Manuscript Title

Multiple foci of spatial attention in multimodal working memory

Authors: Tobias Katus[a,b] & Martin Eimer[a]

[a] Department of Psychology, Birkbeck, University of London, London WC1E 7HX, United Kingdom.

[b] Institut für Psychologie, Universität Leipzig, 04103 Leipzig, Germany.

Corresponding Author: Tobias Katus. Department of Psychology, Birkbeck, University of London, London WC1E 7HX, United Kingdom. Mail: t.katus@bbk.ac.uk

Conflict of Interest: The authors declare no competing financial interests.

Keywords: Selective attention, multisensory (touch / vision), working memory (WM), event-related potentials (ERPs)

Abstract

The maintenance of sensory information in working memory (WM) is mediated by the attentional activation of stimulus representations that are stored in perceptual brain regions. Using event-related potentials (ERPs), we measured tactile and visual contralateral delay activity (tCDA / CDA components) in a bimodal WM task to concurrently track the attention-based maintenance of information stored in anatomically segregated (somatosensory and visual) brain areas. Participants received tactile and visual sample stimuli on both sides, and in different blocks, memorized these samples on the same side or on opposite sides. After a retention delay, memory was unpredictably tested for touch or vision. In same side blocks, tCDA and CDA components simultaneously emerged over the same hemisphere, contralateral to the memorized tactile / visual sample set. In opposite side blocks, these two components emerged over different hemispheres, but had the same sizes and onset latencies as in the same side condition. Our results reveal distinct foci of tactile and visual spatial attention that were concurrently maintained on task-relevant stimulus representations in WM. The independence of spatially-specific biasing mechanisms for tactile and visual WM content suggests that multimodal information is stored in distributed perceptual brain areas that are activated through modality-specific processes that can operate simultaneously and largely independently of each other.

.

1. Introduction

Information that is no longer physically present, but needed for ongoing behavior, is temporarily stored in working memory (WM). The neural basis of WM involves multimodal brain regions such as prefrontal cortex (PFC, Curtis & D'Esposito, 2003; Fuster & Alexander, 1971; Postle, 2006; Sreenivasan et al., 2014) and posterior parietal cortex

(PPC, Xu & Chun, 2006), as well as modality-specific perceptual brain areas (Pasternak & Greenlee, 2005; Supèr et al., 2001; Zhou & Fuster, 1996). According to the sensory recruitment model of WM (Jonides et al., 2005), cortical regions that have encoded sensory signals into WM also mediate the short-term storage of these signals. This hypothesis is supported by fMRI and EEG experiments demonstrating that stimulus-specific WM content can be decoded from neural activity in sensory cortex (Emrich et al., 2013; Harrison & Tong, 2009). Higher-level cortical areas, such as the PFC, which assert top-down influence on perceptual areas are thought to regulate the maintenance of task-relevant stimulus representations in sensory cortex (Awh & Jonides, 2001; Awh et al., 2006; Curtis & D'Esposito, 2003; Postle, 2006; Sreenivasan et al., 2014), but these higher brain regions may also play a role in information storage (Riley & Constantinidis, 2016; Romo & Salinas, 2003; Ester et al., 2015; Mendoza-Halliday et al., 2014).

The attention-based maintenance of WM representations is thought to be governed by a single supramodal control system that operates across all sensory modalities (Cowan, 2011; Cowan et al., 2011). However, this type of supramodal attentional control may be difficult to reconcile with the sensory recruitment model. If the storage of sensory information in working memory is based on the recruitment of perceptual brain areas, the maintenance of this information may also be mediated by modality-specific attentional processes. For example, tactile and visual WM representations have different spatial layouts, because they were encoded into WM by sensory neurons whose receptive fields are organized in a modality-specific fashion (somatotopic versus retinotopic; Katus et al., 2015b; Golomb et al., 2008; Golomb & Kanwisher, 2012). Hence, spatially selective processes that direct focal attention to WM content should rely on such modality-specific coordinate systems, as these index the locus where sensory information is stored in the brain. The top-down attentional control of working memory in different modalities can be

investigated in multimodal WM tasks that require the concurrent maintenance of tactile and visual stimuli. In such tasks, distinct foci of tactile and visual spatial attention may emerge simultaneously over somatosensory and visual cortex However, the hypothesis that spatially selective processes bias modality-specific (tactile/visual) WM representations simultaneously, and perhaps even independently, has so far never been tested empirically.

Previous event-related potential (ERP) studies have uncovered distinct electrophysiological correlates of the attention-based maintenance of visual and tactile WM representations. The contralateral delay activity (CDA) emerges during the retention of visual stimuli over posterior visual areas contralateral to the visual field in which memorized items had been presented (Vogel et al., 2005; Vogel & Machizawa, 2004). The CDA is sensitive to WM load and individual differences in WM capacity, and reflects the spatially selective maintenance of information in visual WM. The tactile CDA component (tCDA) shows a similar response profile as its visual counterpart, but has a modality-specific topography over contralateral somatosensory cortex (Katus & Eimer, 2015; Katus et al., 2015a; Katus & Müller, 2016; for further discussion of the relationship between the tCDA and the somatotopic organization of tactile WM, see Katus et al., 2015b). So far, the CDA and tCDA components have been investigated exclusively with unimodal (visual or tactile) WM tasks. For the first time, we here concurrently measured the tCDA and CDA components in a bimodal WM task to track the maintenance of tactile and visual WM representations simultaneously. To distinguish between the tCDA and CDA, we used current source density (CSD) transforms (Tenke & Kayser, 2012), which minimize volume conduction effects between these components. Note that both the tactile and visual CDA are inherently spatially selective markers of WM maintenance, because these lateralized components are isolated by subtracting ipsilateral from contralateral ERPs (as defined relative to the side where stimuli are memorized). We therefore employed a spatial

manipulation to examine whether the spatially selective biasing of tactile and visual WM representations is mediated by dissociable processes.

Bimodal (tactile/visual) sample sets were simultaneously presented on the left and right sides (Figure 1). Participants memorized the locations of two tactile stimuli and the colors of two visual stimuli, before memory was unpredictably tested for vision or touch. The location where the task-relevant visual and tactile sample stimuli had to be retained alternated across experimental blocks. In half of all blocks, participants memorized tactile and visual stimuli on opposite sides (touch left / vision right, or vice versa). In the other half, their task was to memorize tactile and visual stimuli on the same side. If distinct spatially selective biasing mechanisms maintain focal attention on tactile and visual memory representations, the tCDA and CDA components should emerge over opposite hemispheres in opposite sides blocks, whereas in same sides blocks, both components should manifest over the same hemisphere. The tCDA/CDA components should be statistically reliable (as indexed by amplitudes that differ from zero), and importantly, the polarities of these components should differ between same and opposite sides blocks. Such a pattern of results would strongly support the hypothesis that separate spatially selective biasing mechanisms maintain focal attention on stimulus representations that were encoded into WM through different modalities.

