University of Groningen

Unimodal and Bimodal Access to Sensory Working Memories by Auditory and Visual Impulses

Wolff, Michael J; Kandemir, Güven; Stokes, Mark G; Akyürek, Elkan G

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Link to publication in University of Groningen/UMCG research database

Behavioral/Cognitive

Unimodal and Bimodal Access to Sensory Working Memories by Auditory and Visual Impulses

Michael J. Wolff,[1,2] **Güven Kandemir,**[1] **Mark G. Stokes,**[2] and **Elkan G. Akyürek**[1]

[1]Department of Experimental Psychology, University of Groningen, Groningen, 9712 TS, The Netherlands, and [2]Department of Experimental Psychology, University of Oxford, Oxford OX2 6GG, United Kingdom

It is unclear to what extent sensory processing areas are involved in the maintenance of sensory information in working memory (WM). Previous studies have thus far relied on finding neural activity in the corresponding sensory cortices, neglecting potential activity-silent mechanisms, such as connectivity-dependent encoding. It has recently been found that visual stimulation during visual WM maintenance reveals WM-dependent changes through a bottom-up neural response. Here, we test whether this impulse response is uniquely visual and sensory-specific. Human participants (both sexes) completed visual and auditory WM tasks while electroencephalography was recorded. During the maintenance period, the WM network was perturbed serially with fixed and task-neutral auditory and visual stimuli. We show that a neutral auditory impulse-stimulus presented during the maintenance of a pure tone resulted in a WM-dependent neural response, providing evidence for the auditory counterpart to the visual WM findings reported previously. Interestingly, visual stimulation also resulted in an auditory WM-dependent impulse response, implicating the visual cortex in the maintenance of auditory information, either directly or indirectly, as a pathway to the neural auditory WM representations elsewhere. In contrast, during visual WM maintenance, only the impulse response to visual stimulation was content-specific, suggesting that visual information is maintained in a sensory-specific neural network, separated from auditory processing areas.

Key words: EEG; multivariate pattern analysis; sensory working memory

---

Significance Statement

Working memory is a crucial component of intelligent, adaptive behavior. Our understanding of the neural mechanisms that support it has recently shifted: rather than being dependent on an unbroken chain of neural activity, working memory may rely on transient changes in neuronal connectivity, which can be maintained efficiently in activity-silent brain states. Previous work using a visual impulse stimulus to perturb the memory network has implicated such silent states in the retention of line orientations in visual working memory. Here, we show that auditory working memory similarly retains auditory information. We also observed a sensory-specific impulse response in visual working memory, while auditory memory responded bimodally to both visual and auditory impulses, possibly reflecting visual dominance of working memory.

Introduction

Working memory (WM) is necessary to maintain information without sensory input, which is vital to adaptive behavior. De-

spite its important role, it is not yet fully clear how WM content is represented in the brain, or whether sensory information is maintained within a sensory-specific neural network. Previous research has relied on testing whether sensory cortices exhibit content-specific neural activity during maintenance. While this has indeed been shown for visual memories in occipital areas (e.g., Harrison and Tong, 2009) and, more recently, for auditory memories in the auditory cortex (Huang et al., 2016; Kumar et al., 2016; Uluc et al., 2018), WM-specific activity in the sensory cortex is not always present (Bettencourt and Xu, 2016), fueling an ongoing debate over whether sensory cortices are necessary for WM maintenance (Xu, 2017; Scimeca et al., 2018). However, the neural WM network may not be solely based on measurable neu-

ral activity, and it has been proposed that information in WM may be maintained in an "activity-silent" network (Stokes, 2015), for example, through changes in short-term connectivity (Mongillo et al., 2008). Potentially silent WM states should also be taken into account to better investigate the sensory-specificity account of WM.

Silent network theories predict that its neural impulse response to external stimulation can be used to infer its current state (Buonomano and Maass, 2009; Stokes, 2015). This has been shown in visual WM experiments, in which the evoked neural response from a fixed, neutral, and task-irrelevant visual stimulus presented during the maintenance period of a visual WM task contained information about the contents of visual WM (Wolff et al., 2015, 2017). This not only suggests that otherwise hidden processes can be illuminated, but also implicates the involvement of the visual cortex in the maintenance of visual information, even when no ongoing activity can be detected. It has been suggested that this WM-dependent response profile might be not merely a byproduct of connectivity-dependent WM, but a fundamental mechanism that affords efficient and automatic readout of WM content through external stimulation (Myers et al., 2015).

