Journal of
Cognitive Neuroscience

Intermodal attention shifts in multimodal working memory

Journal:	*Journal of Cognitive Neuroscience*
Manuscript ID	JOCN-2016-0305.R1
Manuscript Type:	Original
Date Submitted by the Author:	n/a
Complete List of Authors:	Katus, Tobias; Birkbeck, University London, Department of Psychological Sciences Grubert, Anna; Birkbeck College, Department of Psychology Eimer, Martin; Birkbeck College, Psychology;
Keywords:	Attention: Somatosensory, Attention: Visual, EEG, Memory: Working memory, Executive functions

Full manuscript title

Intermodal attention shifts in multimodal working memory

Running title

Intermodal attention shifts in working memory

Authors and affiliations: Tobias Katus[a], Anna Grubert[a] & Martin Eimer[a]

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

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.

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 supported by a grant from the Economic and Social Research Council (ESRC), United Kingdom. We thank Sue Nicholas for technical assistance, and Andreas Widmann for providing EEGLab plugins for digital filtering and spherical spline interpolation.

Abstract

Attention maintains task-relevant information in working memory (WM) in an active state. We investigated whether the attention-based maintenance of stimulus representations that were encoded through different modalities is flexibly controlled by top-down mechanisms that depend on behavioral goals. Distinct components of the event-related potential (ERP) reflect the maintenance of tactile and visual information in WM. We concurrently measured tactile (tCDA) and visual contralateral delay activity (CDA) to track the attentional activation of tactile and visual information during multimodal WM. Participants simultaneously received tactile and visual sample stimuli on the left and right sides, and memorized all stimuli on one task-relevant side. After 500 ms, an auditory retro-cue indicated whether the sample set's tactile or visual content had to be compared with a subsequent test stimulus set. tCDA and CDA components that emerged simultaneously during the encoding phase were consistently reduced after retro-cues that marked the corresponding (tactile or visual) modality as task-irrelevant. The absolute size of cue-dependent modulations was similar for the tCDA/CDA components and did not depend on the number of tactile/visual stimuli that were initially encoded into WM. Our results suggest that modality-specific maintenance processes in sensory brain regions are flexibly modulated by top-down influences that optimize multimodal WM representations for behavioral goals.

Introduction

Stimulus-specific information that is needed for ongoing behavior, but no longer physically present, is temporarily represented in working memory (WM). According to the sensory recruitment hypothesis (Curtis & D'Esposito, 2003; D'Esposito, 2007; Jonides, Lacey, & Nee, 2005; Postle, 2006), stimulus representations are stored in the same modality-specific perceptual brain regions that have encoded the original stimulus into WM. These representations are maintained in an active state through the allocation of selective attention, which is controlled in a top-down fashion by higher-level cortical regions (such as the prefrontal cortex, PFC; Gazzaley & Nobre, 2012; Sreenivasan, Curtis, & D'Esposito, 2014). The flexibility of attentional processes that operate within visual WM representations has been demonstrated in experiments where retro-cues were presented after the initial encoding of a visual sample stimulus set (Eimer & Kiss, 2010; Kuo, Rao, Lepsien, & Nobre, 2009; Kuo, Stokes, & Nobre, 2012; Myers, Walther, Wallis, Stokes, & Nobre, 2015). When these retro-cues specified the locations of a subset of stored items that had to be maintained, attention was selectively allocated to these task-relevant items, resulting in benefits for visual WM performance (Griffin & Nobre, 2003; Lepsien & Nobre, 2006). This shows that attention can modulate the activation of specific representations, even after they have been encoded into visual WM. Analogous attentional modulations have also been found for representations in tactile WM (Katus, Andersen, & Müller, 2012; Katus, Müller, & Eimer, 2015b).

While it is clear that top-down attentional control mechanisms can operate on WM representations within a specific sensory modality (vision or touch), it is unknown whether attention can also be flexibly shifted between mnemonic representations that were encoded through different modalities, and hence, are stored in distinct modality-specific cortical regions. In the present study, we tracked goal-dependent activation changes of stimulus representations in somatosensory and visual cortex during the retention period of a multimodal WM task to determine whether attentional maintenance can be selectively switched off for WM contents that are no longer task-relevant. Bimodal sets of tactile and visual sample stimuli were simultaneously presented on the left and right sides, and participants had to memorize the tactile and visual sample sets on one side (block-wise left or right). An auditory retro-cue that was presented 500 ms after the bimodal sample sets indicated whether the memorized visual or tactile samples had to be maintained for a comparison with a subsequent test stimulus set. After this cue, it was no longer necessary to maintain the now task-irrelevant stimuli of the uncued modality.

To track the activation of tactile and visual information in WM before and after the retro-cue, we examined components of the event-related potential (ERP) that reflect the attention-based maintenance of tactile and visual information. The contralateral delay activity (CDA) is elicited over posterior visual areas contralateral to the side where memorized visual stimuli have been presented, and is sensitive to WM load and individual differences in WM capacity (Vogel & Machizawa, 2004; Vogel, McCollough, & Machizawa, 2005). The tactile CDA (tCDA) component is the somatosensory equivalent of the visual CDA, and manifests over somatosensory cortex contralateral to maintained tactile stimuli (Katus & Eimer, 2015; Katus, Grubert, & Eimer, 2015a; Katus & Müller, 2016). Using current source density (CSD; Tenke & Kayser, 2012) transforms of ERP data, we have previously demonstrated that it is possible to dissociate between the tCDA and CDA components by means of their distinct topographical distributions (Katus & Eimer, 2016). In a multimodal WM experiment, participants memorized tactile and visual stimuli on either the same side or on opposite sides. tCDA and CDA components were elicited over somatosensory and visual regions of the same hemisphere, when these multisensory stimuli were memorized on the same side. Memorizing tactile and visual stimuli on opposite sides, in contrast, led to tCDA and CDA components over somatosensory and visual areas of different hemispheres. This finding demonstrates that the tCDA and CDA are distinct ERP components, reflecting the attention-based maintenance of tactile and visual information, respectively.

