

ORIGINAL ARTICLE

Neural Mechanisms for the Benefits of Stimulus-Driven Attention

Katelyn M. Wills¹, Jingtai Liu¹, Jonathan Hakun², David C. Zhu^{1,3}, Eliot Hazeltine⁴, and Susan M. Ravizza¹

¹Department of Psychology, Michigan State University, East Lansing, MI 48824, USA, ²Department of Psychology, Pennsylvania State University, State College, PA 16801, USA, ³Department of Radiology, Michigan State University, East Lansing, MI 48824, USA, and ⁴Department of Psychology, University of Iowa, Iowa City, IA 52242, USA

Address correspondence to Katelyn M. Wills, Department of Psychology, Michigan State University, 316 Physics Road, Room 208A, East Lansing, MI 48824, USA. Email: willska2@msu.edu

Abstract

Stimulus-driven attention can improve working memory (WM) when drawn to behaviorally relevant information, but the neural mechanisms underlying this effect are unclear. The present study used functional magnetic resonance imaging (fMRI) to test competing hypotheses regarding the nature of the benefits of stimulus-driven attention to WM: that stimulus-driven attention benefits WM directly via salience detection, that stimulus-driven attention benefits WM incidentally via cognitive control mechanisms recruited to reduce interference from salient features, or that both mechanisms are co-involved in enhancing WM for salient information. To test these hypotheses, we observed activation in brain regions associated with cognitive control and salience detection. We found 2 cognitive control regions that were associated with enhanced memory for salient stimuli: a region in the right superior parietal lobule and a region in the right inferior frontal junction. No regions associated with salience detection were found to show this effect. These fMRI results support the hypothesis that benefits to WM from stimulus-driven attention occur primarily as a result of cognitive control and top-down factors rather than pure bottom-up aspects of stimulus-driven attention.

Key words: cognitive control, contingent salience, fMRI, working memory

Introduction

Working memory (WM) temporarily maintains information in an activated state, where it can be accessed and manipulated to serve current task goals. Because of its central role in human cognition and its limited capacity, it is crucial that WM is used efficiently to maintain only task-relevant information. Many theories of WM posit a central role for attention as a gatekeeper that serves to prioritize task-relevant information and exert control over the contents of WM (Baddeley 1992; Cowan 1993). Thus, an extensive body of research has focused on exploring the relationship between attention and WM. This literature has established that directing top-down or voluntary attention to

target items benefits the encoding and maintenance of task-relevant information in WM (for reviews, see Awh and Jonides 2001; Awh et al. 2006; Gazzaley and Nobre 2012).

However, it is not always the case that attention is directed in a purposeful, controlled manner toward task-relevant information. Attention can also be captured by salient, novel, or behaviorally relevant properties of a stimulus. Here, we refer to this as stimulus-driven attention—attention that is guided based on the characteristics of a stimulus. Stimulus-driven attention includes aspects of both bottom-up and top-down attention. For example, stimuli that contain novel characteristics such as a unique color or motion trajectory are likely to

interrupt the goal-directed deployment of attention in a bottom-up manner (i.e., “pure” salience; Yantis 1993). In addition, stimuli may also elicit stimulus-driven attention if they share features (e.g., color or shape) with a task goal or target (Folk et al. 2002; Serences et al. 2005). The latter scenario, frequently referred to as the “contingent” capture of attention (Folk et al. 1992), suggests that stimuli with features that are otherwise lacking novelty or uniqueness may draw attention based on the goal state or attention set of the observer. In this way, top-down attentional processes influence stimulus-driven attention.

Stimulus-driven attention elicited by pure salience from unique or novel features and contingent salience related to goal states have both been shown to enhance WM performance (Fine and Minnery 2009; Ravizza and Hazeltine 2013; Santangelo and Macaluso 2013; Gaspelin et al. 2015). Fine and Minnery (2009), for example, found improved visuospatial memory for more visually salient (e.g., color opponency) items on a map. Likewise, Santangelo and Macaluso (2013) showed that visually salient items in a naturalistic scene were prioritized at encoding, and that this prioritization was associated with enhanced memory for those items. Furthermore, this effect has also been shown in contingent salience. Ravizza and Hazeltine (2013) found that participants performing a secondary target detection task while memorizing lists of letters remembered letters matching the color of the target better than letters of other colors.

Although behavioral evidence suggests that stimulus-driven attention can benefit WM, the neural mechanisms underlying these benefits have been less frequently examined. Stimulus-driven attention mechanisms within the brain are thought to play a role in the detection of behaviorally relevant stimuli and act as a “circuit breaker” to redirect attention to salient or important events outside the current focus of attention (Corbetta and Shulman 2002). Several studies have examined the brain system supporting stimulus-driven attention, and they have identified a set of regions including the anterior insula (AI) (Hahn et al. 2006; Seeley et al. 2007; Corbetta et al. 2008; Menon and Uddin 2010), temporoparietal junction (TPJ) (Corbetta et al. 2000; Hahn et al. 2006), and the anterior cingulate cortex (ACC) (Hahn et al. 2006; Seeley et al. 2007; Menon and Uddin 2010). In this paper, we will refer to these regions as salience regions.

In visual search, greater activation within these salience regions has been associated with enhanced performance on a variety of target detection tasks. Shulman et al. (2003) found activation in the TPJ, AI, and ACC for both detected and missed targets, but this activation was greater when targets were successfully detected. Likewise, Pollmann et al. (2003) found regions in the TPJ that were specifically associated with target detection, whereas other parietal regions such as the superior parietal lobe were associated with the inhibition of distracting information. Furthermore, these regions are also thought to be sensitive to novel or salient stimuli and play a role in detecting oddball stimuli (Downar et al. 2000; Bledowski et al. 2004) and unexpected stimulus onsets (Shulman et al. 2009; Greene and Soto 2014).

Salience regions have been implicated in the successful detection of behaviorally relevant information (Corbetta and Shulman 2002; Pollmann et al. 2003; Shulman et al. 2003), but they have rarely been studied in the context of WM. Why might activity in brain regions associated with stimulus-driven attention also be associated with WM enhancement? One possibility is that the capture of attention by salient stimuli determines

what information should be prioritized during encoding (Schmidt et al. 2002; Prinzmetal et al. 2005). Behaviorally relevant information that captures stimulus-driven attention may thus be encoded first, and this earlier encoding may enhance memory for those items (i.e., primacy effects). Therefore, if salience regions are detecting behaviorally relevant information and prioritizing it, then increased activity in these regions may directly benefit WM.