It remains an open question, however, whether information from other modalities in WM is similarly organized. If auditory WM depends on content-specific connectivity changes that include the sensory cortex, we would expect a network-specific neural response to external auditory stimulation. Furthermore, it may be hypothesized that sensory information need not necessarily be maintained in a network that is detached from other sensory processing areas. Direct connectivity (Eckert et al., 2008) and interplay (Martuzzi et al., 2007; Iurilli et al., 2012) between the auditory and visual cortices, or areas where information from different modalities converges, such as the parietal and prefrontal cortices (Driver and Spence, 1998; Stokes et al., 2013), raise the possibility that WM could exploit these connections, even during maintenance of unimodal information. Content-specific impulse responses might be observed not only during sensory-specific but also sensory nonspecific stimulation.

In the present study, we tested whether WM-dependent impulse responses can be observed in visual and auditory WM, and whether that response is sensory specific. We measured EEG while participants performed visual and auditory WM tasks. We show that the evoked neural response of an auditory impulse stimulus reflects relevant auditory information maintained in WM. Visual perturbation also resulted in an auditory WM-dependent neural response, implicating both the auditory and visual cortices in auditory WM. By contrast, visual WM content could only be decoded after visual, but not auditory, perturbation, suggesting that visual information is maintained in a sensory-specific visual WM network with no evidence for a WM-related interplay with the auditory cortex.

Materials and Methods

Participants. Thirty healthy adults (12 female, mean age 21 years, range 18–31 years) were included in the main analyses of the auditory WM experiment and 28 healthy adults (11 female, mean age 21 years, range 19–31 years) of the visual WM experiment. Three additional participants in the auditory WM experiment and 8 additional participants in the visual WM experiment were excluded during preprocessing due to excessive eye movements (>30% of impulse epochs contaminated). The exclusion criterion and resulting minimum number of trials for the multivariate pattern analysis were similar to our previous study (Wolff et al., 2017). Participants received either course credits or monetary compensation (8€ an hour) for participation and gave written informed consent.

Both experiments were approved by the Departmental Ethical Committee of the University of Groningen (approval number 16109-S-NE).

Apparatus and stimuli. Stimuli were controlled by Psychtoolbox, a freely available toolbox for MATLAB. Visual stimuli were generated with Psychtoolbox and presented on a 17-inch (43.18 cm) CRT screen running at 100 Hz refresh rate and a resolution of 1280 × 1024 pixels. Auditory stimuli were generated with the freely available software Audacity and were presented with stereo Logitech computer speakers. The intensity of all tones was adjusted to 70 dB SPL at a fixed distance of 60 cm between speakers and participants in both experiments. All tones had 10 ms ramp up and ramp down time. Responses were collected with a custom two-button response box, connected via a USB interface.

The memory items used in the auditory WM experiment were 8 pure tones, ranging from 270 Hz to 3055 Hz in steps of half an octave. The probes in the auditory experiment were 16 pure tones that were one-third of an octave higher or lower than the corresponding auditory memory items.

The memory items used in the visual WM experiment were 8 sine-wave gratings with orientations of 11.25° to 168.75° in steps of 22.5°. The visual probes were 16 sine-wave gratings that were rotated 20° clockwise or counterclockwise relative to the corresponding visual memory items. All gratings were presented at 20% contrast, with a diameter of 6.5° (at 60 cm distance) and a spatial frequency of 1 cycle per degree. The phase of each grating was randomized within and across trials.

The remaining stimuli were the same in both experiments. The retro-cue was a number (1 or 2) that subtended 0.7°. The visual impulse stimulus was a white circle with a diameter of 12°. The auditory impulse was a complex tone consisting of the combination of all pure tones used as memory items in the auditory task. A gray background (RGB = 128, 128, 128) and a black fixation dot with a white outline (0.25°) were maintained throughout the trials. All visual stimuli were presented in the center of the screen.