In a retro-cue study, we here concurrently measured the tCDA and CDA components to test whether the active maintenance of tactile and visual information adapts to changes in the behavioral relevance of these information. During the early retention period prior to the presentation of the retro-cue, tCDA and CDA components should be triggered simultaneously over somatosensory and visual areas, reflecting the concurrent maintenance of the tactile and visual sample stimuli. The critical question was how these components would be affected by subsequent retro-cues that retrospectively marked one of these two modalities as task-irrelevant. If the activation of tactile and visual WM representations can be flexibly modulated in line with changing behavioral goals, neural activity at somatosensory (tCDA) and visual (CDA) regions of interests (ROIs) should exhibit goal-dependent modulations after retro-cues have been presented (Cued modality x ROI interactions). Visual CDA components should be strongly attenuated following retro-cues that instruct participants to selectively maintain tactile sample stimuli only, whereas tCDA components should be reduced in size after the retrospective cueing of vision. In two experimental sessions, we also manipulated tactile and visual WM load (load 2 for both touch and vision in Session 1; load 1 for touch and load 3 for vision in Session 2) to examine whether the extent of top-down modulations depend on the amplitudes of the tCDA/CDA components in the period before the retro-cue. To ensure that participants would be able to encode and maintain all task-relevant sample stimuli prior to the presentation of the retro-cue, the combined (tactile + visual) WM load was 4 stimuli in each Session.

Materials and Methods

Participants

The study involved two recording sessions run on separate days. Twenty neurologically unimpaired observers were paid to participate in Session 1. Two of these observers were excluded from statistical analyses, and were not re-invited to participate in Session 2. For one participant, error rate in the tactile task exceeded 40%. The other participant was excluded due to excessive EEG artifacts. The remaining 18 participants (mean age 30 years, range 20-44 years, 11 female, 16 right-handed) completed both testing sessions. All participants gave informed written consent prior to testing. The study was conducted in accordance with the Declaration of Helsinki, and was approved by the Psychology Ethics Committee of Birkbeck, University of London.

Stimuli and stimulation hardware

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, using an eight-channel sound card (M-Audio, Delta 1010LT) controlled by Matlab routines

(MathWorks, Natick, MA). All tactile stimuli were 100 Hz sinusoids (duration: 200 ms; intensity: 0.37 N). The auditory cues were presented via headphones for 200 ms. Cues had either a low pitch (600 Hz) or a high pitch (1100 Hz), and consisted of sinusoid waveforms with ramped onset and offset (10 ms ramps). The cues were played on top of white noise that was continuously presented to mask any sounds produced by the tactile stimulators.

Visual stimuli were colored squares (0.63° of visual angle each) presented for 200 ms against a black background on a 22 inch monitor (Samsung wide SyncMaster 2233; 100 Hz refresh rate, 16 ms response time). Six equiluminant colors (~11.8 cd/m^2) were used in the experiment (CIE color coordinates: red = .627/.336; green = .263/.568; blue = .189/.193; yellow = .422/.468; cyan = .212/.350; magenta = .289/.168). A white fixation dot was present on the screen center throughout the experiment. In Session 1, two squares were equidistantly presented on each side of the display (to the left and right of fixation), with 1.26° and 0.52° offset from the x- and y-axes, respectively (measured relative to the squares' centers). In Session 2, each display side contained three squares, the two from Session 1 and an additional one to their left or right side on the left or right display side, respectively (offset from x- and y-axes: 2.22° and 0.52°, respectively).

Task design and randomization procedures

In two sessions, participants performed bimodal WM tasks with identical designs. WM load - i.e., the number of stimuli per side - varied for the tactile and visual tasks across the experimental sessions (Session 1: 2 tactile and 2 visual stimuli; Session 2: 1 tactile and 3 visual stimuli). Figure 1 illustrates the general procedure. A bimodal sample set was presented 500 ms before an auditory cue, which was followed by a bimodal memory test after additional 1500 ms. Vocal responses were recorded via a headset microphone in the 2000 ms period following the memory test, and the next trial began after a jittered interval of 700 to 1000 ms. Observers had to memorize the locations of the tactile sample stimuli and the colors of the visual samples on one side (left or right). This task-relevant side was specified via written instructions on the computer screen at the start of each experimental block, and changed after each block. The relevant side for the first experimental block was randomly determined for each participant. The pitch of the auditory retro-cue (high versus low) indicated on a trial-to-trial basis whether the tactile (50%) or visual (50%) sample stimuli had to be retained in order to be compared with the memory test set. The pitch/modality assignment was counterbalanced across participants. For each modality and on each side, it was equally likely that the test set was identical (match, 50%) or differed (mismatch, 50%) relative to the sample set.

Tactile and visual stimuli were presented bilaterally, and were separately randomized on the left and right sides, as explained below for one side. Two randomly selected stimulators delivered the tactile sample stimuli in Session 1. On memory match trials, the same locations were stimulated. On mismatch trials, one (67% of mismatch trials) or both test stimuli (33%) were delivered to a different location. In Session 2, the sample stimulus was presented by one randomly selected stimulator. The same location was again stimulated at test on match trials, and a different location was stimulated on mismatch trials. In Session 1, two different colors were randomly selected for the visual sample set. The same two colors were shown again at the same locations on match trials. On mismatch trials, one stimulus changed its color between sample and test (67%), or both colored samples swapped their locations in the test set (33%). In Session 2, three different colors were randomly selected for the visual sample set, and these colors were repeated on match trials. On mismatch trials, one randomly selected stimulus changed its color (33%), or two randomly selected stimuli swapped their locations (33%), or all three stimuli swapped their locations in the test set (33%).

Each session comprised twelve 4-minutes blocks with 40 trials each; 60 trials were run for each of the eight combinations of experimental conditions (cued modality: touch vs. vision; task-relevant side: left vs. right; response: match vs. mismatch). Participants were asked to maintain central gaze fixation and to avoid head and body movements during the recording. Instructions emphasized accuracy over speed. Feedback on the percentage of correct responses was provided after each block. One training block was run before the first experimental block.

Insert Figure 1 about here

Processing 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 were used to record lateral eye movements (horizontal electrooculogram, HEOG). Continuous EEG data were online referenced to the left mastoid, and re-referenced offline to the arithmetic mean of both mastoids (electrode sites TP9 and TP10) for data preprocessing. Data were offline filtered with a 30Hz low-pass finite impulse response filter (Blackman window, filter order 500). EEG was segmented into 2200 ms intervals ranging from 200 ms prior to 2000 ms after sample stimulus onset, and were corrected relative to a 200 ms pre-stimulus baseline.