On the other hand, there is evidence that would suggest that salience regions do not directly benefit WM. Santangelo et al. (2013, 2015) found that activation of salience regions such as the right TPJ was at baseline in response to salient stimuli when attention was driven by salient low-level visual features. Instead, these studies found an increase in activity in the posterior parietal cortex (PPC), a region associated with voluntary attention and cognitive control (Corbetta et al. 2000; Esterman et al. 2009). Santangelo and Macaluso (2013) proposed a network including the PPC and the medial temporal lobe that coordinates sensory cues and memory requirements in order to prioritize and encode information in WM. These findings suggest that it is not salience regions that benefit WM, but rather the recruitment of top-down control regions.

Furthermore, it is possible that cognitive control rather than salience detection processes underlie the benefits of stimulus-driven attention to WM. Salient features of an item that captures stimulus-driven attention, even when behaviorally relevant, may disrupt performance (Serences et al. 2005). When behaviorally relevant features of a stimulus compete with current task goals, behavioral salience may create a condition of increased cognitive demand. This type of conflict may occur during a conjunction visual search in which distractors share a feature such as color or shape with a target. For example, if one is looking for a red shirt in a pile of laundry, a red hat may capture their attention and potentially interfere with the goal of finding the red shirt. In other words, the task of finding the red shirt would become more cognitively demanding if other available items share target features. Therefore, cognitive control may be enhanced in the presence of salient stimuli to reduce interference from salient features and maintain task goals. This enhancement of cognitive control may then benefit WM representations incidentally via triggered top-down mechanisms (Rosner and Milliken 2014; Krebs et al. 2015; Rosner et al. 2015). In short, WM may benefit from the stimulus-driven deployment of attention because such stimuli receive additional processing related to cognitive control.

Cognitive control and WM processes are thought to play concurrent roles in supporting goal-related behavior. Harding et al. (2016) found that the brain regions that support cognitive control and WM processes overlapped. These regions include the inferior frontal junction (IFJ), dorsolateral prefrontal cortex (dlPFC), and the superior and inferior parietal lobules (SPL and IPL, respectively). Similarly, Krebs et al. (2015) found that activation of a region in the right dlPFC predicted memory benefits for faces presented during incongruent trials that required more cognitive control. They proposed that this region is part of a cognitive control network that receives conflict signals from the anterior cingulate and triggers enhanced top-down attention to overcome interference, which then enhances memory. Parietal regions whose activation also predicted memory performance were also identified in this study, and other investigators (e.g., Banich et al. 2000) have suggested that the PFC can impose biases upon parietal regions to help select relevant information.

The present study sought to distinguish between alternative hypotheses regarding the neural sources of the benefits of stimulus-driven attention to WM. The first hypothesis is that activity in salience regions benefits WM directly, perhaps by prioritizing the encoding of behaviorally relevant items captured by stimulus-driven attention. The second hypothesis is that the benefits of stimulus-driven attention to WM instead occur incidentally due to top-down processing recruited via cognitive control mechanisms to reduce interference from salient features of a stimulus. A third alternative is that these 2 proposed processes for handling salient information are not exclusive, but rather 2 complimentary mechanisms that work together to detect salient items and enhance their representations in WM.

To test these competing hypotheses, we used functional MRI and a modified version Ravizza and Hazeltine's (2013) behavioral paradigm in which participants memorized differently colored letters while performing a secondary target detection task. To test the neural and cognitive mechanisms implicated in this contingent salience paradigm, we examined activation in regions of interest (ROIs) that were defined a priori. Although the validity of this approach, known as reverse inference, has been debated, when used carefully it can provide a means for functional magnetic resonance imaging (fMRI) to advance knowledge regarding the implementation of brain regions in cognitive processes (Poldrack 2006; Hutzler 2014). Hutzler (2014) demonstrated that the careful selection of regions and the consideration of task-setting can increase the predictive power of reverse inference. Therefore, we selected a small number of ROIs based on literature implicating them in salience detection or cognitive control processes.

We expected our behavioral results to resemble the previous study in that letters matching the color of the target would be remembered better than letters of a nontarget color, which would verify the benefit of stimulus-driven attention to WM performance. Based on the hypotheses regarding the neural mechanisms for this benefit, we predicted 1 of 3 possible results for the fMRI data. First, if salience detection processes benefit WM performance, then we expected to see increased activity in regions such as TPJ, AI, and ACC associated with enhanced memory for the salient stimuli in the behavioral task. Alternatively, if the benefit to WM from stimulus-driven attention is driven by cognitive control, then we instead expected enhanced memory for salient stimuli to be associated with increased activity in control regions such as frontal and parietal areas. Finally, if these processes are co-active, we expected increased activation in both salience and cognitive control regions for salient relative to control letters.

Materials and Methods

Participants

Twenty-five right-handed undergraduate and graduate students (ages 18–30, 11 male) from Michigan State University participated voluntarily in this experiment and were compensated with \$10 per hour and images of their brain. All participants had normal or corrected-to-normal vision. Informed consent was obtained, and procedures were approved by the Human Research Protection Program at Michigan State University. All participants' behavioral data were analyzed, but 3 were excluded from imaging analyses due to excessive head motion, leaving a total of 22 participants whose data underwent full analysis.

Stimuli

The stimuli consisted of a set of 8 phonologically similar letters (B, C, D, G, P, T, V, Z) and pound signs (#), which were presented in 1 of 9 colors (red, yellow, lime, magenta, silver, teal, cyan, blue, and purple). All stimuli were displayed at the center of fixation on a black background and were back-projected into the bore of the scanner, where participants viewed them using an angled mirror attached to the head coil.

Procedure

Behavioral data were collected using E-Prime software and an MR-compatible response glove during each fMRI scan block. Each block began by presenting a reminder screen to indicate the target-colored pound sign, which remained consistent throughout the experiment. Following this screen, the participant was presented with 18 trials in a pseudo-random order. Each trial consisted of an encoding period, a jittered retention interval of either 2000, 4000, or 6000 ms, and a retrieval period, which was followed by another jittered intertrial interval of 2000, 4000, or 6000 ms. Figure 1 illustrates the experimental procedure.