2. Materials and Methods

2.1. Participants

Twenty neurologically unimpaired paid adult participants took part in the experiment. One participant was excluded due to poor behavioral performance (memory accuracy for tactile stimuli was below 60%), another because of excessive alpha activity. The remaining

eighteen participants (mean age 29 years, range 19-42 years, 11 female, 17 right-handed) all had normal or corrected vision. The study was conducted in accordance with the Declaration of Helsinki, and was approved by the Psychology Ethics Committee, Birkbeck College. All participants gave informed written consent prior to testing.

2.2. Stimulation hardware and stimulus materials

Participants were seated in a dimly lit recording chamber with their hands covered from sight. Tactile stimuli were presented by eight mechanical stimulators that were attached to the left and right hands' distal phalanges of the index, middle, ring and small fingers. The stimulators were driven by custom-built amplifiers, controlled by MATLAB routines (The MathWorks, Natick, USA) via an eight-channel sound card (M-Audio, Delta 1010LT). Tactile stimuli were presented in sets of four simultaneous pulses (two to each hand), consisting of 100 Hz sinusoids that were presented for 150 ms with an intensity of 0.37 N. Headphones presented continuous white noise to mask any sounds produced by tactile stimulation.

Visual stimuli were shown for 150 ms at a viewing distance of 100 cm against a black background on a 22 inch monitor (Samsung SyncMaster 2233; 100 Hz refresh rate, 16 ms response time). Four differently colored squares were presented simultaneously (one in each quadrant). Each square had a size of 0.63° of visual angle, and all squares were equidistant from central fixation, with a horizontal eccentricity of 0.64° and a vertical eccentricity of 053° of visual angle (measured relative to the squares' centers). Six equiluminant colors (11.8 cd/m^2) were used in the experiment (red, green, blue, yellow, cyan and magenta). A white fixation dot was constantly present on the screen centre throughout the experiment. At the end of each trial, a question mark was shown centrally for 2000 ms to indicate the response period.

2.3. Stimulation procedure and task

We used a bimodal WM procedure that combined two lateralized change detection tasks for tactile and visual stimuli. Figure 1 illustrates the stimulation procedure. Bimodal (tactile and visual) sample sets were followed after 1500 ms by a unimodal test set (tactile or visual, 50%). The sample sets included two visual stimulus pairs on the left and right side of the monitor, and two tactile stimulus pairs, presented simultaneously to the left and right hands. Participants were instructed to memorize visual and tactile stimulus pairs on one task-relevant side, and to decide whether the (tactile or visual) test stimulus set matched the memorized sample set on the respective task-relevant side. In different blocks, tactile and visual stimuli had to be retained on the *same side* (e.g., memorize visual stimuli on the left side, and tactile stimuli on the left hand), or on *opposite sides* (e.g., visual stimuli on the left side and tactile stimuli on the right hand).

On each trial, two stimulators were randomly and independently selected on each hand to deliver the tactile sample pulses. On those trials where memory was tested for touch after the retention period, the locations of the tactile test stimulus set on the task-relevant hand were either identical to the sample set's locations (match trials, 50%) or differed (mismatch trials, 50%). In two thirds of all mismatch trials, test stimulus pairs were delivered to one previously stimulated location and one new location (where no sample had been presented). In the remaining third of mismatch trials, both test stimuli were presented to new locations. On the task-irrelevant hand, test stimuli were also presented at matching or mismatching locations, independent of whether there was a match or mismatch on the task-relevant hand. Visual sample sets consisted of two squares on the left side and two squares on the right side in four randomly selected colors. On those trials where visual memory was tested, the visual test set was either identical to the sample set on the task-relevant side (match trials, 50%) or differed (mismatch, 50%). In two thirds of all mismatch

trials, one of the two colors changed across sample and test. In the remaining third of mismatch trials, the task-relevant colored squares in the sample set swapped their locations in the test set. Visual test stimuli on the task-irrelevant side could also match or mismatch the sample set on this side, independently of whether there was a match or mismatch on the relevant side.

Since memory was unpredictably tested for touch or vision, participants had to memorize task-relevant tactile and visual stimuli on each trial. They signalled a match or mismatch between sample and test on the relevant hand / side with a vocal response (“a” for match and “e” for mismatch) that was recorded with a headset microphone. A question mark shown on the monitor for 2000 ms indicated the response period, which started 360 ms after test stimulus onset. The interval between the end of the response period and the start of the next trial varied between 720 and 980 ms (average 850 ms). The experiment involved 528 trials, presented during twelve blocks with 44 trials each. The relevant side for the visual task changed after every three blocks, and the relevant side for the tactile task after six blocks. Task instructions specifying the relevant locations for the visual and tactile tasks were shown on the monitor prior to the start of each block. Participants were asked to avoid head and arm movements, to maintain central gaze fixation, and to prioritize accuracy over speed. Feedback on hit and correct rejection rates was provided after each block. Half of the participants performed the same side condition during the first three blocks and during the last three blocks of the experiment. The remaining participants performed the opposite side condition during these blocks (and the same side condition in blocks four to nine). Before the experiment, participants completed training blocks of 25 trials for the same side as well as opposite sides condition.