Blind source separation of EEG data was performed using the *Independent Component Analysis* (ICA) algorithm implemented in the EEGLab toolbox (Delorme & Makeig, 2004; Delorme, Sejnowski, & Makeig, 2007). Independent components (ICs) accounting for blinks were subtracted from the data. Epochs with horizontal eye movements were identified and rejected using a differential step function that ran on the bipolarized HEOG (step width 100 ms, threshold 30 µV). Additionally, ICs accounting for horizontal eye movements were subtracted from EEG epochs to remove residual traces of ocular artifacts that had not exceeded the amplitude threshold of the step function. Epochs were furthermore screened for slow (< 7 Hz) lateralized drifts which would compromise the analysis of the sustained tCDA and CDA components. Difference waves from the 27 lateral electrode pairs (e.g. C3/4) were Fourier transformed, to calculate spectral power in 7 frequency bins between 0.5 and 7 Hz on a single trial level (for a detailed description of this procedure, see Katus & Müller, 2016). Trials where at least two electrode pairs picked up difference waves with unusual spectral profiles were discarded (rejection criterion: 2 electrodes with median z-scores above 2.5). The remaining EEG epochs entered *Fully Automated Statistical Thresholding for EEG Artifact Rejection* (FASTER, Nolan, Whelan, & Reilly, 2010) for the interpolation of noisy electrodes, and were subsequently converted to current source densities (CSDs: iterations = 50, m = 4, lambda = 10^{-5}; compare Tenke & Kayser, 2012). After artifact rejection and elimination of trials with incorrect responses, 89.1% of all epochs were retained for statistical analyses (Session 1: 89.8%, Session 2: 88.4%).

CSDs from three adjacent electrodes were averaged, separately for the hemisphere contralateral and ipsilateral to the memorized sample stimuli on the task-relevant side. Tactile contralateral delay activity (tCDA component) was measured at lateral central scalp regions (C3/4, FC3/4, CP3/4), and visual contralateral delay activity (CDA) was measured at lateral occipital scalp regions (PO7/8, PO3/4, O1/2) (as in Katus & Eimer, 2016). Statistical tests were conducted on difference values of contralateral minus ipsilateral CSDs, averaged between 300 and 600 ms after sample onset for the analysis of delay activity in the period before the cue, and between 800 and 2000 ms after sample onset for the analyses of delay activity after the cue.

The error bars in graphs showing contra- minus ipsilateral difference values indicate 95% within-subject confidence intervals (CIs), which were calculated for each condition by separate 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., $y \neq 0$), and is symbolized by asterisks (* for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$, ns for $p > 0.05$). Topographic voltage maps display spline-interpolated difference values that were obtained by subtracting CSDs ipsilateral to the memorized stimuli from contralateral CSDs. The resulting difference values were collapsed across blocks in which the memory task was performed for stimuli on the left- or right side, by flipping electrode coordinates in left-side memory trials over the midline.

Results

Behavioral data

Participants responded correctly in 93.3% of all trials (93.5% correct in Session 1, 93.2% in Session 2). The sensitivity index d-prime (d') was submitted to a two-way repeated measures ANOVA with the factors *Session* and *Cued modality* (touch vs. vision). There were no significant main effects (all ps > 0.7). As predicted, a Session x Cued modality interaction ($F(1,17) = 55.373$, $p < 10^{-6}$) confirmed that task performance was modulated by tactile/visual WM load. As illustrated in Figure 2, performance in the tactile task was better with Load 1 in Session 2 than Load 2 in Session 1 ($t(17) = 4.589$, $p < 0.001$). Visual task performance was better with Load 2 in Session 1 than with Load 3 in Session 2 ($t(17) = 5.782$, $p < 10^{-4}$).

Insert Figure 2 about here

Electrophysiological data

Early retention period (300-600 ms). Figure 3 shows CSD transforms of ERPs elicited by the bimodal sample set in the early period of the retention period in Session 1 and Session 2. This early time period was defined between 300 and 600 ms after sample onset, as neural responses to the retro-cue did not manifest before 600 ms after the sample onset (see Figure 3, left column). We expected load-dependent modulations for the tCDA and CDA components in this pre-cue period, with larger tCDA components for Load 2 (Session 1) than Load 1 (Session 2), and larger visual CDAs with Load 3 (Session 2) than Load 2 (Session 1). tCDA/CDA mean amplitudes were submitted to a two-way repeated measures ANOVA with the factors Session and ROI (somatosensory vs. visual). The presence of load-dependent amplitude modulations during the pre-cue period was substantiated by a significant Session x ROI interaction ($F(1, 17) = 12.011$, $p = 0.003$). As shown in Figure 3, tCDA amplitudes were larger for two tactile items compared to one tactile item (Session 1 vs. 2, $t(17) = 4.226$, $p < 0.001$), and CDA amplitudes were larger for three relative to two visual items (Session 2 vs. 1, $t(17) = 2.186$, $p = 0.043$). Amplitudes were generally larger at visual ROIs (CDA) relative to somatosensory ROIs (tCDA) (main effect ROI: $F(1,17) = 4.693$, $p = 0.045$). To assess the reliability of lateralized components in the pre-cue period, mean amplitudes were tested against zero. Statistically significant CSD lateralization was found for somatosensory and visual ROIs in both Sessions (Session 1 - tCDA: $t(17) = 5.660$, $p < 10^{-4}$; CDA: $t(17) = 3.007$, $p = 0.008$; Session 2 - tCDA: $t(17) = 2.231$, $p = 0.039$; CDA: $t(17) = 3.824$, $p = 0.001$), confirming that tCDA and CDA components were reliably present in all Load conditions.

Insert Figure 3 about here

Late retention period (800-2000 ms). To examine changes in the activation states of tactile and visual WM representation following the retro-cues, statistical analyses were based on contra- minus ipsilateral difference values, averaged between 800 and 2000 ms after sample onset (i.e., from 300 ms after retro-cue onset to the end of the retention period). Task-dependent modulations of the tCDA (i.e., reduced amplitudes after the cueing of vision, relative to touch) and the CDA (reduced amplitudes after the cueing of touch, rather than vision) would be reflected by a Cued modality x ROI interaction.

The predicted Cued modality x ROI interaction ($F(1,17) = 20.354$, $p < 0.001$) was confirmed by a three-way repeated measures ANOVA on tCDA/CDA mean amplitudes with the factors Session, ROI and Cued modality (touch vs. vision). A main effect of ROI reflected the generally larger amplitude of the CDA as compared to tCDA ($F(1,17) = 17.305$, $p < 0.001$). No further effects or interactions were reliable (all ps > 0.2). The fact that no significant three-way interaction was found between Cued modality, ROI and Session suggests that retro-cues impacted the tCDA/CDA components in a fairly consistent manner in both Sessions, regardless of the load-dependent amplitudes of these components in the early retention period before the cues.

