrossRef Medline

White TP, Jansen M, Doege K, Mullinger KJ, Park SB, Liddle EB, Gowland PA, Francis ST, Bowtell R, Liddle PF (2012) Theta power during encoding predicts subsequent-memory performance and default mode network deactivation. Hum Brain Mapp. Advance online publication. Retrieved June 19, 2012. doi:10.1002/hbm.22114. CrossRef

Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC (1996) A unified statistical approach for determining significant signals in images of cerebral activation. Hum Brain Mapp 4:58–73. CrossRef Medline

Xu Y, Chun MM (2006) Dissociable neural mechanisms supporting visual short-term memory for objects. Nature 440:91–95. CrossRef Medline

Zanto TP, Gazzaley A (2009) Neural suppression of irrelevant information underlies optimal working memory performance. J Neurosci 29:3059–3066. CrossRef Medline