In each encoding trial, 7 stimuli (6 multicolored letters and 1 pound sign) were presented serially at the center of the screen for 500 ms each, separated by fixation crosses that were also presented for 500 ms. Participants were asked to remember the letters in order while simultaneously monitoring for a pound sign of a prespecified color (red, lime, or yellow counterbalanced across participants), which could occur in any position throughout the list. Participants were instructed to respond by pressing their thumb key immediately once a pound sign in the relevant color was detected and to ignore pound signs of any other color. The target pound sign occurred twice in each block of 18 trials.

Each letter in the list was presented in a different color. In half of the trials, 1 letter was presented in the target color (salient color condition) and was either the third, fourth, fifth, or sixth letter presented. Given that letters in the target color appeared more consistently than any other individual color, the other half of the trials consisted of a control condition

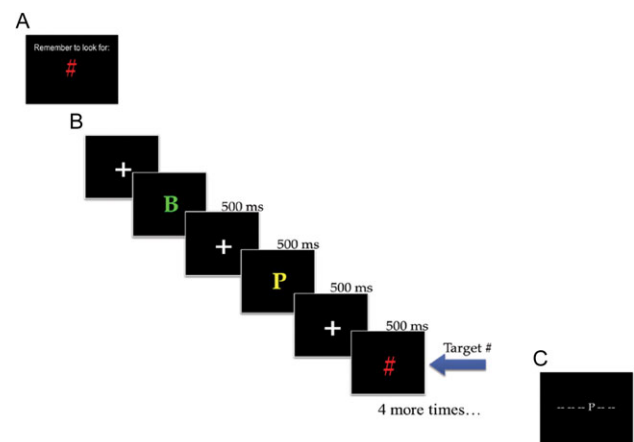


Figure 1. Example of the contingent capture paradigm used in this study. (A) Participants were assigned a target-colored pound sign. (B) Participants viewed a list of 6 randomly colored letters and 1 pound sign, responding with a keypress when the target-colored pound sign appeared while memorizing the letters. (C) Example prompt. Participants responded whether the letter on the screen was the correct letter in the correct position from the prior encoding trial.

(control color condition) in which the color of 1 letter was presented with equal frequency, but was task-irrelevant (i.e., did not match the color of the target). The control color was also chosen from red, lime, or yellow and was counterbalanced across participants. Colors of the other letters and the nonsalient pound signs in the list were chosen randomly without replacement from the color set. Therefore, 2 items of the same color occurred in only 1 trial per block (when both a salient pound sign and salient letter were presented). In all remaining trials, each item was presented in a different color.

Following the presentation of the 7 stimuli for encoding, there was a jittered retention interval, during which participants were instructed to remember the letters in the order presented. At the end of the retention period, a prompt appeared on the screen for 4000 ms, which included a single letter and 5 dashed lines. The participants' task was to respond whether or not the prompted letter was the correct letter presented in the given position during the previous encoding list. Finally, there was a jittered intertrial interval during which only a fixation cross was presented on the screen.

The prompted letter was always displayed in white font, but was a letter that had appeared during presentation in the salient, control, or a random color. The primary contrast of interest was between the salient and control trials because they were controlled for frequency of presentation and testing. Each block included 6 trials in which a salient letter was tested, 6 in which a control letter was tested, 4 in which a randomly colored letter was tested, and 2 trials, during which the target pound sign was presented at encoding, in which any 1 of the 3 letter types was tested. Because these trials required an additional motor response, they were modeled separately and excluded from subsequent statistical analysis.

Participants completed 2 practice blocks of the experiment outside the scanner followed by 8 blocks in the scanner during which fMRI echo-planar images were acquired. Excluding the salient pound trials, this resulted in acquisition of 64 encoding and 48 retrieval trials in the salient and control conditions, and 32 retrieval trials in which the letter tested was neither salient nor control.

Imaging Data Acquisition

Imaging data were acquired using a 3T GE Signa HDx scanner with an 8-channel head coil. An echo-planar imaging (EPI) sequence was used to acquire the functional data (TR = 2 s; TE = 27.7 ms; flip angle = 77°; matrix = 64 × 64; FOV = 220 mm; slice thickness = 3.4 mm; 30 slices). This protocol consisted of 8 functional runs per participant, and 190 EPI volumes were collected in each run. High-resolution T1-weighted images were also acquired using a magnetization prepared, rapid acquisition gradient echo sequence (MPRAGE; FOV = 256 mm; matrix = 256 × 256; slice thickness = 1 mm; 180 slices).

fMRI Data Preprocessing and Analysis

All fMRI data preprocessing and statistical analyses were carried out using AFNI software (Cox 1996; afni.nimh.nih.gov). Preprocessing scripts were generated using the afni_proc.py command. The fMRI volumes were first corrected for time-shifts in slice acquisition. Next, to correct for subject head motion, a 6-parameter (3 translational and 3 rotational) rigid-body transformation was used to spatially align all subsequent volumes to the first. The output from this transformation was examined to determine whether any participants moved

>4 mm or 4° in any direction. Three participants exceeded these criteria and were thus excluded from group analysis, leaving a total of 22 participants' imaging data. EPI images were then aligned to the participant's anatomical images, and then both were warped to Talairach space using a 12-parameter affine transformation to coregister images to the ICBM 452 reference brain. Functional data were spatially normalized and resampled to 2.5 × 2.5 × 2.5 mm³ resolution. Finally, data were spatially smoothed with a Gaussian filter (8 mm FWHM) and the fMRI time series was scaled to a mean of 100 to reflect percent signal change and allow for comparison between participants.

A general linear model (GLM) approach was used to estimate task-related BOLD signal change in each voxel for each participant. The model included regressors of interest for 7 different components of the task: encoding salient letters, encoding control letters, encoding salient pound signs, retrieval of salient letters, retrieval of control letters, retrieval of other colored letters, and retrieval following the presentation of a salient pound sign. For the encoding regressors, activity was modeled beginning at the presentation of the salient or control-colored letter itself rather than at the beginning of the encoding trial. BOLD responses for each condition were convolved with a canonical hemodynamic response function. The model also included 10 noise regressors: 6 of these were taken from the motion estimates output during volume registration in the preprocessing stages and were used to account for residual motion, and the remaining 4 were baseline detrending regressors used to account for constant, linear, quadratic, and cubic trends for each functional run.