To examine whether cue-dependent modulations were equally reliable for tactile and visual ROIs, we submitted the tCDA and CDA components to separate ANOVAs with the factors Session and Cued modality. These analyses revealed main effects of Cued modality for the tCDA ($F(1,17) = 24.776$, $p < 0.001$) and the CDA ($F(1,17) = 6.165$, $p = 0.024$), in the absence of further significant main effects or interactions (all ps > 0.2). The somatosensory tCDA was attenuated when vision rather than touch was cued; likewise, the visual CDA was attenuated when touch rather than vision was cued (see Figure 4).

Figure 4 suggests that the cueing of vision led to a complete drop-to-baseline for the tCDA, whereas the cueing of touch attenuated, but did not fully eliminate the CDA. Formal tests of tCDA/CDA amplitudes against zero demonstrated that there was a statistically significant tCDA after the cueing of touch (Session 1: $t(17) = 3.459$, $p = 0.003$; Session 2: $t(17) = 4.358$, $p < 0.001$), which was completely eliminated after the cueing of vision ($ps > 0.2$). In contrast, CDA components were statistically reliable in the period after retro-cues in both Sessions, not only when vision was cued, but also when retro-cues specified touch as the relevant modality (all $ps < 0.05$).

The bar graphs in Figure 4 show that CDA components were generally larger than tCDA components, but that the absolute size of cue-dependent modulations (i.e., the amplitude differences between trials where the respective modality was marked as relevant versus irrelevant) was similar for the tCDA and CDA. To verify this statistically, attentional modulations were quantified by subtracting tCDA/CDA amplitudes when the corresponding tactile or visual modality was uncued, from amplitudes measured when this modality was cued. When these difference amplitudes were subjected to an ANOVA with the factors Session and ROI, no significant main effects or interactions were obtained (all $ps > 0.2$), suggesting that retro-cues modulated somatosensory and visual delay activity to a comparable degree.

Insert Figure 4 about here

Discussion

Attention-based maintenance processes keep information that has been encoded into working memory in an active state (Awh, Anllo-Vento, & Hillyard, 2000; Awh & Jonides, 2001). If the maintenance of sensory information is controlled in a goal-dependent fashion, it should be possible to selectively de-activate information that has been marked as behaviorally irrelevant, even after this information had been encoded into WM. In a multimodal WM task, we used CSD transforms of ERPs to concurrently track the attentional activation of information stored in somatosensory and visual cortex (see also Katus & Eimer, 2016). Participants initially memorized tactile and visual sample stimuli on one task-relevant side, before a retro-cue indicated whether the tactile or visual stimuli had to be actively maintained for comparison with a subsequent memory test.

Because retro-cues altered the behavioral relevance of tactile and visual WM representations, they should lead to an update of attentional control settings that govern the maintenance of information in somatosensory and visual cortex. If WM maintenance processes are sensitive to such changes in top-down control settings, the tactile and visual CDA components should show modulations that depend on whether retro-cues have instructed participants to selectively retain tactile or visual information. In line with this prediction, a significant ROI x Cued modality interaction was observed for the amplitudes of these components in the period after retro-cues. These tCDA/CDA modulations reveal systematic changes in the attentional activation states of tactile and visual WM representations that mirror their behavioral relevance. Lateralized delay activity, measured over somatosensory and visual ROIs as the difference between electrodes contralateral and ipsilateral to the memorized sample set (compare Figure 4, bottom panel), was consistently reduced in size after retro-cues that marked the respective (tactile or visual) modality as task-irrelevant, as compared to trials where WM content in this modality had to be retained. This finding shows that maintenance processes in modality-specific cortical areas can be flexibly controlled by goal-directed biasing signals from higher-level brain regions.

If the attention-based maintenance of sensory information in modality-specific cortical regions could be perfectly regulated by goal-dependent feedback signals from higher-level control areas, maintenance processes should have been completely de-activated for the modality that was retrospectively marked as task-irrelevant. In this case, tCDA or CDA components should have disappeared following retro-cues that instructed participants to selectively retain stimuli in the other modality. Such a drop-to-baseline was indeed observed for the somatosensory tCDA component after the retrospective cueing of vision. In contrast, the visual CDA remained significantly present when touch was cued, although CDA amplitudes were reliably reduced in size relative to trials where vision was cued. If the elimination of lateralized delay activity marks the de-activation of maintenance processes, the observation that only the tCDA component, but not the CDA, was completely eliminated when the associated modality was task-irrelevant could be interpreted as evidence for an asymmetry in the extent to which tactile and visual maintenance processes are sensitive to top-down control. However, the absolute size of cue-dependent modulations did not differ significantly between the tCDA and CDA components in the period after the retro-cue. Cueing of vision (rather than touch) reduced the tCDA by 0.13 mA/m^3, and the CDA was reduced by 0.10 mA/m^3 when touch (rather than vision) was cued; see bar graph in Figure 4. This suggests that the modulatory effects of goal-dependent feedback signals on maintenance processes in sensory areas may not differ systematically between touch and vision. Given that the visual CDA is generally larger in size than the somatosensory tCDA, a task-dependent reduction in the amplitude of these components by the same absolute amount may completely eliminate the tCDA, while only attenuating the CDA component. Furthermore, the size of cue-dependent modulations of the tactile and visual CDA components did not differ across Sessions 1 and 2, in spite of the fact that visual and tactile WM load differed between these sessions. During the early retention interval, prior to the retro-cue, tCDA and CDA amplitudes reflected the number of items that were initially encoded into tactile and visual WM (see Figure 3), in line with previous observations (e.g., Katus et al., 2015a; McCollough, Machizawa, & Vogel, 2007). Larger tCDA components were measured for tactile Load 2 (Session 1) relative to Load 1 (Session 2), and larger CDA components for visual Load 3 (Session 2) versus Load 2 (Session 1). The absence of a significant Session x ROI x Cued modality interaction for the post-cue period suggests that the changes in the size of tCDA/CDA components after the respective modality was marked as relevant versus irrelevant did not depend on the initial sizes of these components before the retro-cue was presented.