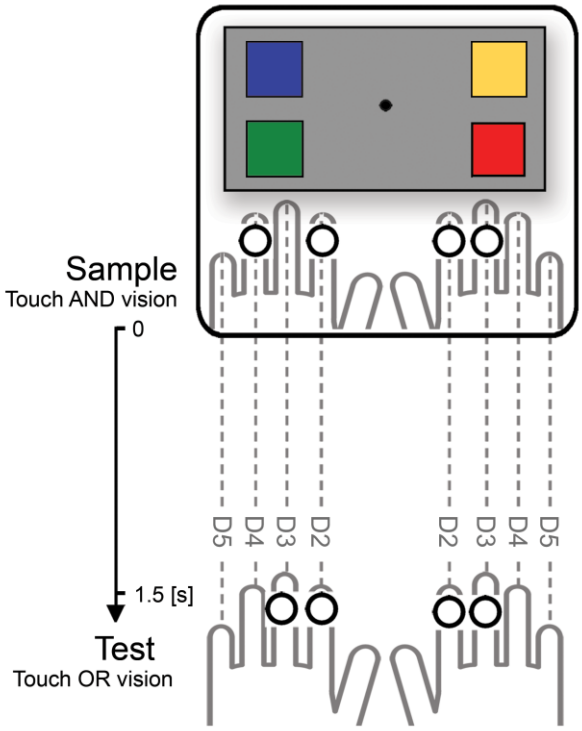

Figure 1. Stimulation procedure and task. A bimodal (tactile-visual) sample set was followed after 1.5 s by a unimodal test set (unpredictably tactile or visual). The locations of the tactile sample stimuli (indicated by circles) were memorized on one task-relevant hand (left or right), and the colors of the visual stimuli were memorized in one visual field (left or right). In *same side* blocks, tactile and visual sample stimuli were memorized on the same side. In *opposite side* blocks, participants memorized tactile samples on the left hand and visual samples on the right side, or vice versa. In each trial participants reported a match or mismatch between sample and test sets (on the task-relevant hand/side).

2.4. Analysis of EEG data

EEG data, sampled at 500 Hz using a BrainVision amplifier, were DC-recorded from 64 Ag/AgCl active electrodes at standard locations of the extended 10-20 system. Two electrodes at the outer canthi of the eyes monitored lateral eye movements (horizontal electrooculogram, HEOG). Continuous EEG data were referenced to the left mastoid during

recording, and were offline re-referenced to the arithmetic mean of both mastoids. Data were submitted to a 30 Hz low-pass finite impulse response filter (Blackman window, filter order 500). Epochs were extracted for the 1500 ms interval after presentation of the sample sets, and were corrected relative to 200 ms pre-stimulus baselines.

Blind source separation of EEG data was performed using the independent component analysis (ICA) algorithm implemented in the EEGLab toolbox (Delorme & Makeig, 2004; Delorme et al., 2007). Independent components (ICs) accounting for eye blinks were subtracted from the data. Epochs with lateral eye movements were identified and rejected using a differential step function that ran on the bipolarized HEOG (step width 100 ms, threshold 30 µV). After exclusion of trials with saccades, we additionally subtracted ICs accounting for horizontal eye movements, to remove residual traces of ocular artifacts that had not exceeded the amplitude threshold of the step function. Because slow lateralized drifts caused by head or body movements can compromise the analysis of sustained lateralized ERP components, epochs with such drifts were identified and rejected in two steps. First, 27 difference waves were computed per trial by calculating the difference between ERPs at corresponding left- and right-hemispheric electrodes (e.g., C3 minus C4) within the time window used for the subsequent ERP analyses (300-1500 ms after sample onset). Epochs that contained difference values exceeding a threshold of +/- 50 µV were rejected. In a second step, we converted single-trial EEG data to current source densities (CSDs) before calculating difference waves for the 27 lateral electrode pairs. Difference values in the time window of interest (300-1500 ms) were standardized across trials via z-transformations. Trials in which at least two electrode pairs showed z-scores exceeding a threshold of +/- 3 were rejected. Note that this procedure was only used to identify epochs with artifacts - the z-scores obtained from CSD-transformed data were not used for statistical analysis. All remaining EEG epochs were submitted to *Fully Automated*

Statistical Thresholding for EEG Artifact Rejection (FASTER, Nolan et al., 2010), and were subsequently converted to CSDs (iterations = 50, m = 4, lambda = 10^{-5}; see Tenke & Kayser, 2012) to minimize effects of volume conduction between the tCDA and CDA components. After artifact rejection, 91.4% of all epochs remained for statistical analysis (same side: 91.5%; opposite sides: 91.3%). These epochs were averaged separately for same side and opposite sides blocks.

EEG data from pairs of three adjacent electrodes were averaged, separately for the hemisphere contralateral and ipsilateral to the currently relevant side for the visual and tactile tasks. Tactile contralateral delay activity (tCDA component) was measured at lateral central scalp regions (C3/4, FC3/4, CP3/4). Visual contralateral delay activity (CDA) was measured at lateral occipital scalp regions (PO7/8, PO3/4, O1/2). Statistical analyses were conducted on CSD amplitudes averaged between 300 ms and 1500 ms relative to sample onset (cf., Katus et al., 2015a).

Error bars in graphs showing contra- / ipsilateral difference values indicate 95% confidence intervals, which were calculated for each condition by t-tests against zero (i.e., no lateralized effect). Statistical significance of difference values is marked by error bars (or colored shadings in CSD plots) that do not overlap with the zero axis (i.e., $y \neq 0$). Topographic voltage maps display spline-interpolated difference values that were obtained by subtracting CSDs ipsilateral to the visual task from contralateral CSDs. The resulting difference values were mirrored to the opposite hemisphere, to obtain symmetrical but inverse voltage values for both hemispheres. As data in these maps are aligned to illustrate lateralized effects for visual sample stimuli that are memorized on the right side, these maps differ as to whether tactile sample stimuli are memorized on the right hand (same side condition) versus left hand (opposite sides condition).

2.5. Statistical analyses

The F- and t-statistics reported in the manuscript were obtained from repeated measures ANOVAs and t-tests. Effect sizes are quantified by partial eta² values (η^2_p) in ANOVAs and by Cohen's d in t-tests. For the jackknife-based procedure (Miller et al., 1998) employed to compare onset latencies of the tCDA and CDA components between same side and opposite sides blocks, we used one-way ANOVAs, with corrected F- and partial eta² values ($F_{corrected}$, $\eta^2_{pcorrected}$), according to Miller et al., 1998 and Ulrich & Miller, 2001.

Because non-significant effects cannot be easily interpreted in the context of conventional null-hypothesis significance testing, we additionally calculated Bayes factors (Wagenmakers et al., 2010; Rouder et al., 2012; Rouder et al., 2009) using the software JASP (JASP team, 2016). The Bayes factor for the null-hypothesis (BF_{01}) denotes the relative evidence in the data supporting the null-hypothesis, as compared with the alternative hypothesis, and corresponds to the inverse of the Bayes factor for the alternative hypothesis (BF_{10}). Depending on whether an effect was statistically significant or non-significant, we here report the Bayes factor for the alternative (BF_{10}) or null-hypothesis (BF_{01}), respectively. Reliable evidence for either hypothesis is indexed by a BF > 3 (Jeffreys, 1961), suggesting that the empirical data is at least 3 times more likely under this hypothesis as compared with the competing hypothesis.