ROI Analysis

To obtain independent ROIs for use in the WM task, activity during the detection of the salient pound sign was used as a localizer. This condition was chosen because it was expected to activate both salience and cognitive control regions. First, the salient pound sign as a rare and behaviorally relevant target should activate salience regions. Second, the presentation of the salient pound sign requires a behavioral "task switch" from passive memorization to perform a behavioral response. This requires an increase in cognitive control and should therefore activate these regions.

Group-mean activity for the salient pound sign condition was tested against zero using a single sample t-test, and ROIs were defined by creating a mask using a sphere with a 5-mm radius centered on the peak coordinate for selected cognitive control and salience regions that showed significant activity with an FDR-corrected threshold of $P < 0.05$. We focused on a set of regions defined a priori based on previous studies (Corbetta et al. 2000; Brass et al. 2005; Seeley et al. 2007; Esterman et al. 2009; Menon and Uddin 2010; Harding et al. 2016). Our salience regions were found in the left and right TPJ, ACC, and AI. Our cognitive control regions included the right dorsolateral PFC, left and right IFJ, IPL, and SPL. To determine whether any of our functionally defined ROIs showed statistically significant differences in activity in response to salient relative to control letters, paired sample t-tests at encoding were conducted using these averaged beta weights.

Connectivity Analysis

To explore changes in functional connectivity for salient and control letters, a beta series correlation analysis (Rissman et al. 2004)

was conducted. For this analysis, a new GLM was constructed that modeled activity for each individual trial as a separate regressor. The resulting GLM yielded separate beta weights for each trial, resulting in a total of 288 regressors of interest for each participant (18 encoding + 18 retrieval \times 8 runs). Beta weights were sorted for each condition and concatenated into 2 separate beta series for each voxel (encoding salient and encoding control). A measure of functional connectivity for each condition was obtained by correlating the average beta series of the seed region with the beta series from all other voxels in the brain. Using Fisher's (1922) r -to- z transform, the resulting correlation coefficients were transformed so that statistical conclusions could be drawn. The resulting z -scores were taken to group-level analyses in which whole-brain connectivity maps for salient and control letters were contrasted to examine the differences in functional connectivity during encoding.

Results

Behavioral Data

Accuracy for the detection of the salient pound sign was high. Participants correctly responded to the salient pound sign 92.25% of the time ($SD = 8.1\%$), with a false alarm rate of 4%.

For the WM task, accuracy in each condition (salient and control) was calculated by subtracting false alarms from hits to correct for guessing (Fig. 2). As in Ravizza and Hazeltine (2013), participants remembered letters of the salient color more accurately (69%) than letters of the control color (64%), $t(24) = 2.12$, $P = 0.04$, indicating that stimulus-driven attention benefitted WM. This was not due to a significant difference in the false alarm rate as there was no difference between the number of false alarms in the salient and control condition, $t(24) = -1.47$, $P > 0.1$. Accuracy for the randomly colored letters fell in between accuracy for salient and control letters, but was not significantly different from salient, $t(24) = 0.29$, $P > 0.10$, or control, $t(24) = -0.88$, $P > 0.10$.

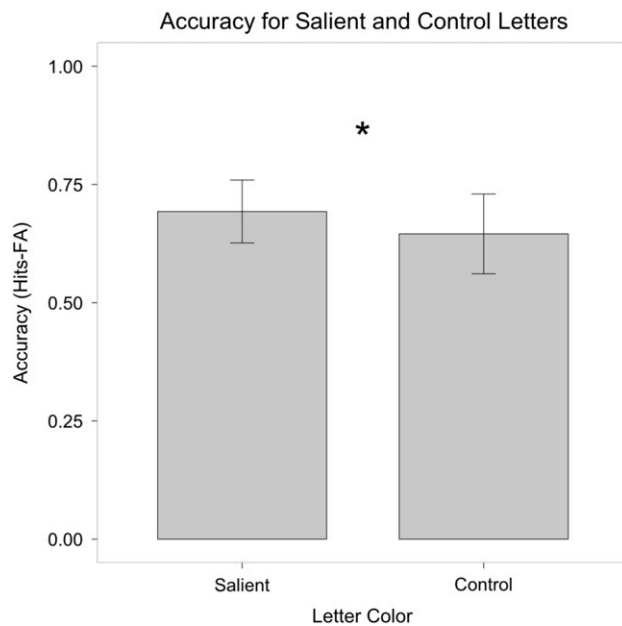


Figure 2. Recall accuracy for salient and control letters. *indicates a significant difference of $P < 0.05$.

ROI Analysis

ROIs for the WM task were determined by activity associated with the salient pound sign. The locations of the ROIs are shown in Figure 3 (see Table 1 for a list of ROI coordinates). Paired t -tests were conducted to explore differences in activity at the encoding of salient and control letters for each of these ROIs. None of the salience regions reached the statistical threshold for significance when comparing salient with control letter trials ($P > 0.05$) nor is there a trend in that direction. In fact, the right TPJ, for example, showed an overall negative response during encoding of both salient and control letters and the magnitude of deactivation was greater in the salient condition.

However, 2 cognitive control ROIs showed a significant increase in activity for salient relative to control letters (Fig. 4). One region resided in right SPL (Talairach coordinates: 29, -60, 44; BA 7), $t(21) = 2.72$, $P = 0.01$, and the other in the right IFJ (Talairach coordinates: 41, 0, 36; BA 6), $t(21) = 2.37$, $P = 0.02$, as defined by Derrfuss et al. (2005).

Correlations with Performance

The ROIs found in the right SPL and right IFJ showed an overall increase in activity for salient relative to control letters. To further examine the nature of the relationship between activity in these significant ROIs and the behavioral data, a Pearson's correlation was calculated between activity in the IFJ and SPL and task accuracy in the salient and control conditions across participants. There was a significant negative correlation between salient accuracy and right IFJ activity at the encoding of salient letters, $r = -0.56$, $P = 0.007$ (Fig. 5). Individuals who showed greater activity in the right IFJ during encoding of salient letters were more likely to have poorer WM for those letters.

