Competition in Working Memory Reduces Frontal Guidance of Visual Selection

David Soto1, Ciara M Greene1, Anum Chaudhary1 and Pia Rotshtein2

1Centre for Neuroscience, Department of Medicine, Imperial College London, London W6 8RP, UK and 2Behavioural Brain Sciences Centre, School of Psychology, University of Birmingham, Birmingham B15 2TT, UK

Address correspondence to David Soto. Email: d.soto@imperial.ac.uk.

Working memory (WM) representations can bias visual selection to matching stimuli in the field. WM biases can, however, be modulated by the level of cognitive load, with WM guidance reduced as memory load increases. Here, we used functional magnetic resonance imaging to distinguish between competing hypotheses for the reduction of WM guidance under load: 1) poor neural representations of memory contents under high load, 2) strategic control at high loads to direct attention away from search distractors matching the WM content, and 3) reduction of frontal top-down biasing of visual areas with increasing memory loads. We show that matching between WM contents and the visual display appeared to be well represented in visual areas under high memory loads, despite a lack of WM guidance at the behavioral level. There was little engagement of “cognitive control” areas in the prefrontal cortex during search at high loads. More importantly, WM guidance at low loads engaged a set of frontal regions in the superior and inferior ventral frontal cortex. Functional connectivity analyses revealed frontal regions working in concert with occipital areas at low memory loads, but this coupling was disrupted by increased memory load. We discuss the implications for understanding the mechanisms supporting the interplay between WM and attention.

Keywords: attention, prefrontal, vision, working memory

Introduction

Working memory (WM) representations can guide the deployment of attention to visually matching items in the environment (Duncan and Humphreys 1989; Folk et al. 1992; Desimone and Duncan 1995; Soto et al. 2005; Olivers et al. 2006). There has been recent controversy as to whether attentional guidance by WM content can occur automatically. Several studies have observed attentional biasing by the contents of WM, independently of their relevance for the attention task (cf. nonstrategic WM guidance; for reviews, see Soto et al. 2008; Olivers et al. forthcoming). Soto et al. (2005) asked participants to hold an object cue in WM throughout a trial; this WM cue was followed by a search task for a tilted line. On valid trials, the WM cue matched the object surrounding the critical search target, while on invalid trials the WM cue appeared in the search display surrounding a distractor. Relative to a neutral baseline, where the WM stimulus did not appear in the search display, search performance was slower in the invalid condition and faster in the valid condition, demonstrating that attention was biased to a stimulus that matched the content of WM. This pattern held even on conditions where the cue would never provide information relevant to the search, a finding that stands in contrast with studies that failed to observe nonstrategic effects of WM biases on selection (Downing and Dodds 2004; Woodman and Luck 2007).

A critical factor that may help to reconcile this apparent discrepancy across studies is the level of processing load in WM. As cognitive load increases (as in Woodman and Luck 2007, and Downing and Dodds 2004, where several items had to be kept in memory simultaneously) the strength of WM guidance of attention may decrease (see also Houtkamp and Roelfsema 2006; Olivers 2009). When the availability of WM is not taxed strong WM guidance effects are observed (for reviews and discussion on this point, see Soto and Humphreys 2008; Soto et al. 2008; Olivers et al. forthcoming). Using functional magnetic resonance imaging (fMRI), we investigated how a stressed WM may disrupt the guidance of visual selection by memory representations. Specifically, we employed the combined WM and attention paradigm from Soto et al. (2005) and varied the number of memory items that observers held in WM as they performed the visual search task. In the low-load condition, a single WM item had to be maintained, whereas in the high load, observers were required to hold 3 items in WM, one of which could reappear in the search display to interfere with selection of the target. In keeping with prior evidence that cognitive load may attenuate the expression of WM guidance (Soto and Humphreys 2008), we predicted that WM effects on selection ought to be weakest in the high-load case. From a neural perspective, we aimed to distinguish between 3 competing accounts of the absence of WM guidance under load. Specific predictions arising from each of the 3 accounts are outlined below.

One account of the absence of WM guidance under high-load contends that cognitive load may lead to increased competition between WM items and therefore to degraded memory representations as a result, thus attenuating any biasing of attention by the WM contents (Soto and Humphreys 2008). Evidence suggests that the precision of WM representations is inversely related to the degree of memory load (Bays and Husain 2008). A related account contends that WM is composed of discrete slots of fixed resolution (Zhang and Luck 2009); high levels of processing load in WM might therefore prevent the representation of some of the memory items. Either way, these accounts predict that weak neural representations or no representation of the memory item at all may explain the lack of WM guidance of the search process at high loads. Previous research (Soto et al. 2007, 2011a, 2011b) indicates that memory traces are marked by increased neural signal in occipital and frontal regions on trials where the WM item reappears in the search array relative to trials when it does not. Thus, if higher memory loads attenuate the quality of the WM trace, diminished neural reappearance effects in the presence of memory-matching items should be observed in the high relative to the low memory load.