3. Results

3.1. Behavioral performance

Participants responded correctly on 91.1% of all trials. The percentage of correct responses and mean reaction times (RTs) were virtually identical in same side and opposite sides blocks (91.0% versus 91.1%; 871 ms versus 863 ms). Full factorial ANOVAs examined whether RTs and memory accuracy (d') were influenced by the factors *attended sides*

(same vs. opposite) and *tested modality* (touch vs. vision). RTs were significantly faster on trials in which visual WM was tested (815 ms versus 918 ms when touch was tested; $F(1, 17) = 23.091$, $p < 0.001$, $\eta^2_p = 0.576$, $BF_{10} = 180.959$), but accuracy was not significantly increased on these trials (d' = 3.2 versus 2.8; $F(1, 17) = 3.347$, $p = 0.085$, $\eta^2_p = 0.164$, $BF_{01} = 1.040$). The factor *attended sides* did neither influence RTs ($F(1, 17) = 0.463$, $p = 0.505$, $\eta^2_p = 0.027$, $BF_{01} = 3.350$) nor memory accuracy ($F(1, 17) = 0.220$, $p = 0.645$, $\eta^2_p = 0.013$, $BF_{01} = 3.729$), and no significant interactions were found between *attended sides* and *tested modality* (RTs: $F(1, 17) = 1.280$, $p = 0.274$, $\eta^2_p = 0.070$, $BF_{01} = 2.362$; Accuracy: $F(1, 17) = 0.001$, $p = 0.971$, $\eta^2_p = 0.000$, $BF_{01} = 4.112$).

3.2. Event-related potentials

Lateralized effects were present in CSDs recorded at visual and somatosensory regions of interest (ROIs), both in same side and opposite sides blocks; see Figure 2. The visual CDA component was found contralateral to the side where visual stimuli were memorized. The polarity of the somatosensory tCDA component (defined relative to the task-relevant side for the visual task) reversed between blocks where both tasks were performed on the same side as opposed to opposite sides. This tCDA polarity reversal is displayed in the CSDs and difference waves in Figure 2, as well as in the topographical maps, which show tCDA and CDA components over lateral central and posterior regions of the same hemisphere in same side blocks, and over opposite hemispheres in opposite sides blocks.

Statistical analyses were conducted on CSD amplitudes that were averaged for the time period between 300 and 1500 ms after the sample set. A three-way repeated measures ANOVA with the factors *attended sides*, *ROI* and *contralaterality* (now defined independently for tactile and visual ROIs relative to the task-relevant hand and the task-

relevant visual field, respectively) assessed contralateral and ipsilateral CSDs at somatosensory and visual ROIs in same side and opposite sides blocks. Contralateral CSD amplitudes were more negative than CSDs measured ipsilateral to the task-relevant hand / side, as reflected by a significant main effect of contralaterality ($F(1, 17) = 58.782$, $p < 10^{-6}$, $\eta^2_p = 0.776$, $BF_{10} > 10^4$). Lateralized effects were more pronounced over visual as compared to tactile ROIs (contralaterality x ROI interaction: $F(1,17) = 29.949$, $p < 10^{-4}$, $\eta^2_p = 0.638$, $BF_{10} = 619.679$), and this result suggests that the visual CDA component was larger in size than its somatosensory counterpart. No further main effects or interactions were statistically significant (all ps > 0.1). Note that the absence of a significant interaction between the factors contralaterality and attended sides ($F(1, 17) = 0.000$, $p = .984$, $\eta^2_p = 0.000$, $BF_{01} = 4.114$) implies that tCDA and CDA components had similar sizes in blocks of the same side and opposite sides conditions (see bar graphs in Figure 2). Importantly, t-tests against zero confirmed that the simultaneously elicited tCDA / CDA components were statistically reliable in same side blocks (tCDA: $t(17) = 3.117$, $p = 0.006$, $d = 0.735$, $BF_{10} = 7.796$; CDA: $t(17) = 6.527$, $p < 10^{-4}$, $d = 1.538$, $BF_{10} > 10^3$), as well as in opposite sides blocks (tCDA: $t(17) = 4.211$, $p = 0.001$, $d = 0.992$, $BF_{10} = 59.313$; CDA: $t(17) = 6.668$, $p < 10^{-4}$, $d = 1.572$, $BF_{10} > 10^3$). The difference waveforms in Figure 2 suggest that there were no systematic differences in the onset of lateralized components over somatosensory and visual cortex between same side and opposite sides blocks. To test this formally, we submitted contra-/ipsilateral difference waveforms to a jackknife-based procedure (Miller et al., 1998). Onset latencies were defined as the point in time where amplitudes of tCDA and CDA difference waveforms exceeded an absolute criterion of -0.1 mA/m³. There were no significant differences of tCDA / CDA onset latencies between same side and opposite sides blocks (tCDA: $F_{corrected}(1, 17) = 0.371$, $p = 0.551$, $\eta^2_{pcorrected} = 0.021$, $BF_{01} = 3.489$; CDA: $F_{corrected}(1, 17) = 0.368$, $p = 0.552$, $\eta^2_{pcorrected} = 0.021$, $BF_{01} = 3.494$), indicating that

WM maintenance was not delayed when tactile and visual samples were memorized on opposite sides.