No significant correlations were found for in the control condition ($r = -0.21$, $P > 0.10$), or in the right SPL region for either salient ($r = -0.32$, $P > 0.10$) or control ($r = -0.11$, $P > 0.10$) conditions. This analysis was repeated for the retrieval conditions, but no significant differences were found in any of the ROIs at retrieval for salient and control letters.

Connectivity Analysis

Using the right SPL and right IFJ as seed regions, separate beta series correlation analyses were conducted. The resulting correlation maps for salient and control letters were compared in order to examine whether some regions showed stronger connectivity with the right SPL or right IFJ during encoding in one condition or the other. This analysis yielded no significant results—there were no regions exhibiting stronger connectivity in the salient relative to the control condition.

Discussion

The present study examined the neural mechanisms supporting the benefits of stimulus-driven attention to WM. Using fMRI, we sought to distinguish between 3 alternative hypotheses: that salience detection might benefit WM by leading to prioritized encoding for salient items, that the WM benefit from stimulus-driven attention stems from cognitive control processes, or that both processes are implemented simultaneously. Our results provide evidence for the second hypothesis such that WM enhancement was due to the effects of cognitive control. Thus, a feature that evoked contingent salience was likely to engage further cognitive processing such as deploying

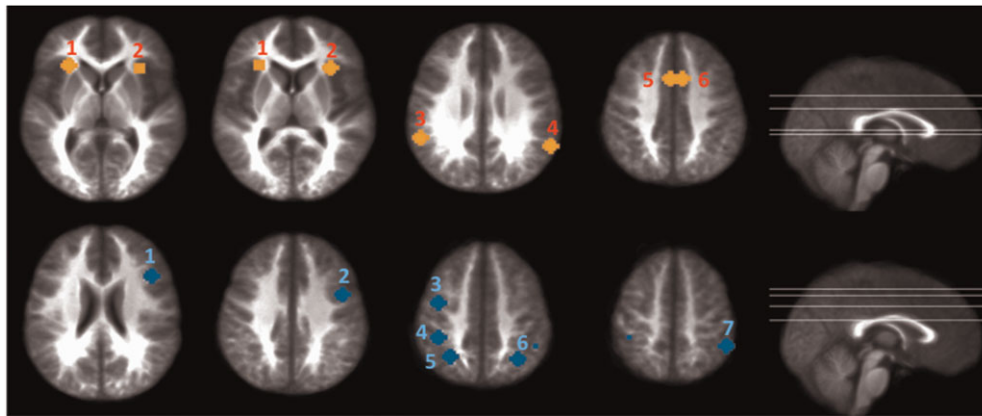


Figure 3. Regions of interest. Saliency regions (orange) included the left anterior insula (1), right anterior insula (2), left temporoparietal junction (3), right temporoparietal junction (4), left anterior cingulate cortex (5), and right anterior cingulate cortex (6). Cognitive control regions (blue) included the right dorsolateral prefrontal cortex (1), right inferior frontal junction (2), left inferior frontal junction (3), left inferior parietal lobule (4), left superior parietal lobule (5), right inferior parietal lobule (6), and right superior parietal lobule (7).

Table 1 ROI Talairach coordinates

Region		X	Y	Z	T-value
Left anterior insula	Saliency	-29	24	4	18.59
Right anterior insula	Saliency	31	21	9	18.29
Left temporoparietal junction	Saliency	-54	-43	26	9.55
Right temporoparietal junction	Saliency	55	-52	27	11.65
Right anterior cingulate cortex	Saliency	5	12	42	14.18
Left anterior cingulate cortex	Saliency	-5	12	42	10.66
Left inferior frontal junction	Cognitive control	-39	-6	44	21.22
Right inferior frontal junction	Cognitive control	41	0	36	11.24
Right dorsolateral prefrontal cortex	Cognitive control	42	18	23	10.84
Right inferior parietal lobule	Cognitive control	44	-46	51	10.68
Left inferior parietal lobule	Cognitive control	-39	-38	47	14.59
Right superior parietal lobule	Cognitive control	29	-60	44	10.70
Left superior parietal lobule	Cognitive control	-29	-56	43	10.62

Coordinates listed are the peak voxel for each ROI.

greater top-down attention to the stimulus or increased activation of task set representations. In turn, this enhanced WM for that item.

The results of the ROI analysis yielded no evidence that saliency regions in the brain are associated with better memory for the salient letters in our paradigm. Several saliency regions were activated by the salient pound sign, but none showed significant differences in activity during the encoding of salient and control letters. In fact, one of these regions, the right TPJ, tended to be suppressed in the salient letter condition. This result is similar to studies showing TPJ deactivation when distracting information is present in WM tasks (see Anticevic et al. 2010). The right TPJ was similarly deactivated here even though all items were behaviorally relevant. Furthermore, there is also evidence suggesting that saliency and cognitive control networks may compete with one another during memory tasks, and thus one may be deactivated while the other is active (Uncapher et al. 2011; Majerus et al. 2012). Our finding of deactivation in the TPJ supports this conception of the interplay between these networks during verbal WM.

On the other hand, 2 ROIs known to be involved in cognitive control did show significant differences in activity between the

2 conditions of interest. Both the right SPL and the right IFJ showed increased activity during encoding of the salient letters relative to the control letters.