A second possibility for the absence of WM guidance at high cognitive loads is that observers engage in strategic control and
intentionally "suppress" the WM content to direct attention away from the irrelevant WM distracters (cf. Woodman and Luck 2007; see also Han and Kim 2009). If the attenuation of WM guidance under high load is due to the implementation of cognitive control to suppress or ignore the WM distracters, increased neural activity in "control" regions of the prefrontal cortex (Miller and Cohen 2001) should be observed in the high relative to the low-load condition. We note that under low memory loads cognitive control over WM biases is not optimal and effects of WM guidance can be observed (Soto et al. 2005). However, it remains possible that cognitive control over WM biases may become more relevant at high cognitive loads. Considering that competition between memory items at high loads may degrade the representations being maintained (Bays and Husain 2008), cognitive control processes to suppress or direct attention away from the WM distracters may operate more efficiently over these impoverished representations than when a single item is held in WM. The "strategic suppression" hypothesis therefore predicts increased activity in frontal regions in the high load relative to the low memory load.

An alternative hypothesis is that the reduction in WM guidance under load is due to a reduction in frontal top-down biasing of lower level regions in the visual hierarchy. Unlike the previous proposal, this account does not require the presence of strategic control applied to the irrelevant WM contents be invoked in order to explain the reduction of WM guidance under high cognitive load. We hypothesize that frontoparietal regions may become fully engaged in WM processing at high loads (Curtis and D'Esposito 2003; Todd and Marois 2004), the consequence being that fewer neural resources may be available to bias perceptual selection in visual regions relative to when a single item is maintained in WM (for a similar argument, see Soto and Humphreys 2008). Prior research indicates that attentional guidance by WM representations may be supported by a network of frontal, subcortical (i.e., thalamus), and temporal regions, linked to early visual areas (Soto et al. 2007; Grecucci et al. 2010; Rotshtein et al. 2011; Soto, Mok, et al. 2011; Soto, Rotshtein, et al. 2011). We predict that increasing the competition between items in WM should attenuate activity in the frontal-temporal-occipital network. Furthermore, it is predicted that functional coupling between frontal and earlier visual processing regions should be higher at low memory loads—when WM guidance is maximal—than at high memory loads when WM guidance is disrupted.

### Materials and Methods

#### Participants

Twelve healthy volunteers aged between 22 and 26 years (7 females, all right handed) were recruited from the postgraduate student population at Imperial College London. None of the participants had prior history of neurological or psychiatric disorders. Local research ethics committee approval was granted for the experimental procedures. All participants provided written informed consent, had normal or corrected-to-normal visual acuity, and received a payment of £20 for their participation.

#### Experimental Procedure

Stimuli and task were controlled by E-Prime v2.0 (Psychology Software Tools Inc., Pittsburgh, PA; http://www.pstnet.com/eprime.cfm). A repeated-measures full-factorial design was used with the following factors: memory type (low WM load, high WM load, and passive priming) and the validity if the memory cue for search (valid, invalid, and neutral). A schematic of the experimental procedure is presented in Figure 1. Each trial began with a fixation cross lasting 500 ms, followed by a cue display consisting of 3 vertically aligned colored shapes on a gray background. The cue display was presented for 100 ms, followed by a brief presentation (100 ms) of a blank screen, and another 500 ms presentation of the cue display. In the high WM load condition (WM3), participants were instructed to hold the 3 colored shapes in the memory cue display in memory, whilst during the low-load condition (WM1) they were instructed to remember only the central colored shape. During a third priming condition, participants were instructed to attend to the central shape in the memory cue display in order to detect a change in either color or shape between the first and the second presentations of the 3 items held in memory. It is important that in the WM condition, the cue display was identical across the 2 presentations, while in the priming condition there were catch trials where features of the cue could vary across presentations; these catch trials were excluded from the analyses (see below). Each stimulus was unique in color and shape and was formed by the conjunction of 1 of 5 possible shapes (square, triangle, diamond, circle, or hexagon) and 1 of 5 possible colors (red, blue, green, yellow, or pink). Following the cue display, a blank screen was presented for 700 ms and was followed by the search display. Items held in memory was a group of 3 horizontally aligned colored shapes, each containing a black line which could either be tilted 26° to the right or left (the search target) or vertical (search distractors). There were 3 types of trials, in which the validity of the memory cues for search was varied. In valid trials, the target was surrounded by one of the items which had been held in memory (WM3 condition), by the single item held in WM (WM1 condition), or by the attended cue (priming condition). The matching item had the same color and shape as one of the memory cues, and the features of the other memory cues were never represented in the search display. In invalid trials, the precued object was presented but contained a distractor line instead of the search target. In neutral trials, the cued object did not appear in the search display. The 3 trial types occurred randomly and with equal probability. The search display remained onscreen for 175 ms and was followed by a 750 ms response period during which a blank screen was presented. Participants indicated the orientation of the target (right tilted or left tilted) by means of a button press.