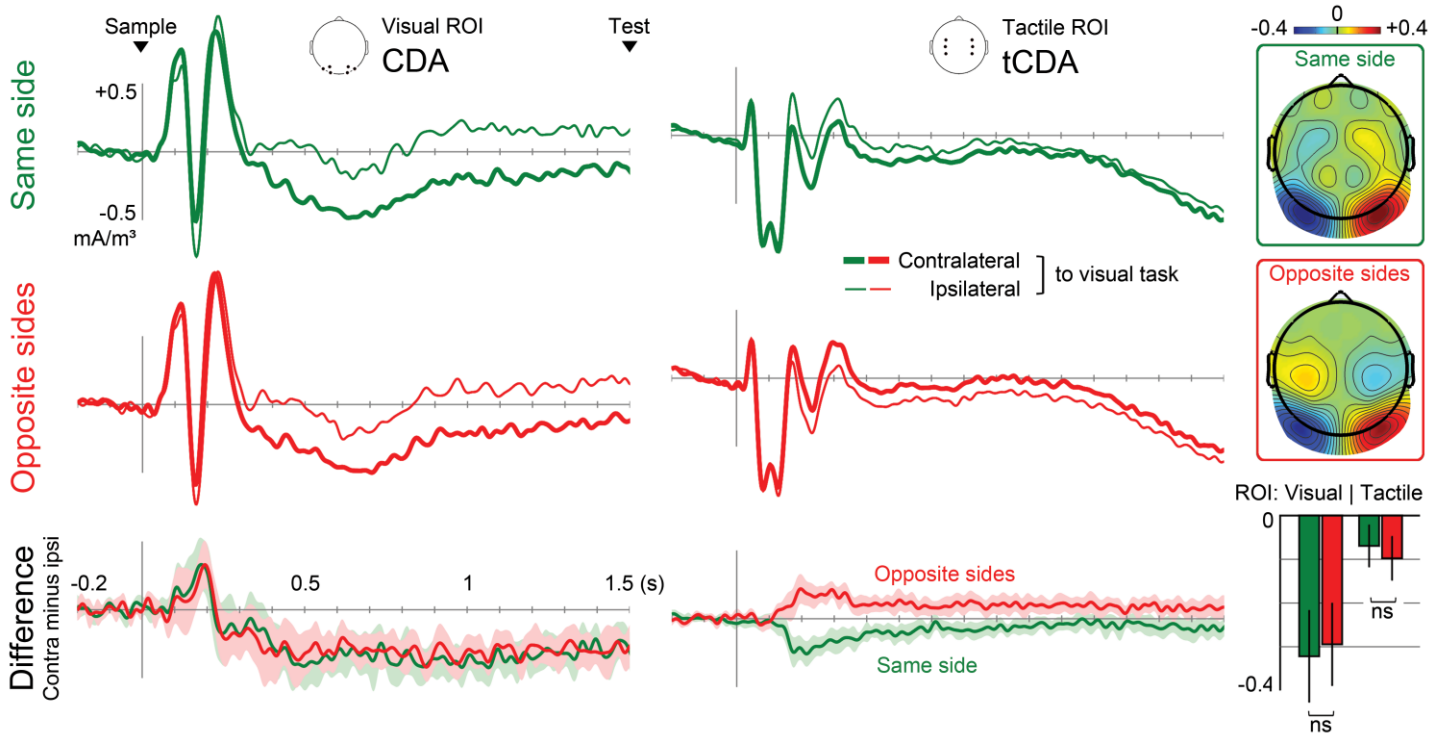

Figure 2. Lateralized delay activity. Grand mean CSD-transformed ERPs evoked by the bimodal sample set in blocks where tactile and visual stimuli were memorized on the same side (green) and on opposite sides (red). Results are shown for lateral visual (CDA component) and somatosensory (tCDA component) regions of interest (ROIs). Contralateral and ipsilateral electrodes (thick versus thin lines) were defined relative to the task-relevant side for the visual WM task. The bottom panel shows contra- minus ipsilateral difference waveforms. Shaded areas represent 95% confidence intervals (CIs) for tests of difference values against zero (i.e. no lateralized effect). Topographical maps show the scalp distribution of spline-interpolated difference values obtained by subtracting ipsilateral from contralateral mean amplitude values between 300 - 1500 ms after sample onset. Notably, tCDA and CDA components were triggered over the same hemisphere in same side blocks,

and over opposite hemispheres in opposite sides blocks. Bar graphs (bottom right) show mean amplitudes of lateralized components between 300 and 1500 ms after sample onset for visual and somatosensory ROIs, in same side (green) and opposite sides (red) blocks, with laterality now defined relative to the task-relevant side in each task (i.e. relative to the visual task for visual ROIs, and tactile task for tactile ROIs). Statistically reliable lateralized effects are marked by error bars that do not overlap the zero line ($y \neq 0$).

3.3. Behavioral control experiment

The absence of behavioral costs in opposite sides relative to same sides blocks in the main experiment may indicate that the demands of the task were too low. This could have resulted in ceiling effects that may have obscured potential performance costs when tactile and visual stimuli had to be maintained on opposite sides. To assess this possibility, we conducted an additional behavioral control experiment that used the same procedures as the main experiment, except that visual WM load was doubled from 2 to 4. Thus, participants had to memorize 6 simultaneously presented stimuli (2 tactile plus 4 visual stimuli), exceeding the suggested WM capacity limit of 4 items (Cowan, 2001), which is assumed to apply even when these items are encoded through different sensory modalities (Cowan, 2011).

On each side of the monitor, two visual stimuli that appeared at the same locations as in the main experiment (horizontal and vertical eccentricity relative to the fixation cross: 0.64° and 053° of visual angle) were accompanied by two additional stimuli (horizontal and vertical eccentricity: 1.17° and 0.53°). In visual mismatch trials, one randomly selected sample stimulus changed its color at memory test. Memory was again unpredictably tested

for touch or vision (50% each), and memory matches and mismatches (50% each) were equally likely for the task-relevant and -irrelevant sides.

13 volunteers participated in the control experiment. One participant was excluded due to chance performance in the tactile task. The remaining 12 participants (mean age 30 years, range 21-42 years, 6 female, 9 right-handed) responded correctly on 85.3% of all trials (tactile task: 90.8% correct, visual task: 79.9% correct). Importantly, and analogous to the main experiment, accuracy was not impaired in opposite sides blocks (opposite vs. same sides: 85.8% vs. 84.9% correct). A formal ANOVA tested memory accuracy (d') for the factors *attended sides* (same vs. opposite) and *tested modality* (touch vs. vision). This analysis confirmed that memory performance did not differ in same sides and opposite sides blocks (*attended sides*: $F(1, 11) = 0.194$, $p = 0.668$, $BF_{01} = 3.199$). Accuracy was higher for the tactile as compared to visual task (*tested modality*: $F(1, 11) = 16.823$, $p = 0.002$, $BF_{10} = 24.940$), but there was no reliable interaction (*attended sides* x *tested modality*: $F(1, 11) = 0.503$, $p = 0.493$, $BF_{01} = 2.290$).

4. Discussion

The current experiment has demonstrated for the first time that the attentional activation of information stored in somatosensory and visual brain areas is mediated by distinct spatially selective processes. Observers simultaneously maintained task-relevant visual and tactile sample stimuli for a subsequent comparison with a test stimulus set. The concurrent attentional maintenance of tactile and visual WM representations was reflected by lateralized tCDA and CDA components with modality-specific topographies. When observers memorized tactile and visual stimuli on the same side, statistically reliable tCDA

and CDA components emerged over somatosensory and visual cortex within the same hemisphere, contralateral to the task-relevant stimuli. This finding shows that tactile and visual WM representations can be activated simultaneously in anatomically segregated brain regions, and demonstrates the feasibility of our concurrent tCDA/CDA measurement approach. Even stronger evidence for a dissociation between tactile and visual WM maintenance processes was obtained when tactile and visual stimuli were memorized on opposite sides, resulting in tCDA and CDA components that were simultaneously elicited over different hemispheres (see topographical maps in Figure 2). This result reveals distinct foci of tactile and visual spatial attention, and leads to the conclusion that spatial attention operates in a modality-specific fashion during the maintenance of multimodal WM representations. In spite of the reversed polarity of the tCDA and CDA components in opposite side blocks, their absolute amplitudes and onset latencies did not differ between opposite sides and same side blocks. This observation further bolsters the interpretation that the spatially selective activation of tactile and visual information is mediated by separate modality-specific processes which operate within the same perceptual systems that have accomplished the storage of information in WM.