The SPL is a region of the PPC and has been implicated in top-down control of attention (Coull and Frith 1998; Corbetta et al. 2000; Corbetta and Shulman 2002; Giesbrecht et al. 2003). Similar ROIs in the SPL have been found in tasks probing cognitive control and WM (Harding et al. 2016). Corbetta and Shulman (2002) also suggested a role for dorsal regions of the PPC in associating stimuli and responses within a task. Furthermore, the right SPL has been implicated in nonspatial attention and WM specifically (Coull and Frith 1998). Based on these ideas, we suggest that the SPL region in the present study is deploying top-down attention toward the task and aiding in stimulus-response mapping to ensure that the participant responds appropriately to each stimulus type (i.e., memorize the letters, respond to the salient pound sign, ignore nonsalient pound signs). Furthermore, Santangelo et al. (2015) found evidence that parietal regions help to integrate information about stimulus saliency with behavioral goals and prior knowledge. Thus, the right SPL region found in the present study could likewise play a role in integrating relevant information about the current stimulus about color, identity, and the associated

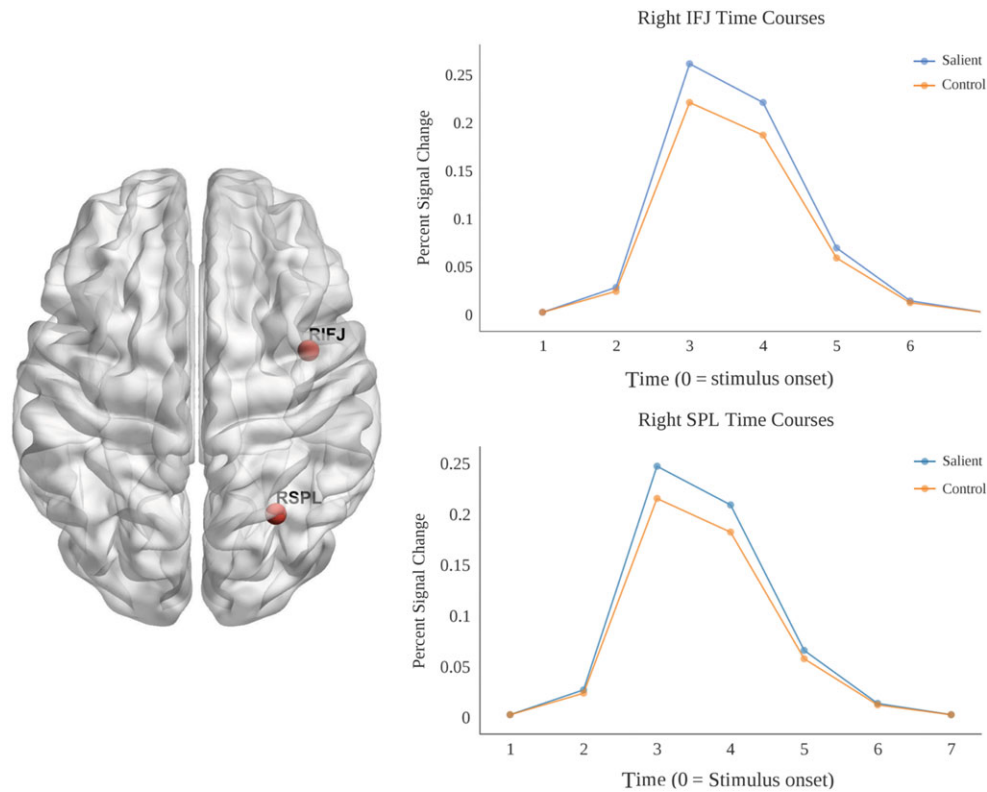


Figure 4. ROIs in the right IFJ and the right SPL and their respective time courses during encoding for salient versus control-colored letters. ROIs were visualized with the BrainNet Viewer (Xia et al. 2013, <http://www.nitrc.org/projects/bnv/>).

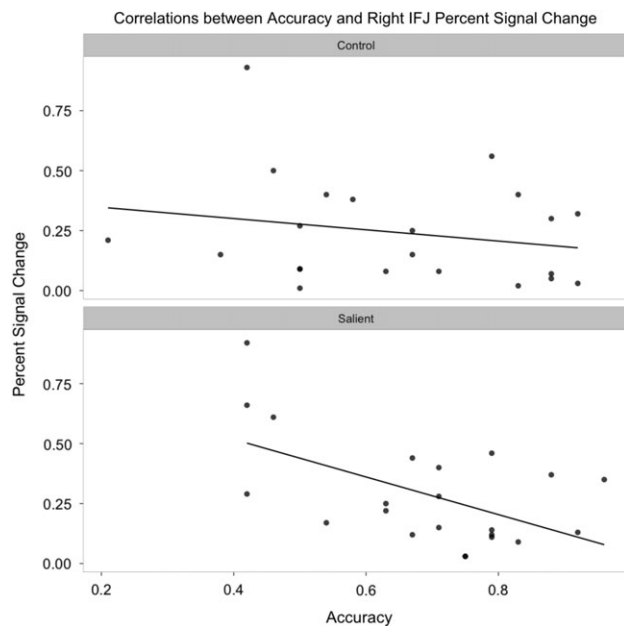


Figure 5. Scatterplot of the correlation between activity in the right IFJ and participants' accuracy during control (top) versus salient (bottom) trials.

behavioral task, which, taken together, help participants to execute appropriate behavioral responses.

The IFJ region that we observed is located at the junction between the premotor and prefrontal cortices and has been reported across several different cognitive control tasks including task and set switching, Stroop, and N-back tasks

(Brass et al. 2005; Asplund et al. 2010). Brass et al. (2005) suggest that the IFJ plays a central role in cognitive control by activating task representations. This region has also been described as a central “hub” for cognition due to its robust involvement in a variety of attentional control tasks (Cole et al. 2013). Furthermore, Asplund et al. (2010) postulate a role for the IFJ as a region that coordinates between stimulus-driven and goal-directed attention.

In the current study, the pattern of activation in the IFJ is consistent with multiple roles. First, the IFJ may serve to activate task representations and respond to distinct sets of behaviors associated with different stimuli within the task. This means that the IFJ might signal whether to memorize (letters), ignore (nontarget pound signs), or respond (salient pound signs) to the presented stimulus during encoding. Second, the IFJ might, as Asplund et al. (2010) suggest, help to coordinate the information coming from stimulus-driven attention areas as elicited by the target color and the goal-driven attention to the task. Given the diverse range of attention and cognitive control functions associated with the IFJ, it is possible that this region serves multiple functions within the present paradigm.

Paradoxically, the present study found a significant negative correlation between participants' WM performance and average BOLD response in the IFJ during the encoding of salient letters. However, when participants with an accuracy score below chance ($N = 4$) were excluded, this negative correlation was nullified, consistent with the correlation analyses for the SPL in both conditions and the IFJ in the control condition. We therefore propose that task difficulty contributes to the activation of this region during the salient condition. In future studies, it may be useful to select only high performing participants to

further tease apart the role of the IFJ in the benefits of stimulus-driven attention to WM.