Experimental design. The trial structure was the same in both experiments, as shown in Figure 1A, C. In both cases, participants completed a retro-cue WM task. Only the memory items and probes differed between experiments. Memory items and probes were pure tones in the auditory WM task and sine-wave gratings in the visual WM task. Each trial began with the presentation of a fixation dot, which stayed on the screen throughout the trial. After 1000 ms, the first memory item was presented for 200 ms. After a 700 ms delay, the second memory item in the same modality as the first item was presented for 200 ms. Each memory item was selected randomly without replacement from a uniform distribution of 8 different tonal frequencies or grating orientations (see above) for the auditory and visual experiment, respectively. After another delay of 700 ms, the retro-cue was presented for 200 ms, indicating to participants whether the first or second memory item would be tested at the end of the trial. After a delay of 1000 ms the impulse stimuli (the visual circle and the complex tone) were presented serially for 100 ms each with a delay of 900 ms in-between. The order of the impulses was fixed for each participant but counterbalanced between participants. Impulse order was fixed within participants for two reasons: First, it removed the effect of surprise by making the order of events within trials perfectly consistent and predictable (Wessel and Aron, 2017), ensuring minimal intrusion by the impulse stimuli during the maintenance period. Second, random impulse order might have resulted in qualitatively different neural responses of each impulse, depending on when it was presented, due to different trial histories and elapsed maintenance duration at the time of impulse onset (Buonomano and Maass, 2009). This would have necessitated splitting the neural data by impulse order for the decoding analyses, resulting in reduced power. The probe stimulus followed 900 ms after the second impulse offset and was presented for 200 ms. In the auditory WM experiment, the probe was a pure tone and the participant's task was to indicate via button press on the response box whether the probe's frequency was lower (left button) or higher (right button) than the cued memory item. In the visual task, the probe was another visual grating, and the participants indicated whether it was rotated counterclockwise (left button) or clockwise (right button) relative to the cued memory item. The direction of the tone or tilt was selected randomly without replacement from a uniform distribution. After each response, a smiley

face was shown for 200 ms, which indicated whether the response was correct or incorrect. The next trial began automatically after a randomized, variable delay of 700–1000 ms after response input. Each experiment consisted of 768 trials in total and lasted ~2 h.

EEG acquisition and preprocessing. The EEG signal was acquired from 62 Ag/AgCls sintered electrodes laid out according to the extended international 10–20 system. An analog-to-digital TMSI Refa 8–64/72 amplifier and Brainvision recorder software were used to record the data at 1000 Hz using an online average reference. An electrode placed just above the sternum was used as the ground. Bipolar EOG was recorded by electrodes placed above and below the right eye, and to the left and right of the left and right eye, respectively. The impedances of all electrodes were kept <10 kΩ.

Offline the data were downsampled to 500 Hz and bandpass filtered (0.1 Hz high-pass and 40 Hz low-pass) using EEGLAB (Delorme and Makeig, 2004). The data were epoched relative to the onsets of the memory items (−150 ms to 900 ms) and to the onsets of the auditory and visual impulse stimuli (−150 to 500 ms). The signal's variance across channels and trials was visually inspected using a visualization tool provided by the MATLAB extension FieldTrip (Oostenveld et al., 2010), and especially noisy channels were removed and replaced through spherical interpolation. This led to the interpolation of 1 channel in 3 participants and 2 channels in 1 participant in the auditory WM task, and 1 channel in 5 participants and 5 channels in 1 participant in the visual WM task. Noisy epochs were removed from all subsequent electrophysiological analyses. Epochs containing any artifacts related to eye movements were identified by visually inspecting the EOG signals and also removed from analyses. The following percentage of trials were removed for each epoch in the auditory WM experiment: item 1 epoch (mean ± SD, 13.39 ± 6.08%), item 2 epoch (9.28 ± 4.42%), auditory impulse epoch (11.53 ± 7.03%), and visual impulse epoch (9.81 ± 5.44%). The following percentage of trials were removed for each epoch in the visual WM experiment: item 1 epoch (19.81 ± 5.91%), item 2 epoch (20.69 ± 5.88%), auditory impulse epoch (18.51 ± 5.73%), and visual impulse epoch (19.33 ± 4.94%).

Multivariate pattern analysis of neural dynamics. We wanted to test whether the electrophysiological activity evoked by the memory stimuli and impulse stimuli contained item-specific information. Since event-related potentials (ERPs) are highly dynamic, we used an approach that is sensitive to such changing neural activity within predefined time windows, by pooling relative voltage fluctuations over space (i.e., electrodes) and time. This approach has two key benefits: First, pooling information over time (in addition to space) multivariately can boost decoding accuracy (Grootswagers et al., 2017; Nemrodov et al., 2018). Second, by removing the mean-activity level within each time window, the voltage fluctuations are normalized. This is similar to taking a neutral prestimulus baseline, as is common in ERP analysis. Notably, this also removes stable activity traces that do not change within the chosen time window, making this approach ideal to decode transient, stimulus-evoked activation patterns, while disregarding more stationary neural processes. The following details of the analyses were the same for each experiment, unless explicitly stated.