A memory probe test followed completion of the search task. A single colored shape was presented, and observers were required to match it to the memory cue display. In responding to the memory probe, participants had to indicate whether or not the memory probe matched both the color and the shape of any of the 3 items held in memory (in the high-load case) or whether it matched the single item held in memory (in the low-load conditions). Participants had 1500 ms to respond. In the no WM load (priming) condition, participants had been instructed to monitor the 2 presentations of the cue display for a change in the color or shape of the central item in order to ensure that the cues were attended. On 25% of the priming trials (catch trials), the central item changed and participants were to withhold all further response during the remainder of the trial. These catch trials were not included in the analyses. The memory test was not presented in the priming condition; however, after completion of the search task, participants were asked to detect the location of a gap in the outline of the shape, which occurred at the top or bottom with equal probability. This was done to equate the stimulus presentation and the response requirements across the different conditions. Participants had 1500 ms to respond in the gap discrimination task. We note that the task in the priming condition had 3 components: 1) compare cues, 2) search tilted line, and 3) detect a gap in the shape contour. Hence, this condition was relatively "costly" in terms of attentional control processes and task switching. Importantly, however, no specific items were required to be held in memory across the entire trial. Responses were entered via button press. Participants were encouraged to respond as quickly and as accurately as possible in the search, memory, and gap discrimination tasks. They were also encouraged to fixate their gaze on the center of the display throughout the experiment to maximize performance and to reduce eye movements. Note that search display time was very brief (175 ms) to encourage fixation. Intertrial intervals were jittered between 1.5 and 5.5 s to facilitate the estimation of the HRF. Participants were familiarized with the task prior to scanning and were explicitly trained to avoid eye movements.
The imaging protocol followed a mixed block and event-related design with memory load varied across blocks and memory validity varied as random events within each block. Each fMRI run included 3 different blocks (WM1, WM3, priming). Four fMRI runs were conducted each lasting about 12 min, and the order of presentation of the 3 blocks within each run was randomized. There were 32 valid, 32 invalid, and 32 neutral trials on each WM1, WM3, and priming conditions (a total of 288 trials plus 24 catch trials in the no load priming condition).

**fMRI Data Acquisition**

A Siemens Verio 3-T fMRI scanner and a 32-channel head coil were used to acquire blood oxygenated level-dependent (BOLD) contrast-weighted echo-planar images during the functional scans. We acquired 38 slices in a descending order of 3 mm thickness with a distance factor of 0%, 30 ms echo time, and 2200 ms slice repetition time. In-plane resolution was 2.4 x 2.4, and the orientation of slices was set up with regard to the line between the anterior and posterior commissures (AC/PC).

**fMRI Data Analysis**

The data were analyzed using SPM8 (www.filion.ucl.ac.uk/~spm). The first 6 volumes from each session were removed to reduce the magnetic saturation effects on the BOLD response. The images were spatially realigned to correct for movement artifacts, unwrapped to correct for movement by magnetic field interactions, subsequently normalized to Montreal Neurological Institute (MNI) standard space and smoothed with an 8-mm Gaussian kernel to account for residual intersubject differences, enhance signal to noise ratio, and adhere to

---

**Figure 1.** Experimental design during WM conditions (high and low load) and priming condition.
the smoothness requirement of the random field theory (Worsley and Friston 1995).

Voxel-Based Analysis
Analysis was performed using a random-effects model. For each subject, a first-level model that included a regressor of the onsets of each trial across the 9 different experimental conditions was used. The design was 3 (task) × 3 (validity). A separate regressor was used to model the onset of the catch trials and another for error trials in the memory, gap discrimination, and search tasks. The reaction times in the search and memory tests were included as parametric modulators to control for RT differences. All the regressors were convolved with the canonical hemodynamic response function. To further account for residual signal changes because of head movement, the realignment parameters were included in the design matrix. Low-frequency harmonics (1/128 Hz) were also included in the model to account for slow wave changes typically associated with biological and scanner noise. The average estimated response across 4 sessions was computed for each subject for each condition. Consistent effects across subjects (random-effects second-level analysis) were then tested using analysis of variance (ANOVA) with 3 (low load, high load, and priming) × 3 (valid, invalid, and neutral) within subject factors. Independence and equal variance between the conditions were not assumed.

We report clusters false discovery rate (FDR) corrected for multiple comparisons (with voxel significance of at least $P < 0.01$, $Z > 2.33$), unless otherwise indicated. Further post hoc statistical analyses (tests of simple effects) were carried out on the estimated effect size of the maxima from the voxel-based analysis.

Psychophysiological Interaction Analyses
To further delineate the neural circuitry supporting WM guidance and the role of load, we computed psychophysiological interactions (PPIs), placing a seed in a region of the left ventral anterior frontal cortex that showed a validity effect. Effects of the validity manipulation were observed in proximity to this region in our previous studies (Soto et al. 2007; Grecucci et al. 2010; Soto, Mok, et al. 2011). This analysis assessed changes in the degree of functional coupling during valid trials in the WM1 (i.e., when WM guided attention toward the search target) relative to the WM3 condition (i.e., when WM guidance was attenuated). Eigenvectors were extracted within a 3 mm sphere centered on the seed region (MNI: −50, 18, 10) defined based on the maxima of the group effect (see Results). The underlying neuronal responses for these regions were estimated using deconvolution (Gitelman et al. 2003). A new model was then estimated for each subject, including a psychological regressor of our critical comparison (valid WM1 trials vs. valid WM3 trials), a physiological regressor for the time course of the region of interest, and a psychological × physiological regressor for the PPI. These 3 new regressors were added to the previous first-level model. The increase in positive or negative coupling because of head movement, the realignment parameters were included in the design matrix. Low-frequency harmonics (1/128 Hz) were also included in the model to account for slow wave changes typically associated with biological and scanner noise. The average estimated response across 4 sessions was computed for each subject for each condition. Consistent effects across subjects (random-effects second-level analysis) were then tested using analysis of variance (ANOVA) with 3 (low load, high load, and priming) × 3 (valid, invalid, and neutral) within subject factors. Independence and equal variance between the conditions were not assumed.