The behavioral results of the current study replicate the findings of Ravizza and Hazeltine (2013) showing that when participants monitored a stream of multicolored letters for a target-colored pound sign, they remembered letters of the target color best. Our replication of this finding lends further support to the idea that contingent salience can benefit WM when items are task relevant. One remaining question, however, regards the effect of the pound sign color. Unlike the letter controls, there were no pound signs controlled for color and frequency. Furthermore, a trial never occurred in which a nontarget pound sign was the same color as another letter that could then be probed to examine whether or not the color would show a similar effect as the target pound sign. We speculate, however, that the priming effect that a pound sign color may have on the WM representation of a letter of the same color is exclusive to the target color because of its distinct importance to the task (i.e., the requirement of a response to the target color versus the instruction to ignore the remaining pound signs).

Although we found no evidence that salience detection regions were associated with more accurately recalled stimuli, given the behavioral effect showing enhanced WM for salient stimuli, we cannot rule out the possibility that information about salience is used to improve encoding and, therefore, the possibility that both salience and cognitive control mechanisms are implemented in the enhanced memory for salient stimuli. For example, activity in salience regions might be critical in triggering cognitive control regions, but then further activity of these regions may be quickly suppressed. Because our behavioral results suggest enhanced WM for salient stimuli, it is likely the case that cognitive control and top-down attention mechanisms implemented in the SPL and IFJ play a vital role in integrating and utilizing salience information for the benefit of WM. Methods such as event-related brain potential may be useful for reaching a stronger conclusion about the how these processes are coordinated in WM by providing information about the temporal characteristics of encoding salient versus nonsalient items.

Although our proposal resembles dual attention networks such as that of Corbetta and Shulman (2002), unlike Corbetta and colleagues, who focus on a top-down attention network that includes the intraparietal sulcus and frontal eye fields, we propose a larger set of frontal and parietal cognitive control regions triggered by a need to enhance the representation of task-relevant information and suppress the distracting effects of salient features. However, our account does not stand in opposition to dual network conceptualizations of bottom-up and top-down attention, and may instead lend support to these conceptualizations as a means of understanding the control of attention under various circumstances within the brain.

A strength of the present study is that it tested competing hypotheses regarding the nature of the brain mechanisms that support enhanced WM for items that elicit stimulus-driven attention via contingent salience. These results clearly favored one hypothesis over the others, namely that enhanced WM for salient items appears to be supported by more top-down control regions as opposed to salience detection regions. However, these results were obtained by relying on previous work to make inferences about the roles of various regions, and there are clear limitations to this approach that should be taken into consideration when interpreting these findings. One such limitation lies in our ability to predict the underlying cognitive processes associated with activity in a particular brain region.

In effort to overcome some of these limitations, our ROIs were carefully selected based on a large literature demonstrating their involvement in salience and cognitive control (Corbetta et al. 2000; Brass et al. 2005; Seeley et al. 2007; Esterman et al. 2009; Menon and Uddin 2010; Harding et al. 2016). We cannot rule out alternative functions for these regions, however. For example, Han and Marois (2014) include the IFJ as a region in the stimulus-driven attention network that responds to salient odd-ball stimuli. However, their interpretation of the IFJ's role in stimulus-driven attention in both attention to and evaluation of salient stimuli is consistent with cognitive control.

Conclusion

The results of our study demonstrate that stimulus-driven attention can benefit WM performance when drawn to behaviorally relevant items, and that the neural mechanisms of this effect are best understood as a result of enhanced cognitive control rather than salience detection itself. However, although we found no evidence that activity in salience regions is associated with benefits to memory for behaviorally relevant information, our behavioral results support enhanced memory for salient stimuli. Therefore, it is likely the case that salience information is taken into account during encoding, and that stimulus-driven attention and goal-directed attention must be coordinated via cognitive control in order to optimize WM performance

Funding

National Science Foundation Early Development CAREER award (#1149078 to S.R.).

Notes

Conflict of Interest: None declared.

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.
- Asplund CL, Todd JJ, Snyder AP, Marois R. 2010. A central role for the lateral prefrontal cortex in goal-directed and stimulus-driven attention. *Nat Neurosci*. 13:507–512.
- Awh E, Jonides J. 2001. Overlapping mechanisms of attention and spatial working memory. *Trends Cogn Sci*. 5:119–126.
- Awh E, Vogel EK, Oh SH. 2006. Interactions between attention and working memory. *Neuroscience*. 139:201–208.
- Baddeley A. 1992. Working memory. *Science*. 255:556–559.
- Banich MT, Milham MP, Atchley R, Cohen NJ, Webb A, Wszalek T, Kramer AF, Liang ZP, Wright A, Shenker J. et al. 2000. fMRI studies of Stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. *J Cogn Neurosci*. 12:988–1000.
- Bledowski C, Prvulovic D, Goebel R, Zanella FE, Linden DEJ. 2004. Attentional systems in target and distractor processing: a combined ERP and fMRI study. *Neuroimage*. 22:530–540.
- Brass M, Derrfuss J, Forstmann B, von Cramon DY. 2005. The role of the inferior frontal junction area in cognitive control. *Trends Cogn Sci*. 9:314–316.
- Cole MW, Reynolds JR, Power JD, Repovs G, Anticevic A, Braver TS. 2013. Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat Neurosci*. 16:1348–1355.