For the time course analysis, we used a sliding window approach that takes into account the relative voltage changes within a 100 ms window. The time points within 100 ms of each channel and trial were first downsampled by taking the average every 10 ms, resulting in 10 voltage values for each channel. Next, the mean activity within that time window of each channel was subtracted from each individual voltage value. All 10 voltage values per channel were then used as features for the eightfold cross-validation decoding approach.

We used Mahalanobis distance (De Maesschalck et al., 2000) to take advantage of the potentially parametric neural activity underlying the processing and maintenance of orientations and tones. The distances between each of the left-out test-trials and the averaged, condition-specific patterns of the train trials (tones and orientations in the auditory and visual experiment, respectively), were computed, with the covariance matrix estimated from the train trials using a shrinkage estimator (Ledoit and Wolf, 2004). To acquire reliable distance estimates, this process was repeated 50 times, where the data were randomly partitioned into 8 folds using stratified sampling each time. The number of trials of each condition (orientation/tone frequency) of the 7 train-folds were equalized by randomly subsampling the minimum number of condition-specific trials to ensure an unbiased training set. The average was then taken of these repetitions. For each trial, the 8 distances (one of each stimulus condition) were sign-reversed for interpretation purposes, so that higher values reflect higher pattern similarity between test and train trials. For visualization, the sign-reversed distances were furthermore mean-centered by subtracting the mean distance of all distances of a given trial and ordered as a function of tone difference, in 1 octave steps by averaging over adjacent half-octave differences, and orientation differences.

To summarize the expected positive relationship between tone similarity and neural activation similarity (indicative of tone-specific information in the recorded signal) into a single value in the auditory WM experiment, the absolute tonal differences were linearly regressed against the corresponding pattern similarity values for each trial. The obtained β values of the slopes were then averaged across all trials to represent "decoding accuracy," where high values suggest a strong positive effect of tone similarity on neural pattern similarity. To summarize the tuning curves in the visual WM experiment, we computed the cosine vector means (Wolff et al., 2017), where high values suggest evidence for orientation decoding.

The approach described above was repeated in steps of 8 ms across time (−52 to 900 ms relative to item 1 and 2 onset, and −52 to 500 ms relative to auditory and visual onset). The decoding values were averaged over trials, and the decoding time course was smoothed with a Gaussian smoothing kernel (SD 16 ms). Within the time window, information was pooled from −100 to 0 ms relative to a specific time point. By only including data points from before the time point of interest, it is ensured that decoding onsets can be more easily interpreted, whereas decoding offsets should be interpreted with caution (Grootswagers et al., 2017). In addition to the sliding window approach, we also pooled information multivariately across the whole time window of interest (Nemrodov et al., 2018). As before, the data were first downsampled by taking the average every 10 ms, and the mean activity from 100 to 400 ms relative to impulse onset was subtracted. The resulting 30 values per channel were then provided to the multivariate decoding approach in the same way as above, resulting in a single decoding value per participant. The time window of interest was based on previous findings showing that the WM-dependent impulse response is largely confined within that window (Wolff et al., 2017). Additionally, items in the item-presentation epochs were also decoded using each channel separately, using the data from 100 to 400 ms relative to onset. Decoding topographies were visualized using FieldTrip (Oostenveld et al., 2010).

Cross-epoch generalization analysis. We also tested whether WM-related decoding in the impulse epochs generalized to the memory presentation. Instead of using the same epoch (100–400 ms) for training and testing, as described above, the classifier was trained on the memory item epoch and tested on the impulse epoch that contained significant item decoding (and vice versa). In the auditory task, we also tested whether the different impulse epochs cross-generalized by training on the visual and testing on the auditory impulse (and vice versa).

Representational similarity analysis (RSA). While the decoding approach outlined above takes into account the potentially parametric relationship of pitch/orientation difference, it is not an explicit test for the presence of a parametric relationship. Indeed, decodability could theoretically be solely driven by high within stimulus-condition pattern similarity, and equally low pattern similarities of all between stimulus-condition comparisons. To explicitly test for a linear/circular relationship between stimuli, and explore additional stimulus coding schemes, we used RSA (Kriegeskorte et al., 2008).