The fact that the visual CDA component remained reliably present after the retrospective cuing of touch may seem surprising, since it suggests that visual WM representations were still actively maintained even though this was no longer required. One possibility is that the CDA is not exclusively linked to visual WM, but may to some degree also reflect the maintenance of tactile stimuli. Neural generators of the CDA are assumed to be located in posterior parietal cortex (PPC; Becke, Müller, Vellage, Schoenfeld, & Hopf, 2015; Robitaille, Grimault, & Jolicoeur, 2009), consistent with fMRI evidence that the intraparietal sulcus (IPS) in the PPC shows load-dependent modulations in visual WM tasks (Todd & Marois, 2004; Xu & Chun, 2006). Since the

PPC receives multimodal sensory input and appears to be involved in multimodal WM (Cowan et al., 2011), as well as multisensory spatial attention (e.g., Macaluso, Frith, & Driver, 2000; Macaluso, Frith, & Driver, 2002), the active maintenance of task-relevant tactile sample stimuli could in principle be reflected by a CDA-like component, generated in the PPC, and/or in multimodal areas of occipitotemporal cortex (compare Amedi et al., 2001; Sathian et al., 2011). However, in all previous experiments of unimodal tactile WM that reported tCDA components during the maintenance of tactile stimuli (Katus & Eimer, 2015; Katus et al., 2015a; Katus & Müller, 2016; Katus et al., 2015b), no evidence was found for the simultaneous presence of a posterior CDA component. This suggests that the visual and tactile CDA components mirror dissociable maintenance processes for visual and tactile information, respectively (Katus & Eimer, 2016; for further discussion of the tCDA as a neural marker of somatosensory processing, see Katus et al., 2015b). Here, the sustained presence of a visual CDA after the retrospective cueing of touch may thus indicate generic limitations in the ability to regulate the activation states of visual stimulus representations that had been attended during encoding, but were subsequently marked as task-irrelevant. Once activated, such representations may retain an above-baseline level of activation, even when they are no longer needed for ongoing behavior (see also Rerko & Oberauer, 2013, for corresponding behavioral evidence).

The finding that the tCDA, but not the CDA, disappeared after the corresponding modality was cued as task-irrelevant, could also be linked with differences in the demands of our tactile and visual tasks. The visual task required memory for colors at specific locations, whereas the tactile task was a purely spatial memory task. Instead of reflecting general differences between touch and vision in the control of WM representations that are no longer relevant, the current pattern of tCDA and CDA results may indicate that the ability to de-activate task-irrelevant WM content is more limited for non-spatial attributes than for stimulus locations. This could be tested in future experiments with bimodal WM tasks where the same attributes have to be memorized in touch and vision (e.g., two purely spatial memory tasks, or two tasks requiring memory for a conjunction of spatial and non-spatial attributes). If results indicated that only the maintenance of spatial stimulus coordinates can be fully de-activated in a top-down fashion, this may suggest that a spatial indexing system that selectively maintains spatial pointers for behaviorally relevant memory content (compare Ikkai et al., 2010) is the main source of retrospective cueing effects in WM.

Previous behavioral and neuroimaging experiments demonstrated that changes in the allocation of attention after retro-cues optimize the activation states of WM representations in a goal-dependent manner. EEG studies have shown that retro-cues signaling the locations of task-relevant WM content guide spatial selection within unimodal tactile (Katus et al., 2015b) or visual WM representations (Griffin & Nobre, 2003; Kuo et al., 2012; Myers et al., 2015). Spatially selective modulations of WM content have not only been observed with spatial retro-cues, but also after the retrospective cueing of non-spatial stimulus attributes (i.e., stimulus intensity in tactile studies: Katus et al., 2012; color or shape in visual studies: Eimer & Kiss, 2010; Kuo et al., 2009); such effects indicate the selection of feature or object information, which is stored in cortical maps that are organized in a spatially specific manner (somatotopic vs. retinotopic for tactile vs. visual WM). There is also evidence that the retrospective cueing and subsequent attentional selection of object categories in WM leads to goal-dependent adjustments in the activation states of WM representations in distinct category-selective visual brain areas. fMRI studies reported that changes in neural activity in fusiform and parahippocampal areas reflect the behavioral relevance of retrospectively cued faces and scenes, respectively (Lepsien & Nobre, 2007; Lepsien, Thornton, & Nobre, 2011). These findings show that unimodal WM representations can be optimized through the retrospective selection of locations, features or objects, as mirrored by goal-dependent activation changes in functionally and anatomically distinct brain areas (for a review, see Lepsien & Nobre, 2006). Using a multimodal WM task, we here demonstrated for the first time that attentional feedback signals also control the activation level of WM representations across sensory modalities. The observation that dissociable modulations of the tCDA and CDA components mirrored the behavioral relevance of tactile and visual information supports the interpretation that these components reflect functionally distinct maintenance processes for somatosensory and visual information, respectively (Katus & Eimer, 2016).

Conclusion. The maintenance of sensory information in WM is mediated by processes that activate task-relevant representations at the site where this information is stored in the brain (i.e., in sensory cortex). Using a multimodal WM task, we showed that changes in the behavioral relevance of tactile / visual WM contents lead to an update of top-down control settings that are used to bias the activation states of information in somatosensory and visual cortical regions. This suggests that modality-specific maintenance processes are regulated by top-down influences that modulate multimodal WM representations in a goal-directed fashion.

References

Amedi, A., Malach, R., Hendler, T., Peled, S., & Zohary, E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nature Neuroscience, 4*(3), 324-40.

Awh, E., Anllo-Vento, L., & Hillyard, S. A. (2000). The role of spatial selective attention in working memory for locations: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, *12*(5), 840–847.

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, *5*(3), 119–126.

Becke, A., Müller, N., Vellage, A., Schoenfeld, M. A., & Hopf, J.-M. (2015). Neural sources of visual working memory maintenance in human parietal and ventral extrastriate visual cortex. *NeuroImage*, *110*, 78–86.

Cowan, N., Li, D., Moffitt, A., Becker, T. M., Martin, E. A., Saults, J. S., & Christ, S. E. (2011). A neural region of abstract working memory. *Journal of Cognitive Neuroscience*, *23*(10), 2852–2863.

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, *7*(9), 415–423.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage*, *34*(4), 1443–1449.

D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1481), 761–772.

Eimer, M., & Kiss, M. (2010). An electrophysiological measure of access to representations in visual working memory. *Psychophysiology*, *47*(1), 197–200.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: bridging selective attention and working memory. *Trends in Cognitive Sciences*, *16*(2), 129–135.

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, *15*(8), 1176–1194.

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. *Journal of Neurophysiology, 103*(4), 1963-8.

Jonides, J., Lacey, S. C., & Nee, D. E. (2005). Processes of Working Memory in Mind and Brain. *Current Directions in Psychological Science*, *14*(1), 2–5.

Katus, T., Andersen, S. K., & Müller, M. M. (2012). Nonspatial cueing of tactile STM causes shift of spatial attention. *Journal of Cognitive Neuroscience*, *24*(7), 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. *Journal of Neuroscience*, *35*(17), 6689–6695.