- Corbetta M, Kincade JM, Ollinger JM, McAvoy MP, Shulman GL. 2000. Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nat Neurosci.* 3:292–297.
- Corbetta M, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci.* 3: 201–215.
- Corbetta M, Patel G, Shulman GL. 2008. The reorienting system of the human brain: from environment to theory of mind. *Neuron.* 58:306–324.
- Coull JT, Frith CD. 1998. Differential activation of right superior parietal cortex and intraparietal sulcus by spatial and non-spatial attention. *Neuroimage.* 8:176–187.
- Cowan N. 1993. Activation, attention, and short-term memory. *Mem Cognit.* 21:162–167.
- Cox RW. 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res.* 29:162–173.
- Derrfuss J, Brass M, Neuman J, von Cramon Y. 2005. Involvement of the inferior frontal junction in cognitive control: Meta-analysis of switching and Stroop studies. *Hum Brain Mapp.* 25:22–34.
- Downar J, Crawley AP, Mikulis DJ, Davis KD. 2000. A multimodal cortical network for the detection of changes in the sensory environment. *Nat Neurosci.* 3:277–283.
- Esterman M, Chiu YC, Tamber-Rosenau B, Yantis S. 2009. Decoding cognitive control in the human parietal cortex. *Proc Natl Acad Sci U S A.* 106:17974–17979.
- Fine MS, Minnery BS. 2009. Visual salience affects performance in a working memory task. *J Neurosci.* 29:8016–8021.
- Fisher RA. 1922. On the mathematical foundations of theoretical statistics. *Philos Trans R Soc Lond.* 222:309–368.
- Folk C, Leber A, Egeth H. 2002. Made you blink! Contingent attentional capture produces a spatial blink. *Percept Psychophys.* 64:741–753.
- Folk C, Remington RW, Johnston JC. 1992. Involuntary covert orienting is contingent on attentional control settings. *J Exp Psychol Hum Percept Perform.* 18:1030–1044.
- Gaspelin N, Leonard CJ, Luck SJ. 2015. Direct evidence for active suppression of salient-but-irrelevant sensory inputs. *Psychol Sci.* 26:1740–1750.
- Gazzaley A, Nobre AC. 2012. Top-down modulation: bridging selective attention and working memory. *Trends Cogn Sci.* 16:129–135.
- Giesbrecht B, Woldorff MG, Song AW, Mangun GR. 2003. Neural mechanisms of top-down control during spatial and feature attention. *Neuroimage.* 19:496–512.
- Greene CM, Soto D. 2014. Functional connectivity between ventral and dorsal frontoparietal networks underlies stimulus-driven and working memory-driven sources of visual distraction. *Neuroimage.* 84:290–298.
- Hahn B, Ross TJ, Stein EA. 2006. Neuroanatomical dissociation between bottom-up and top-down processes of visuospatial selective attention. *Neuroimage.* 32(2):842–853.
- Han SW, Marois R. 2014. Functional fractionation of the stimulus-driven attention network. *J Neurosci.* 34:6958–6969.
- Harding IH, Harrison BJ, Breakspear M, Pantelis C, Yucel M. 2016. Cortical representations of cognitive control and working memory are dependent yet non-interacting. *Cereb Cortex.* 26:557–565.
- Hutzler F. 2014. Reverse inference is not a fallacy per se: cognitive processes can be inferred from functional imaging data. *Neuroimage.* 84:1061–1069.
- Krebs RM, Boehler CN, De Belder M, Egnér T. 2015. Neural conflict-control mechanisms improve memory for target stimuli. *Cereb Cortex.* 25:833–843.
- Majerus S, Attout L, D’Argembeau A, Degueldre C, Fias W, Maquet P, Martinez Perez T, Stawarczyk D, Salmon E, Van der Linden M, et al. 2012. Attention supports verbal short-term memory via competition between dorsal and ventral attention networks. *Cereb Cortex.* 22:1086–1097.
- Menon V, Uddin LQ. 2010. Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct.* 214:655–667.
- Poldrack RA. 2006. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn Sci.* 10:59–63.
- Pollmann S, Weidner R, Humphreys GW, Olivers CNL, Müller K, Lohmann G, Wiggins CJ, Watson DG. 2003. Separating distractor rejection and target detection in posterior parietal cortex – an event-related fMRI study of visual marking. *Neuroimage.* 18:310–323.
- Prinzmetal W, McCool C, Park S. 2005. Attention: reaction time and accuracy reveal different mechanisms. *J Exp Psychol Gen.* 134:73–92.
- Ravizza SM, Hazeltine E. 2013. The benefits of stimulus-driven attention for working memory encoding. *J Mem Lang.* 69: 384–396.
- Rissman J, Gazzaley A, D’Esposito M. 2004. Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage.* 23:752–763.
- Rosner TM, Milliken B. 2014. Context reinstatement and the desirable difficulty effect. *Can J Exp Psychol.* 68:294.
- Rosner TM, D’Angelo MC, MacLellan E, Milliken B. 2015. Selective attention and recognition: Effects of congruency on episodic learning. *Psychol Res.* 79:411–424.
- Santangelo V, Macaluso E. 2013. Visual salience improves spatial working memory via enhanced parieto-temporal functional connectivity. *J Neurosci.* 33:4110–4117.
- Santangelo V, Di Francesco SA, Mastroberardino S, Macaluso M. 2015. Parietal cortex integrates contextual and saliency signals during the encoding of natural scenes in working memory. *Hum Brain Mapp.* 36:5003–5017.
- Schmidt BK, Vogel EK, Woodman GF, Luck SJ. 2002. Voluntary and automatic attentional control of visual working memory. *Percept Psychophys.* 64:754–763.
- Seeley W, Menon V, Schatzberg A, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD. 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. *J Neurosci.* 27:2349–2356.
- Serences JT, Shomstein S, Leber AB, Golay X, Egeth HE, Yantis S. 2005. Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychol Sci.* 16: 114–122.
- Shulman GL, McAvoy MP, Cowan MC, Astafiev SV, Tansy AP, d’Avossa G, Corbetta M. 2003. Quantitative analysis of attention and detection signals during visual search. *J Neurophysiol.* 90:3384–3397.
- Shulman GL, Astafiev SV, Franke D, Pope DL, Snyder AZ, McAvoy MP, Corbetta M. 2009. Interaction of stimulus-driven reorienting and expectation in ventral and dorsal frontoparietal and basal ganglia-cortical networks. *J Neurosci.* 29:4392–4407.
- Uncapher MR, Hutchinson JB, Wagner AD. 2011. Dissociable effects of top-down and bottom-up attention during episodic encoding. *J Neurosci.* 31:12613–12628.
- Xia M, Wang J, He Y. 2013. BrainNet Viewer: a network visualization tool for human brain connectomics. *PLoS One.* 8: e68910.
- Yantis S. 1993. Stimulus-driven attentional capture. *Curr Dir Psychol Sci.* 2:156–161.