We report clusters false discovery rate (FDR) corrected for multiple comparisons (with voxel significance of at least $P < 0.01$, $Z > 2.33$), unless otherwise indicated. Further post hoc statistical analyses (tests of simple effects) were carried out on the estimated effect size of the maxima from the voxel-based analysis.

Results
Only trials with correct responses in both memory and search tasks were included in the analyses of the response times behavioral and neural differences across conditions cannot therefore be a reflection of differences in memory accuracy across conditions.

Behavioral Results
A 3 × 3 repeated-measures ANOVA was conducted to establish the effect of load (WM1, WM3, or priming) and cue validity (valid, invalid, or neutral) on search efficiency, defined as reaction time/accuracy (Townsend and Ashby 1983). Reaction time and accuracy data across conditions are indicated in Table 1. A significant effect of cue validity was observed ($F_{2,22} = 11.45$, $P < 0.001$) and planned contrasts indicated significant differences in search efficiency between valid and invalid trials ($F_{1,11} = 13.34$, $P = 0.004$) and between invalid and neutral trials ($F_{1,11} = 13.28$, $P = 0.004$). There was no significant difference in efficiency between valid and neutral trials ($F_{1,11} = 0.493$, $P > 0.5$). No effect of load on search efficiency was observed, although there was a significant interaction between load and search validity ($F_{1,44} = 9.875$, $P < 0.001$), such that the validity effect was strongest in the low-load WM1 condition (see Fig. 2) compared with the high-load and priming conditions. One-way ANOVAs confirmed that the validity effect was present during the WM1 condition ($F_{2,22} = 23.028$, $P < 0.001$) but not during the WM3 ($F_{2,22} = 1.694$, $P = 0.2$) or priming ($F_{2,22} = 1.416$, $P = 0.26$) conditions.

Next, we assessed the influence of load (high vs. low) and cue validity conditions upon performance in the subsequent memory test. A significant effect of load was observed ($F_{1,11} = 218.08$, $P < 0.001$), such that memory accuracy decreased from low to high load. Furthermore, the validity of the cue for search influenced memory test accuracy ($F_{2,22} = 10.055$, $P = 0.001$); post hoc contrasts showed that memory accuracy was significantly higher following both valid and invalid trials than following neutral trials ($P < 0.001$ and $P < 0.05$, respectively). There was no significant difference in accuracy between valid and invalid trials. No interaction between load and validity was observed. These data are depicted in Figure 3.

fMRI Results
Effects of Load (WM3 > WM1)
The 2 memory conditions (WM1 and WM3) differed solely with regard to the memory load requirement. A comparison of these 2 conditions is therefore the most appropriate method of examining WM load effects in this data set. The priming condition is not included in this contrast as the task requirements (perceptual comparison of successive presentation of cues, search, and gap discrimination test) are likely to have taxed the executive control mechanism and hence its relative cognitive load was assumed to be high. At a threshold of $P < 0.05$ (FDR-corrected cluster level), 3 large clusters in frontal,

Table 1
Accuracy and RTs across load and validity conditions

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>Invalid</th>
<th>Neutral</th>
<th>Valid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primed</td>
<td>0.94 (0.08)</td>
<td>0.96 (0.06)</td>
<td>0.96 (0.06)</td>
</tr>
<tr>
<td>WM1</td>
<td>0.89 (0.1)</td>
<td>0.95 (0.07)</td>
<td>0.95 (0.05)</td>
</tr>
<tr>
<td>WM3</td>
<td>0.96 (0.07)</td>
<td>0.95 (0.04)</td>
<td>0.96 (0.04)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.93</td>
<td>0.95</td>
<td>0.96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reaction time</th>
<th>Invalid</th>
<th>Neutral</th>
<th>Valid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primed</td>
<td>546.62 (50.06)</td>
<td>539.7 (57.52)</td>
<td>536 (55.72)</td>
</tr>
<tr>
<td>WM1</td>
<td>571.58 (48.83)</td>
<td>526.24 (49.2)</td>
<td>518.30 (57.42)</td>
</tr>
<tr>
<td>WM3</td>
<td>544.24 (55.97)</td>
<td>526.76 (59.24)</td>
<td>528.29 (58.27)</td>
</tr>
<tr>
<td>Mean</td>
<td>554.15</td>
<td>530.91</td>
<td>527.53</td>
</tr>
</tbody>
</table>
temporal, and occipital regions were observed which displayed increased activation during the WM3 condition relative to the WM1 condition. These clusters are listed in Table 2 and in Figure 4.