Katus, T., & Eimer, M. (2016). Multiple foci of spatial attention in multimodal working memory. *NeuroImage*, doi: 10.1016/j.neuroimage.2016.08.019

547 Katus, T., Grubert, A., & Eimer, M. (2015a). Electrophysiological evidence for a sensory recruitment model of somatosensory working memory. *Cerebral Cortex*, *25*(12), 4697–4703.

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

552 Katus, T., Müller, M. M., & Eimer, M. (2015b). Sustained maintenance of somatotopic information in brain regions recruited by tactile working memory. *Journal of Neuroscience*, *35*(4), 1390–1395.

555 Kuo, B. C., Rao, A., Lepsien, J., & Nobre, A. C. (2009). Searching for targets within the spatial layout of visual short-term memory. *Journal of Neuroscience*, *29*(25), 8032–8038.

558 Kuo, B.-C., Stokes, M. G., & Nobre, A. C. (2012). Attention Modulates Maintenance of Representations in Visual Short-term Memory. *Journal of Cognitive Neuroscience*, *24*(1), 51–60.

561 Lepsien, J., & Nobre, A. C. (2006). Cognitive control of attention in the human brain: Insights from orienting attention to mental representations. *Brain Research*, *1105*(1), 20–31.

564 Lepsien, J., & Nobre, A. C. (2007). Attentional modulation of object representations in working memory. *Cerebral Cortex*, *17*(9), 2072–2083.

566 Lepsien, J., Thornton, I., & Nobre, A. C. (2011). Modulation of working-memory maintenance by directed attention. *Neuropsychologia*, *49*(6), 1569–1577.

568 Macaluso, E., Frith, C., & Driver, J. (2000). Selective spatial attention in vision and touch: unimodal and multimodal mechanisms revealed by PET. *Journal of Neurophysiology*, *83*(5), 3062–3075.

Macaluso, E., Frith, C. D., & Driver, J. (2002). Directing attention to locations and to sensory modalities: multiple levels of selective processing revealed with PET. *Cerebral Cortex*, *12*(4), 357–368.

McCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex*, *43*(1), 77–94.

Myers, N. E., Walther, L., Wallis, G., Stokes, M. G., & Nobre, A. C. (2015). Temporal dynamics of attention during encoding versus maintenance of working memory: complementary views from event-related potentials and alpha-band oscillations. *Journal of Cognitive Neuroscience*, *27*(3), 492–508.

Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: Fully automated statistical thresholding for EEG artifact rejection. *Journal of Neuroscience Methods*, *192*(1), 152–162.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*(1), 23–38.

Rerko, L., & Oberauer, K. (2013). Focused, unfocused, and defocused information in working memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(4), 1075–1096.

Robitaille, N., Grimault, S., & Jolicoeur, P. (2009). Bilateral parietal and contralateral responses during maintenance of unilaterally encoded objects in visual short-term memory: evidence from magnetoencephalography. *Psychophysiology*, *46*(5), 1090–1099.

Sathian, K., Lacey, S., Stilla, R., Gibson, G. O., Deshpande, G., Hu, X., Laconte S., & Glielmi, C. (2011). Dual pathways for haptic and visual perception of spatial and texture information. *NeuroImage, 57*(2), 462-75.

Sreenivasan, K. K., Curtis, C. E., & D'Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. *Trends in Cognitive Sciences*, *18*(2), 82–89.

Tenke, C. E., & Kayser, J. (2012). Generator localization by current source density (CSD): implications of volume conduction and field closure at intracranial and scalp resolutions. *Clinical Neurophysiology*, *123*(12), 2328–2345.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, *428*(6984), 751–754. doi:10.1038/nature02466

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748–751.

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503.

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*(7080), 91–95.

Figure Legends

Figure 1 Stimulation procedure and task. A bimodal (tactile-visual) sample set was presented before an auditory retro-cue, which was followed by a bimodal test set. Participants memorized the locations of the tactile sample stimuli (symbolized by black dots) and the colors of the visual sample stimuli on one task-relevant side (left or right, varied across blocks). On each trial, the pitch of the retro-cue indicated whether the memorized tactile or visual stimuli (unpredictably 50%) had to be retained and compared with the test stimulus set.

Figure 2. Behavioral performance, quantified in d-Prime (d'), in the tactile task (red bars) and visual task (green bars), in Session 1 (blue outlines) and Session 2 (brown outlines).

Figure 3. Grand mean CSDs in the early period of the retention delay measured at somatosensory (tCDA, left) and visual ROIs (CDA, right) in Session 1 (blue) and Session 2 (brown). CSDs were recorded contralateral (thick line) and ipsilateral (thin line) to the memorized sample set. The bottom panels show contra-ipsilateral difference waves, with shaded areas indicating 95% within-subject confidence intervals (CIs) for tests against zero (i.e., no lateralized effect). CSDs were collapsed across the factor levels of Cued modality. Note that negativity is plotted downwards, and that different scales were used for somatosensory and visual CSDs (as indicated by the length of length of the y-axes representing ±0.5 mA/m^3). Bar graphs display mean amplitudes of the tCDA/CDA averaged for the time period before neural responses were triggered by the retro-cue (300 to 600 ms after sample onset); error

bars represent 95% CIs for tests against zero. Topographical maps illustrate the scalp distribution of the central tCDA and the posterior CDA components that were elicited during the concurrent maintenance of tactile and visual sample stimuli in Session 1 and Session 2.

Figure 4. Grand mean CSDs measured at somatosensory (left) and visual ROIs (right), in trials in which touch (red) or vision (green) was cued. CSDs were recorded contralateral (thick line) and ipsilateral (thin line) to the memorized sample set, and were collapsed across Session 1 and 2. Note the negativity is plotted downwards, and different scales were used for somatosensory and visual ROIs. The bottom panels show contra- minus ipsilateral difference waves for the tCDA and CDA; shaded areas indicate 95% CIs for tests against zero. Bar graphs display tCDA/CDA mean amplitudes (i.e., contralateral minus ipsilateral amplitude differences, with more negative values reflecting larger tCDA/CDA components) averaged between 800 and 2000 ms after sample onset (i.e., 300 ms after the retro-cue, until the end of the retention delay); error bars represent 95% CIs for tests against zero. Topographical maps illustrate the scalp distribution of the central tCDA and posterior CDA components, for trials where touch (top) or vision (bottom) was cued.