Lateralized ERP components elicited during the delay period of WM tasks mark the spatially selective allocation of attention to WM representations that are stored in perceptual brain regions. Top-down control signals generated in multimodal areas, such as PFC and/or PPC, regulate the maintenance of information in WM by biasing neural activity in sensory cortex in a task-dependent fashion (Curtis & D'Esposito, 2003; Jonides et al., 2005; Postle, 2006; Sreenivasan et al., 2014). When behavioral goals change, sensory cortex exhibits corresponding changes in neural activity (Lepsien & Nobre, 2006; Katus et al., 2015b), suggesting that the activation of WM content can be flexibly modulated through the selective allocation of attention to currently task-relevant representations in perceptual brain

areas. It has previously been argued that the focus of attention in WM is controlled by a single central / supramodal system that is shared with perception, and also shared between sensory modalities (Cowan, 2011). If this supramodal mechanism operates in a space-based fashion, directing attention to tactile and visual WM representations on opposite sides should lead to costs in behavioral and EEG measures (see evidence from perception research: e.g., Eimer, 2001). However, tCDA and CDA components were neither attenuated nor delayed in opposite sides blocks relative to same side blocks, and WM accuracy was virtually identical in both types of blocks. The absence of any costs for WM performance in opposite sides blocks could have been a result of the bimodal WM task not being sufficiently demanding in the main experiment. In a behavioural follow-up experiment where six stimuli (two tactile and four visual stimuli) had to be simultaneously maintained, performance was again identical in same side and opposite sides blocks (see section 3.3), thereby ruling out this possibility. Overall, these results suggest that the spatially selective allocation of attention to multimodal WM representations is mediated by independent processes for tactile and visual information.

To demonstrate the spatial independence of maintenance processes for tactile and visual information, we here used a spatial manipulation, and focused on spatially-selective markers of WM maintenance. We showed that the polarities of the sustained tCDA / CDA components can vary independently of each other, suggesting that these components index modality-specific spatial biasing processes that operate concurrently and independently. However, this conclusion does not necessarily imply that tactile and visual WM rely on independent resources, which would entail independent capacity limitations. To confirm an independence of WM resources for touch and vision, what has to be shown is that the number of items that can be successfully retained in one modality is not affected by the number of items maintained in another modality. Future behavioral and electrophysiological

studies hence need to manipulate WM load separately for each modality, with multisensory sample sets sizes that exceed the capacity limits of unimodal WM (cf. Cowan, 2001; Vogel & Machizawa, 2004). Further, while we here employed the lateralized tCDA / CDA components to track the focus of spatial attention in multimodal WM, we do not claim that spatial attention is the only mechanism involved in the activation of WM representations. Attentional mechanisms that operate in a feature- and/or object-based manner may also contribute to the maintenance of information in WM. Recent evidence has linked the visual CDA component with object-based attentional mechanisms (Luria & Vogel, 2011; Ikkai et al., 2010), and it is possible that such mechanisms were also activated in our study, in particular, because the visual task required memory for features (i.e., colors) at specific locations. To shed light on the roles of feature- or object-based attention mechanisms for the maintenance of multimodal information in WM, future experiments could separately manipulate the type of information maintained in touch and vision, and compare tCDA / CDA amplitudes between purely spatial WM tasks and tasks that require WM for features or objects. The novel finding in this study is that spatial attention operates in a modality-specific fashion during WM maintenance. The importance of this finding is owed to the fact that WM representations are inherently spatially specific. Stimulus locations are obligatorily stored in tactile (Katus et al., 2012) and visual WM (Kuo et al., 2009), even for tasks that do not explicitly require memory for locations. The spatial layout of WM representations is a direct consequence of the map-like organization of sensory cortical regions that were recruited to store information (Franconeri et al., 2013; Cavanagh et al., 2010). Spatially selective mechanisms play a vital role in maintaining focal attention on WM content, because this content needs to be activated at the site where it is stored in the brain.

The apparent independence of spatial biasing mechanisms for visual and tactile WM may seem inconsistent with previous behavioral and ERP experiments that investigated

crossmodal links in perceptual attention (Spence & Driver, 1996; Spence et al., 2000; Eimer, 2001; Eimer & Driver, 2000; Eimer & Schröger, 1998). Directing spatial attention to one side in a primary modality resulted in a corresponding spatial bias for a different secondary modality, even when stimuli in this secondary modality were task-irrelevant or equally likely to appear on either side. It remains possible to deploy auditory and visual attention simultaneously to opposite sides, though not as effectively as directing attention to the same side in both modalities (Spence & Driver, 1996; Eimer, 2001), suggesting that the control mechanisms responsible for allocating spatial attention to sensory stimuli in different modalities are separable but linked. The presence of such crossmodal links has been explained by assuming that perceptual attention operates within a spatial reference frame that is shared across modalities, and is based on external spatial coordinates (Driver & Spence, 1998; Eimer et al., 2001; Eimer & Driver, 2001; for further discussion, see Heed et al., 2015). If spatial synergies in crossmodal perceptual attention are the result of a shared reference frame, the absence of crossmodal interactions during the spatially selective attentional maintenance of visual and tactile WM representations in our study is not surprising, because these representations use different spatial coordinate systems. Stimuli in tactile WM are indexed in somatotopic, rather than allocentric / retinotopic coordinates, as demonstrated by the observation that tCDA components emerge over somatosensory cortex contralateral to the hand where a tactile stimulus is memorized, regardless of whether this hand is placed on the left or right side in external space (Katus et al., 2015b). The incommensurability of spatial coordinate systems for tactile and visual WM representations (somatotopic versus retinotopic) may be the main reason why distinct foci of spatial attention can be simultaneously maintained on multimodal WM content.