Effects of Cue Repetition
To test for item specific neural markers of the WM content, we contrasted activity when the memory cue was represented in the search display, regardless of validity (i.e., in both valid and invalid trials) relative to the neutral baseline where there was no matching between memory and search display. Following Soto et al. (2007, 2011) we tested for enhanced neural responses to repetition when the cue was held in WM relative to the priming condition, where neural repetition suppression was expected. This was tested using the following interaction t-contrast: \[\text{WM1: valid(1) + invalid(1) > baseline(-2), WM3: valid(1) + invalid(1) > baseline(-2), priming: valid(-1) + invalid(-1) < baseline(2)}\]. This pattern of interaction was observed in a large bilateral cluster in the posterior occipital cortices (Fig. 5 and Table 3) including Brodmann areas (BAs): 17, 18, 19, 23, and 31; with peak at the calcarine sulcus (MNI: -20, -100, 20) and the precuneus (MNI: 8, -50, 24). These regions showed increased response for the reappearance of the WM item in the search display relative to the neutral baseline (independent of load or validity) and showed attenuated response relative to baseline when an item was merely repeated in the priming condition. In line with previous observations (Soto et al. 2007, 2011a), this indicates the presence of an active memory trace during both WM conditions independent of memory load. More interestingly, cue reappearance effects were not lower in the high relative to the low WM load conditions (Fig. 5b). This suggests that the strength of the active memory trace alone cannot account for the observed effects of memory validity on behavior.

Effects of Cue Validity
We first assessed cue validity effects as a function of the memory requirement of the task to find neural regions showing validity effects in the WM conditions (WM1 and WM3) but not in the priming condition (WM1: valid(1) > invalid(-1), WM3: valid(1) > invalid(-1); priming: valid < invalid). Areas in the superior temporal gyrus, inferior frontal gyrus, insula, lingual gyrus, and precuneus showed validity effects that were common to both the WM1 and the WM3 conditions (Table 4a, Fig. 6a, top left). A region around the cingulate gyrus showed a validity effect in the priming condition (see Table 5). We next identified distinct regions that showed a validity effect (Valid > Invalid) separately for each condition and tested whether the observed effects were significantly different from the 2 other conditions. For example, we delineated the validity effects in WM1 and subsequently tested for validity-by-task interaction effects in the observed regions. A network involving superior frontal, inferior, and anterior prefrontal cortex (PFC) showed enhanced activity on valid relative to invalid trials in the low-load condition (WM1). The ventral anterior PFC (vaPFC) cluster showed an interaction between validity and memory type, with stronger validity effects in the WM1 than in the WM3 condition (Table 4b). An interaction effect between validity and load (WM1 vs. WM3) was detected in both the superior frontal gyrus (SFG) and the vaPFC at the peak level \((P < 0.001, \text{uncorrected, see Table 4})\). In keeping with previous studies (Soto et al. 2007; Grecucci et al. 2011), we also found validity-driven activation for the WM1 task in the left thalamus \((P < 0.001, \text{peak uncorrected; MNI: -10, -4, 8})\) and in rostral PFC \((P < 0.001, \text{uncorrected, left BA17 MNI: -34, 50, -6} \text{and right BA11 MNI: 16, 46, -12})\). Under high memory loads, validity effects on neural responses were greatly attenuated and only one significant cluster in the occipital cortex remained (Table 4) where validity-driven activation was significantly higher than in the WM1 and priming conditions. Note that the occipital cluster activated by cue reappearance was anatomically superior to this, activated by cue validity (see Figs 6 and 6a, top right).

In line with previous observations (Soto et al. 2007; Grecucci et al. 2010), there were no above threshold responses showing increased activity on invalid relative to valid trials either in the low or in the high WM load \((P < 0.05, \text{uncorrected})\).

Importantly, the results show that frontal regions typically associated with executive control processes were more involved in search validity effects during the low WM load than in the high WM load. This suggests that lack of behavioral validity effects under the high-load condition cannot be explained as increase in control processes.
PPI Results

Our third hypothesis contends that WM guidance diminishes as memory load increases due to attenuation of top-down modulation of sensory processing. To test this hypothesis, we conducted a functional connectivity analysis to provide supporting evidence for the notion that WM guidance involves the interplay between higher and lower level regions. We predicted increased functional coupling between frontal and occipital regions on valid relative to invalid trials in the low-load relative to the high-load condition. A PPI analysis was conducted using the vaPFC as seed region (MNI: $-50, 18, 10$). This region displayed WM-validity effects in prior studies (Grecucci et al. 2010). The results showed that functional coupling between the vaPFC and posterior occipital areas around the lingual gyrus during valid trials was enhanced in the WM1 case relative to the WM3 case (Fig. 6c, red blobs). The bilateral occipital cluster (cluster size = 4210 voxels, FDR and Family Wise Error corrected, $P < 0.05$) included calcarine, cuneus, lingual gyrus, and precuneus, extending into the right middle temporal cortex. To ensure that the coupling between vaPFC and posterior occipital cortex was associated with the validity manipulation, as opposed to WM load, we conducted a second PPI analysis. Here, using an identical seed we contrasted change in coupling of valid versus invalid trials only during the low-load condition (WM1). We found that coupling strength between the left vaPFC and posterior occipital

<table>
<thead>
<tr>
<th>Region</th>
<th>BA</th>
<th>Hemisphere</th>
<th>MNI coordinates</th>
<th>Cluster size (voxels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Middle occipital gyrus; cuneus</td>
<td>18, 17</td>
<td>Bilateral</td>
<td>$-18, -102, 16$</td>
<td>2676</td>
</tr>
<tr>
<td>2. Middle temporal gyrus; insula</td>
<td>21, 13</td>
<td>R</td>
<td>$44, -12, -18$</td>
<td>2748</td>
</tr>
<tr>
<td>3. Superior frontal gyrus; medial frontal gyrus; cingulate gyrus</td>
<td>6, 32</td>
<td>Bilateral</td>
<td>$8, 24, 42$</td>
<td>2527</td>
</tr>
</tbody>
</table>

Figure 4. Load effects on neural responses. Regions showing increased activation in the WM3 relative to the WM1 condition (WM3 > WM1).
cortices was larger during valid relative to invalid trials. This pattern was not observed in the high WM load case. These results, depicted in Figure 6c (blue blobs), demonstrate that coupling of frontal and posterior-occipital during valid (vs. invalid) trials was disrupted when memory load increased.