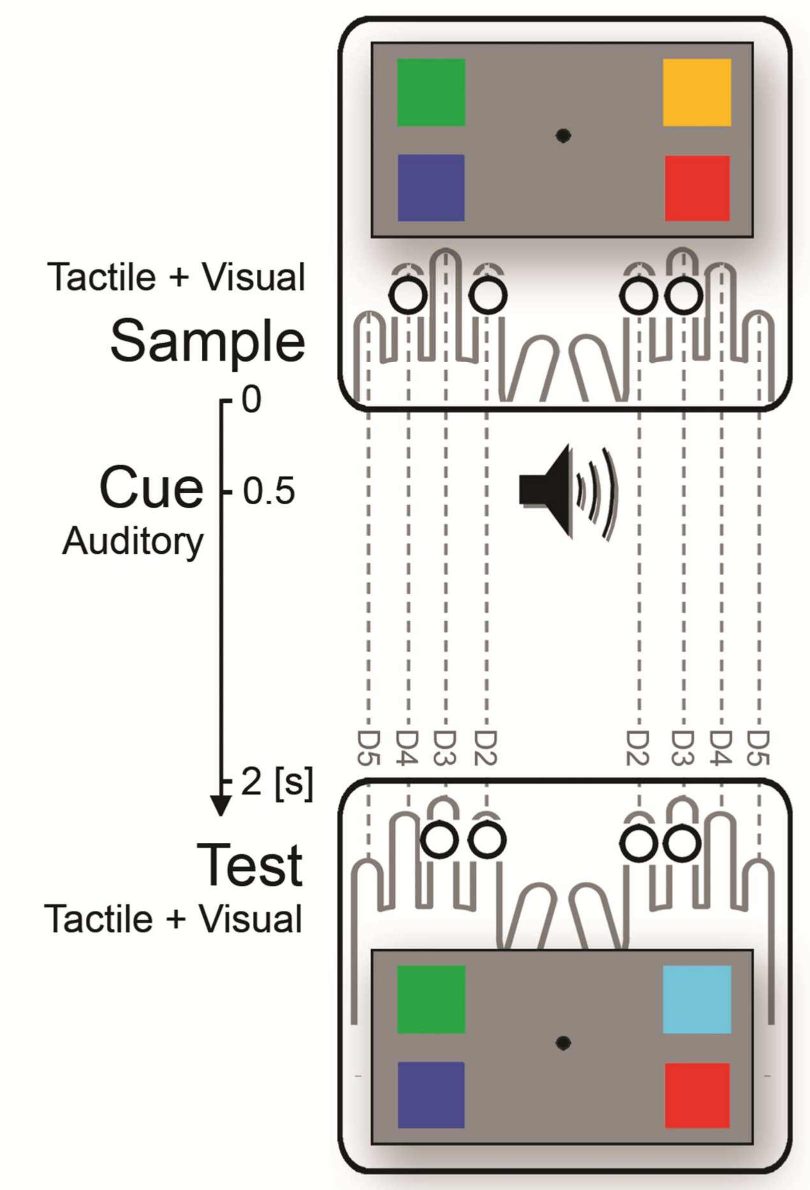

Figure 1 Stimulation procedure and task. A bimodal (tactile-visual) sample set was presented before an auditory retro-cue, which was followed by a bimodal test set. Participants memorized the locations of the tactile sample stimuli (symbolized by black dots) and the colors of the visual sample stimuli on one task-relevant side (left or right, varied across blocks). On each trial, the pitch of the retro-cue indicated whether the memorized tactile or visual stimuli (unpredictably 50%) had to be retained and compared with the test stimulus set.

90x130mm (300 x 300 DPI)

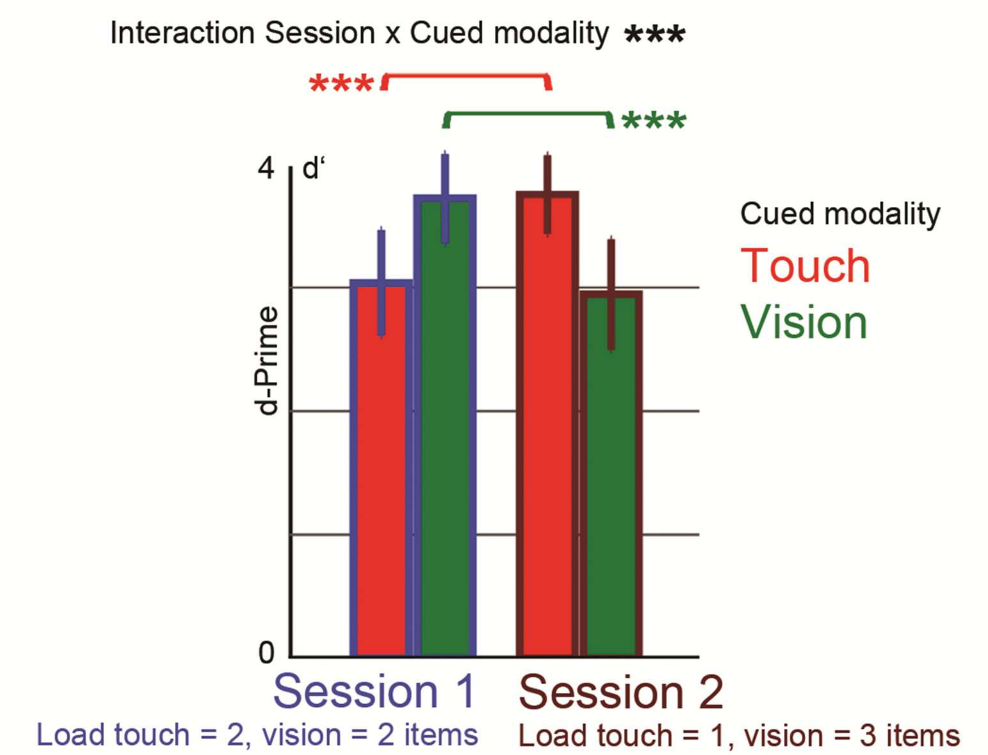

Figure 2. Behavioral performance, quantified in d-Prime (d'), in the tactile task (red bars) and visual task (green bars), in Session 1 (blue outlines) and Session 2 (brown outlines).