How might these modality-specific spatial biasing mechanisms for tactile and visual WM contents be implemented at the neural level? There are extensive reciprocal

connections between higher-order control regions such as PFC and/or PPC and tactile and visual cortical areas (Andersen et al., 1997; Barbas, 2000). In these control regions, persistent activity of neurons with receptive fields that match the locations of memorized stimuli during WM retention may represent stable activation patterns that are centred on task-relevant coordinates in spatial priority maps (Compte et al., 2000; Wang, 2001; Ikkai & Curtis, 2011; Jerde & Curtis, 2013). The PPC is a zone of multisensory convergence that plays a central role in coordinate transformations, such as the remapping of tactile stimuli into an external, supramodal, frame of reference (Azañón et al., 2010), but it is still controversial whether spatial maps in PPC are consistently referenced to external space (Silver & Kastner, 2009; Medendorp et al., 2011). Neurons in ventral intraparietal area (VIP) of macaque cortex encode stimuli using a variety of modality-specific and intermediate frames of reference (Avillac et al., 2005). These spatial maps may provide pointers to visual and tactile WM representations that employ different modality-specific coordinate systems (cf. Cavanagh et al., 2010). We hypothesize that the spatially selective maintenance of visual and tactile WM representations, as reflected by lateralized delay activity, is mediated by modality-specific mechanisms that bridge the gap between top-down control areas such as PFC and/or PPC, and WM storage systems in sensory cortex. More precisely, we suggest that the recruitment of modality-specific cortical regions for the storage of information is accompanied by a recruitment of modality-specific functions that implement the attentional biasing of WM content at the site where this information is stored in the brain. This interpretation does not rule out the possibility of genuinely supramodal control functions at central levels. For example, connectionist models (e.g., Fuster, 2009) assume that central and modality-specific mechanisms are both critical for WM, which depends on the interplay between executive networks (in frontal cortex) and sensory networks (in posterior cortex). The assumption that modality-specific mechanisms are implicated in WM

is further consistent with hierarchical theories, which posit that WM encompasses modality-specific processing systems that are controlled by a central mechanism in a top-down fashion (e.g., Baddeley, 2003).

Conclusion WM emerges due to the attentional activation of brain regions that store stimulus-specific information. We observed distinct foci of tactile and visual spatial attention during the concurrent maintenance of multimodal stimuli in WM. This suggests that multimodal WM representations are stored in distributed brain regions which are subject to separate spatially-specific biasing mechanisms that operate simultaneously and independently during WM retention.

Acknowledgments This work was funded by the Deutsche Forschungsgemeinschaft (DFG Grants KA 3843/1-1, KA 3843/1-2 and KA 3843/2-1), the Leverhulme Trust (Grant RPG-2015-370), and was supported by a grant from the Economic and Social Research Council (ESRC), United Kingdom. We thank Sue Nicholas for help in setting up the tactile stimulation hardware, and Andreas Widmann for providing EEGLab plugins for digital filtering and spherical spline interpolation. We furthermore thank John McDonald, Tobias Heed and five anonymous reviewers for their constructive comments on an earlier version of this manuscript.

References

Andersen, R. A., Snyder, L. H., Bradley, D. C., & Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annu. Rev. Neurosci.*, *20*, 303–330.

Avillac, M., Denève, S., Olivier, E., Pouget, A., & Duhamel, J.-R. (2005). Reference frames for representing visual and tactile locations in parietal cortex. *Nat. Neurosci.*, *8*, 941–949.

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends Cogn. Sci.*, *5*, 119–126.

Awh, E., Vogel, E. K., & Oh, S. H. (2006). Interactions between attention and working memory. *Neuroscience*, *139*, 201–208.

Azañón, E., Longo, M. R., Soto-Faraco, S., & Haggard, P. (2010). The posterior parietal cortex remaps touch into external space. *Curr. Biol.*, *20*, 1304–1309.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci., 4*, 829–839.

Barbas, H. (2000). Connections underlying the synthesis of cognition, memory, and emotion in primate prefrontal cortices. *Brain Res. Bull.*, *52*, 319–330.

Cavanagh, P., Hunt, A. R., Afraz, A., & Rolfs, M. (2010). Visual stability based on remapping of attention pointers. *Trends Cogn. Sci.*, *14*, 147–153.

Compte, A., Brunel, N., Goldman-Rakic, P. S., & Wang, X. J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb. Cortex*, *10*, 910–923.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain. Sci., 24*, 87-114; discussion 114-85.

569 Cowan, N. (2011). The focus of attention as observed in visual working memory tasks:
570 making sense of competing claims. *Neuropsychologia*, *49*, 1401–1406.
571 Cowan, N., Li, D., Moffitt, A., Becker, T. M., Martin, E. A., Saults, J. S., & Christ, S. E.
572 (2011). A neural region of abstract working memory. *J. Cogn. Neurosci.*, *23*, 2852–2863.
573 Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during
574 working memory. *Trends Cogn. Sci.*, *7*, 415–423.
575 Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-
576 trial EEG dynamics including independent component analysis. *J. Neurosci. Methods*,
577 *134*, 9–21.
578 Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG
579 data using higher-order statistics and independent component analysis. *NeuroImage*,
580 *34*, 1443–1449.
581 Driver, J., & Spence, C. (1998). Attention and the crossmodal construction of space. *Trends*
582 *Cogn. Sci.*, *2*, 254–262.
583 Eimer, M. (2001). Crossmodal links in spatial attention between vision, audition, and touch:
584 evidence from event-related brain potentials. *Neuropsychologia*, *39*, 1292–1303.
585 Eimer, M., Cockburn, D., Smedley, B., & Driver, J. (2001). Cross-modal links in
586 endogenous spatial attention are mediated by common external locations: evidence
587 from event-related brain potentials. *Exp. Brain Res.*, *139*, 398–411.
588 Eimer, M., & Driver, J. (2000). An event-related brain potential study of cross-modal links in
589 spatial attention between vision and touch. *Psychophysiology*, *37*, 697–705.
590 Eimer, M., & Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial
591 attention: evidence from event-related brain potential studies. *Neurosci. Biobehav. Rev.*,
592 *25*, 497–511.