**Discussion**

Our behavioral results indicated that visual attention is guided by the contents of WM. Search efficiency was modulated by the validity of the WM contents for search. Replicating prior findings, we found that the WM bias was stronger in the low WM load condition (i.e., when a single item was maintained) relative to the high WM load condition (when 3 items were held in memory), indicating that a stressed WM can disrupt the influence of the WM content on selection (Soto and Humphreys 2008).

How does cognitive load disrupt selection of stimuli matching the WM content? One hypothesis is that the absence of WM guidance under high cognitive load is due to the imprecise nature of the memory representations being maintained at high loads (Bays and Husain 2008) or perhaps due to the absence of memory representations under load (Zhang and Luck 2009). These 2 proposals would predict insignificant neural memory traces with increased WM load, however, we found robust repetition and validity effects in occipital areas in both high and low memory loads. Independent of memory load, holding an item in WM enhanced responses in sensory cortices when this item reappeared in the search display (Fig. 5). This indicates that neural markers of active memory traces were present for both the high- and the low-load conditions. We also note that only correct memory trials were included in the behavioral and neuroimaging analyses. Although the WM...
content did not impinge upon search efficiency in the high load, matching between memory and search display did influence subsequent performance in the memory recognition test. This pattern of results can be explained if we assume the existence of a postperceptual stage that follows identification of the search target where observers may use the search display to “refresh” the WM representations to aid performance in the subsequent recognition test (cf. Woodman and Luck 2007). The behavioral pattern of memory responses mirrors the effects observed in the occipital cortex, where response increased when a WM item reappeared in the search display independent of load and validity (Fig. 3). Under this hypothesis, WM matching may not impinge on target identification processes during search itself but it may still influence the resampling of the WM content against the display. Together these results indicate that the lack of WM guidance on selection is unlikely to be the consequence of a weak memory trace.

Could the reduction in WM guidance under load be due to the application of strategic control over the irrelevant WM content in an effort to optimize target search? (cf. Woodman et al. 2001; see also Han and Kim 2009). We found little support for this proposal in the current data as no activity in regions of the prefrontal cortex associated with cognitive control (MacDonald et al. 2000; Miller and Cohen 2001; Egner and Hirsch 2005) was observed as a function of memory validity during the high-load condition. As expected an overall effect of memory load was observed in superior medial frontal cortex, the anterior cingulate, occipital, and superior temporal areas (Fig. 4). Somewhat unexpectedly, we did not find increased activity in the posterior parietal cortex with increased load (Todd and Marois 2004). We did, however, find memory load effects in lingual gyrus extending into the lateral occipital cortex, which have been previously identified as substrates of WM capacity (Xu and Chun 2006). More importantly, in line with our previous observations (Soto et al. 2007; Soto, Mok, et al. 2011), we did not find any control regions in the PFC or the parietal cortex that showed increased activity on invalid relative to valid trials; invalid trials ought to recruit the highest level of cognitive control in order to direct attention away from the irrelevant distracter and optimize search efficiency. Even in the presence of invalid WM distracters, there was no evidence of activity in these latter regions associated with cognitive control. We therefore suggest the activation found here in the anterior cingulate under high memory load reflects processing of effort (Walton et al. 2003; Rushworth et al. 2004) including WM processing (Courtney et al. 1996) rather than being associated with a cognitive control network that reacts in the presence of WM distracters.

In contrast to load effects, areas in the left SFG and the anterior inferior frontal gyrus (IFG) extending into orbitofrontal cortex (BA4/12, according to Petrides and Pandya 1994) displayed enhanced activity when the WM content surrounded the search target (on valid relative to invalid trials), but this was observed more strongly in the low load compared with the high-load or priming conditions. Increasing memory load attenuated the WM-validity effects in frontal cortex (though some validity effects of high load were observed in occipital cortex). This pattern of results suggests that the attenuation of WM guidance of selection by cognitive load is not due to the implementation of strategic control processes. In fact, we found that regions conventionally associated with control processing (e.g., the PFC) were more active in valid low memory load trials compared with high memory load trials.