90x68mm (300 x 300 DPI)

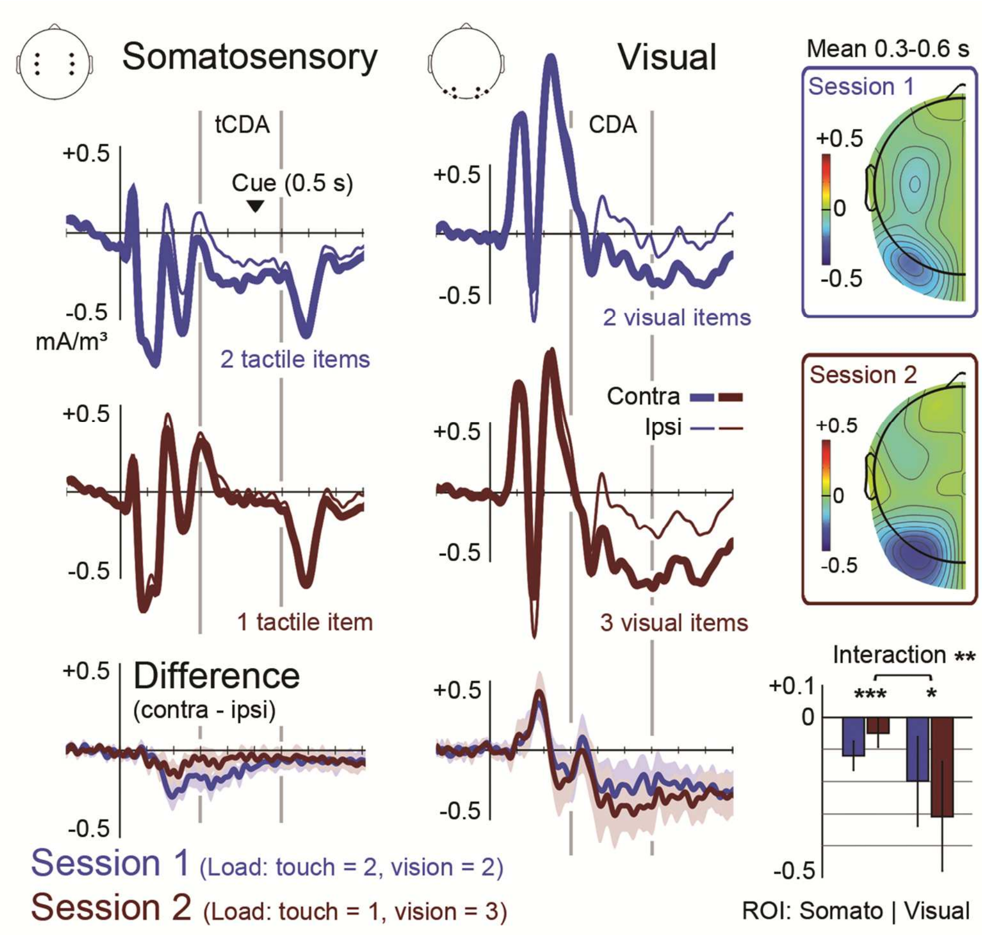

Figure 3. Grand mean CSDs in the early period of the retention delay measured at somatosensory (tCDA, left) and visual ROIs (CDA, right) in Session 1 (blue) and Session 2 (brown). CSDs were recorded contralateral (thick line) and ipsilateral (thin line) to the memorized sample set. The bottom panels show contra-ipsilateral difference waves, with shaded areas indicating 95% within-subject confidence intervals (CIs) for tests against zero (i.e., no lateralized effect). CSDs were collapsed across the factor levels of Cued modality. Note that negativity is plotted downwards, and that different scales were used for somatosensory and visual CSDs (as indicated by the length of length of the y-axes representing ±0.5 mA/m3). Bar graphs display mean amplitudes of the tCDA/CDA averaged for the time period before neural responses were triggered by the retro-cue (300 to 600 ms after sample onset); error bars represent 95% CIs for tests against zero. Topographical maps illustrate the scalp distribution of the central tCDA and the posterior CDA components that were elicited during the concurrent maintenance of tactile and visual sample stimuli in Session 1 and Session 2.

114x108mm (300 x 300 DPI)

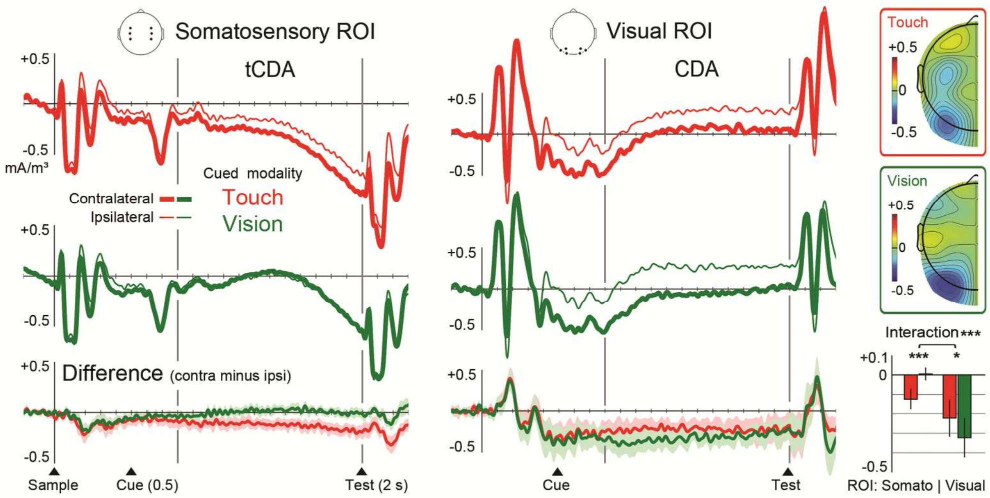

Figure 4. Grand mean CSDs measured at somatosensory (left) and visual ROIs (right), in trials in which touch (red) or vision (green) was cued. CSDs were recorded contralateral (thick line) and ipsilateral (thin line) to the memorized sample set, and were collapsed across Session 1 and 2. Note the negativity is plotted downwards, and different scales were used for somatosensory and visual ROIs. The bottom panels show contra- minus ipsilateral difference waves for the tCDA and CDA; shaded areas indicate 95% CIs for tests against zero. Bar graphs display tCDA/CDA mean amplitudes (i.e., contralateral minus ipsilateral amplitude differences, with more negative values reflecting larger tCDA/CDA components) averaged between 800 and 2000 ms after sample onset (i.e., 300 ms after the retro-cue, until the end of the retention delay); error bars represent 95% CIs for tests against zero. Topographical maps illustrate the scalp distribution of the central tCDA and posterior CDA components, for trials where touch (top) or vision (bottom) was cued.

189x95mm (300 x 300 DPI)