593 Eimer, M., & Schröger, E. (1998). ERP effects of intermodal attention and cross-modal links
594 in spatial attention. *Psychophysiology*, *35*, 313–327.
595 Emrich, S. M., Riggall, A. C., LaRocque, J. J., & Postle, B. R. (2013). Distributed patterns of
596 activity in sensory cortex reflect the precision of multiple items maintained in visual
597 short-term memory. *J. Neurosci*, *33*, 6516–6523.
598 Ester, E. F., Sprague, T. C., & Serences, J. T. (2015). Parietal and Frontal Cortex Encode
599 Stimulus-Specific Mnemonic Representations during Visual Working Memory. *Neuron,*
600 *87*, 893-905.
601 Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources:
602 competitive content maps for attention and memory. *Trends Cogn. Sci., 17*, 134-141.
603 Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory.
604 *Science*, *173*, 652–654.
605 Fuster, J. M (2009). Cortex and memory: emergence of a new paradigm. *J. Cognitive*
606 *Neurosci.*, *21*, 2047–2072.
607 Golomb, J. D., Chun, M. M., & Mazer, J. A. (2008). The native coordinate system of spatial
608 attention is retinotopic. *J. Neurosci.*, *28*, 10654–10662.
609 Golomb, J. D., & Kanwisher, N. (2012). Retinotopic memory is more precise than
610 spatiotopic memory. *Proc. Natl. Acad. Sci. U.S.A.*, *109*, 1796–1801.
611 Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory
612 in early visual areas. *Nature*, *458*, 632–635.
613 Heed, T., Buchholz, V. N., Engel, A. K., & Röder, B. (2015). Tactile remapping: from
614 coordinate transformation to integration in sensorimotor processing. *Trends Cogn. Sci.*,
615 *19*, 251–258.
616 Ikkai, A., & Curtis, C. E. (2011). Common neural mechanisms supporting spatial working
617 memory, attention and motor intention. *Neuropsychologia*, *49*, 1428–1434.

Ikkai, A., McCollough, A. W., & Vogel, E. K. (2010). Contralateral Delay Activity Provides a Neural Measure of the Number of Representations in Visual Working Memory. *J. Neurophysiol., 103*, 1963–1968.

JASP Team (2016). JASP (Version 0.7.5.5) [Computer software].

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press. UK: Oxford University Press.

Jerde, T. A., & Curtis, C. E. (2013). Maps of space in human frontoparietal cortex. *J. Physiol. Paris*, *107*, 510–516.

Jonides, J., Lacey, S. C., & Nee, D. E. (2005). Processes of Working Memory in Mind and Brain. *Curr. Dir. Psychol. Sci.*, *14*, 2–5.

Katus, T., Andersen, S. K., & Müller, M. M. (2012). Nonspatial cueing of tactile STM causes shift of spatial attention. *J. Cogn. Neurosci., 24*, 1596–1609.

Katus, T., & Eimer, M. (2015). Lateralized delay period activity marks the focus of spatial attention in working memory: evidence from somatosensory event-related brain potentials. *J. Neurosci.*, *35*, 6689–6695.

Katus, T., Grubert, A., & Eimer, M. (2015a). Electrophysiological evidence for a sensory recruitment model of somatosensory working memory. *Cereb. Cortex*, *25*, 4697–4703.

Katus, T., & Müller, M. M. (2016). Working memory delay period activity marks a domain-unspecific attention mechanism. *NeuroImage*, *128*, 149–157.

Katus, T., Müller, M. M., & Eimer, M. (2015b). Sustained maintenance of somatotopic information in brain regions recruited by tactile working memory. *J. Neurosci.*, *35*, 1390–1395.

Kuo, B. C., Rao, A., Lepsien, J., & Nobre, A. C. (2009). Searching for targets within the spatial layout of visual short-term memory. *J. Neurosci., 29*, 8032–8038.

Lepsien, J., & Nobre, A. C. (2006). Cognitive control of attention in the human brain: Insights from orienting attention to mental representations. *Brain Res.*, *1105*, 20–31.

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia, 49*, 1632-1639.

Medendorp, W. P., Buchholz, V. N., Van Der Werf, Jurrian, & Leoné, Frank T M (2011). Parietofrontal circuits in goal-oriented behaviour. *Eur. J. Neurosci.*, *33*, 2017–2027.

Mendoza-Halliday, D., Tores, S., & Martinez-Trujillo, J. C. (2014). Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nat. Neurosci., 17*, 1255-1262.

Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology*, *35*, 99–115.

Riley, M. R., & Constantinidis, C. (2016). Role of Prefrontal Persistent Activity in Working Memory. *Front. Syst. Neurosci., 9*, 181.

Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: Fully automated statistical thresholding for EEG artifact rejection. *J. Neurosci. Methods*, *192*, 152–162.

Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nat. Rev. Neurosci.*, *6*, 97–107.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*, 23–38.

Romo, R., & Salinas, E. (2003). Flutter discrimination: neural codes, perception, memory and decision making. *Nat. Rev. Neurosci., 4*, 203-218.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *J. Math. Psychol., 56,* 356-374.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev., 16*, 225-237.

667 Silver, M. A., & Kastner, S. (2009). Topographic maps in human frontal and parietal cortex.
668 *Trends Cogn. Sci.*, *13*, 488–495.
669 Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *J.*
670 *Exp. Psychol. Hum. Percept. Perform.*, *22*, 1005–1030.
671 Spence, C., Pavani, F., & Driver, J. (2000). Crossmodal links between vision and touch in
672 covert endogenous spatial attention. *J. Exp. Psychol. Hum. Percept. Perform.*, *26*,
673 1298–1319.
674 Sreenivasan, K. K., Curtis, C. E., & D'Esposito, M. (2014). Revisiting the role of persistent
675 neural activity during working memory. *Trends Cogn. Sci.*, *18*, 82–89.
676 Supèr, H., Spekreijse, H., & Lamme, V. A. (2001). A neural correlate of working memory in
677 the monkey primary visual cortex. *Science*, *293*, 120–124.
678 Tenke, C. E., & Kayser, J. (2012). Generator localization by current source density (CSD):
679 implications of volume conduction and field closure at intracranial and scalp resolutions.
680 *Clin. Neurophysiol.*, *123*, 2328–2345.
681 Ulrich, R., & Miller, J. O. (2001). Using the jackknife-based scoring method for measuring
682 LRP onset effects in factorial designs. *Psychophysiology*, 38, 816 – 827.
683 Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in
684 visual working memory capacity. *Nature*, *428*, 748–751.
685 Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal
686 individual differences in controlling access to working memory. *Nature*, *438*, 500–503.
687 Wagenmakers, E. J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian
688 hypothesis testing for psychologists: A tutorial on the Savage-Dickey method. *Cogn.*
689 *Psych., 60*, 158-189.
690 Wang, X. J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends*
691 *Neurosci.*, *24*, 455–463.

692 Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term
693 memory for objects. *Nature*, *440*, 91–95.
694 Zhou, Y. D., & Fuster, J. M. (1996). Mnemonic neuronal activity in somatosensory cortex.
695 *Proc. Natl. Acad. Sci. U.S.A.*, *93*, 10533–10537.