Our third hypothesis contended that high memory load reduced top-down effects on sensory processing. To formally test this hypothesis, we performed a functional connectivity analysis to show how coupling with the PFC changes during the memory load and validity manipulations. We observed that functional coupling between left vaPFC and occipital visual regions when the WM matched the target of selection was higher in the low-load case—when WM guidance was maximal—relative to the high-load condition (Fig. 6c). Furthermore, the coupling between left vaPFC and occipital cortex was affected by the validity of the WM content for selection, with a higher coupling when the WM content matched the sought after target relative to when it matched a distracter. This later pattern was observed during the low load but not during the high-load condition. The joint action of the ventral PFC and posterior occipital cortex appears to be critical for WM guidance to emerge. We note that neural repetition effects for memory-matching items in the occipital cortex were strong in the both high- and low-load conditions. It might be argued that high processing loads in WM may also constrain availability of neural resources, precluding visual representations from gaining access to frontal regions. As frontal activity at high loads

---

**Table 4**

Cue validity effects on neural responses (P < 0.01, FDR-corrected cluster level)

<table>
<thead>
<tr>
<th>Region</th>
<th>BA</th>
<th>Hemisphere</th>
<th>MNI coordinates</th>
<th>Cluster size (voxels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superior temporal gyrus; inferior frontal gyrus</td>
<td>22, 48</td>
<td>L</td>
<td>−56, −38, 8</td>
<td>892</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>−60, 10, 6</td>
<td></td>
</tr>
<tr>
<td>Insula</td>
<td>48</td>
<td>R</td>
<td>34, −24, 16</td>
<td>865</td>
</tr>
<tr>
<td>Lingual gyrus; precuneus</td>
<td>18, 19</td>
<td>R</td>
<td>4, −84, −10</td>
<td>1104</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10, −50, −12</td>
<td></td>
</tr>
</tbody>
</table>

(a) Interaction between validity and memory requirement (Valid > Invalid in WM1 and WM3; Valid < Invalid in Priming)

(b) Validity effects in WM1

<table>
<thead>
<tr>
<th>Region</th>
<th>BA</th>
<th>Hemisphere</th>
<th>MNI coordinates</th>
<th>Cluster size (voxels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inferior frontal gyrus</td>
<td>44</td>
<td>L</td>
<td>−50, 18, 10</td>
<td>941</td>
</tr>
<tr>
<td>Orbitofrontal</td>
<td>47</td>
<td>L</td>
<td>−46, −28, −8</td>
<td></td>
</tr>
<tr>
<td>Superior frontal gyrus; middle frontal gyrus</td>
<td>6</td>
<td>L</td>
<td>−10, 32, 52</td>
<td>1125</td>
</tr>
<tr>
<td>Precentral gyrus; postcentral gyrus; central sulcus</td>
<td>4, 3</td>
<td>R</td>
<td>44, −24, 84</td>
<td>858</td>
</tr>
<tr>
<td>Middle temporal gyrus; cuneus; precuneus</td>
<td>39, 19, 7</td>
<td>R</td>
<td>46, −66, 34</td>
<td>1353</td>
</tr>
</tbody>
</table>

(c) Validity effects in WM3

<table>
<thead>
<tr>
<th>Region</th>
<th>BA</th>
<th>Hemisphere</th>
<th>MNI coordinates</th>
<th>Cluster size (voxels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lingual gyrus; fusiform gyrus</td>
<td>19</td>
<td>R</td>
<td>10, −50, −12</td>
<td>3119</td>
</tr>
</tbody>
</table>

---

*aClusters which survived Family Wise Error correction at the same threshold.
*bRegions showing a significant interaction between cue validity in the WM1 relative to the priming case, with higher validity effects in the WM1.
*cRegions where peak-level activity survived an interaction contrast for higher validity effects in the WM1 relative to the WM3 (P < 0.001, uncorrected).
*dSurvived an interaction contrast for higher validity effects in the WM3 relative to the WM1 and priming conditions.
may be primarily devoted to WM processing, the frontal cortex may be less available to process incoming sensory information from the search display and hence it may fail to modulate selection. In line with this proposal, WM-validity effects in frontal cortices were stronger in the low-load condition, but it is worth noting that WM-validity effects in the IFG were observed even under high memory loads (Table 4a, Fig. 6a, top left) though to a lesser extent. We propose that WM guidance of selection emerges through interactive loops involving the posterior visual cortex and the vaPFC. Visual cortex may transmit the presence of memory-matching information in the environment to the vaPFC, which may then trigger a biasing

Figure 6. Cue validity effects on neural responses. (a) Left: Interaction between cue validity effects and memory requirements (WM1 and WM3 vs. PR) on neural responses. Right: Simple effects of cue validity effects in the WM conditions. Variation in hot colors (red-yellow) illustrates WM-validity effects (Valid > Invalid) in the WM1 condition; variation in blue-green illustrates the validity effect in the WM3. Slices according to neurology convention: left is on the left. (b) Validity effect size (Valid minus Invalid). (c) PPI results. Brain regions showing increased functional coupling with vaPFC. Variation in hot colors depicts increased coupling on valid WM1 trials relative to valid WM3 trials. Variation in winter colors illustrates increased functional coupling on valid relative to invalid trials in the WM1 condition alone.
signal that boosts subsequent sensory processing toward items that match the contents held in WM. If the prefrontal cortex is "preoccupied" by maintaining multiple items in memory this interactive loop may be disrupted and WM guidance attenuated. Although our PPI analyses do not allow inferences regarding the direction of information flow throughout the brain (though see Stephan et al. 2003), we suggest that impaired attentional guidance by WM may be due to diminished top-down processing under high memory load. This argument is supported by neurophysiological evidence that top-down feedback influences from the superior and prefrontal regions to occipital areas play an important role in visual processing (Schall et al. 1995; Moore and Armstrong 2003; Zanto et al. 2011). An interesting venue for future research would be to address whether a high processing load in WM may attenuate feed-forward processing in addition to reentrant processing.

Anterior regions of the prefrontal cortex have been involved in memory encoding (Haxby et al. 1996), the control of memory representations (Badre and Wagner 2007), and explicit decision making in recognition memory (Petrides et al. 1993). It is unclear how the WM-validity effects in the anterior PFC reported here may relate to these prior findings. Other studies showed activation of the PFC in visual tasks that involve expectation of a critical target, for example, when expected and observed information coincide (Summerfield and Koechlin 2008). There are obvious commonalities between the process of holding something in memory and the process of expecting a target (for evidence that attentional set and memory maintenance can be dissociated at a behavioral level see Olivers and Eimer 2011). We speculate that WM-validity effects in anterior PFC reflect the operation of a process that is similar to the perceptual matching between internal and external signals that occurs in expectancy-based processing. It is possible that holding a feature cue in memory may generate a form of prediction signal despite not being associated with the target. In line with this, anterior PFC regions have been implicated in reward processing (for a review, see Wallis 2007), and it is thus possible that anterior PFC activity additionally reflects a reward signal when the WM content matches the sought after target. Further work ought to investigate this hypothesis further. The fact that working memory guidance in the current study was associated with enhanced functional coupling between the vPFC and occipital areas is consonant with a role of the vPFC in biasing lower level activity in visual processing pathways.

It has been proposed that a frontal-temporal-subcortical network perhaps linked to visual areas may support non-strategic WM guidance of selection (Soto et al. 2007; Grecucci et al. 2010; Rotshtein et al. 2011; Soto, Mok, et al. 2011; Soto, Rotshtein, et al. 2011). We found evidence for activity related to WM guidance in all these regions including thalamic nuclei, which is also consistent with the above prior work. More importantly, we showed that at high WM loads, validity effects in this network are greatly attenuated.

In previous papers (Soto et al. 2007, 2011; Grecucci et al. 2010), we have argued that WM guidance of visual selection involves 2 neural mechanisms. The first mechanism is sensitive to the reappearance on the memory contents in the external environment and typically displays enhanced response to the reappearance of a WM-matching stimulus in the search display relative to when there is no matching between memory and search. The second mechanism is sensitive to the relationship between the WM content and the target of search; here, increases in neural responses are observed when the WM item matches the sought after target relative to a neutral baseline without matching, and attenuation of neural response is found when the WM item matches a search distracter rather than the target. Accumulating evidence suggests that these 2 neural mechanisms are not mutually exclusive and neural responses in similar brain regions can reflect both processes. For example, SFG involvement in the sort of paradigms has been reported in prior work, responding both as a function of the mere reappearance of the WM content in search and also as a function of the its validity (Soto et al. 2007, 2011a, 2011b; Grecucci et al. 2010); more recent evidence also indicates that SFG interference by transcranial magnetic stimulation reduces the WM bias on search performance (Soto et al. 2011b). These 2 neural mechanisms may not always be dissociated from an anatomical perspective, but they may reflect different functional operations during the time course of WM guidance. It is possible that WM reappearance effects on sensory processing precede subsequent operation of WM-validity effects, but the low time resolution of fMRI precludes us from testing this hypothesis.

The absence of WM biases at high cognitive loads is consistent with the view that WM may hold several items in parallel for subsequent access, but only a single item can be the current focus of internal attention (Oberauer 2002). Under certain conditions, irrelevant memory contents might therefore remain shielded from the processes that are used to guide selection. Recent research has illustrated how this "shielding" process in WM may be accomplished as a function of task goals. When the task goal enforces the prioritization of information that is most relevant for search goals, memory information that is currently task irrelevant may not influence selection (Olivers 2009; see also Downing and Dodds 2000). Thus, it appears that WM contents may adopt different states depending on task requirements (for an overview, see Olivers et al. forthcming); the amount of processing load in WM (Soto and Humphreys 2007) or the priority of the search template over the irrelevant memory content (Downing and Dodds 2000; Houtkamp and Roelfsema 2006; Olivers 2009) will then determine whether or not WM biases attention.

According to the biased competition model of visual selection (Desimone and Duncan 1995), WM plays a critical role in biasing competitive interactions between neural representations of objects in visual cortex; the present data, however, showed occipital engagement in the presence of memory matches in the high memory load case despite there being no evidence of WM guidance. This finding indicates that WM-related activity in visual regions may not be sufficient to bias selection. Rather it appears that a fundamental feature of the cognitive state that promotes WM effects on attention is the presence of functional crosstalk between high- and low-level regions in the visual hierarchy—a state that is disrupted by stressing WM capacity. The attenuation
of WM guidance under load does not appear to be a consequence of strategic control on the part of the observer, and it cannot be merely due to imprecise memory representations. Rather, we suggest that the marriage and the divorce between memory and attention reflects a natural property of the brain which is determined by a dynamic allocation of limited resources according to the load constraints imposed by the task at hand.

**Funding**

Medical Research Council (UK, 89631).

**Notes**

Imaging data may be available for sharing upon request. We thank Chris Olivers for comments and Rebecca Quest and Adam Waldman for their help with the imaging protocol. Conflict of Interest: None declared.

**References**