The RSA was based on the Mahalanobis distances between all stimulus conditions (unique orientations and frequencies) in both experiments using the same time window of interest as in the decoding approach described above (100–400 ms relative to stimulus onset). For each participant, the number of trials of each stimulus condition were equalized by randomly subsampling the minimum number of trials of a condition before taking the average across all same stimulus condition trials and
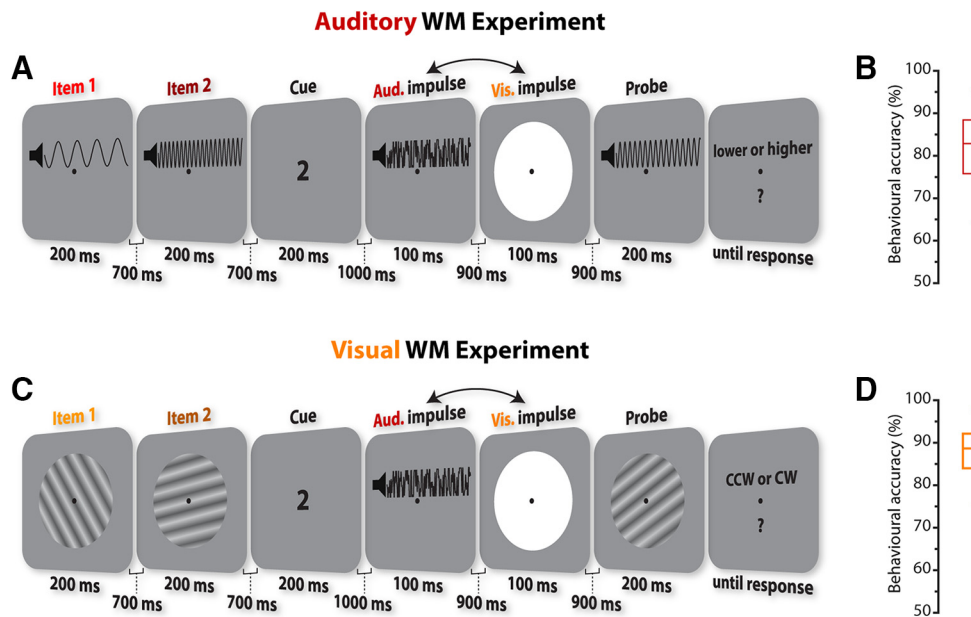
Auditory WM Experiment

Visual WM Experiment

Figure 1. Task structure and behavioral performance. **A**, Trial schematic of auditory task. Two randomly selected pure tones (270 –3055 Hz) were serially presented, and a retro-cue indicated which of those tones would be tested at the end of the trial. In the subsequent delay, two irrelevant impulse stimuli (a complex tone and a white circle) were serially presented. At the end of each trial, another pure tone was presented (the probe), and participants were instructed to indicate whether the frequency of the previously cued tone was higher or lower than the probe's frequency. **B**, Boxplot represents auditory task accuracy. Middle line indicates the median. Box outlines indicate 25th and 75th percentiles. Whiskers indicate 1.5× the interquartile range. Superimposed circles represent mean. Error bars indicate 95% CI. **C**, Trial schematic of visual task. The trial structure was the same as in the auditory task. Instead of pure tones, memory items were randomly oriented gratings. The probe was another orientation grating, and participants were instructed to indicate whether the cued item's orientation was rotated clockwise or counterclockwise relative to the probe's orientation. **D**, Visual task performance.

computing all pairwise Mahalanobis distances. This procedure was repeated 50 times, with random subsamples each time, before averaging them all into a single representation dissimilarity matrix (RDM). The covariance matrix was computed from all trials using the shrinkage estimator (Ledoit and Wolf, 2004). Since each experiment contained 8 unique memory items, this resulted in an 8 × 8 RDM for each participant and epoch of interest.

For the RSA in the auditory WM experiment, we considered two models: a positive linear relationship between absolute pitch height difference (i.e., the more dissimilar pitch frequency, the more dissimilar the brain activity patterns), and a positive relationship of pitch chroma (i.e., higher similarity between brain activity patterns of the same pitch chromas). The tone frequencies used in the experiment increased in half-octave steps. Every other tone thus had the same pitch chroma (i.e., the same note in a different octave). The model RDMs are shown for illustration in Figure 4A. The model RDMs were z-scored to make the corresponding model fits between them more comparable, before entering both of them into a multiple regression analysis with the data RDM.

In the visual WM experiment, we also considered two models. The first model was designed to capture the circular relationship between absolute orientation difference (i.e., the more dissimilar the orientation, the more dissimilar the brain activity patterns). The second model was designed to capture the specialization of cardinal orientations (i.e., horizontal and vertical) that could reflect the "oblique effect," where orientations close to the cardinal axes are discriminated and recalled more accurately than more oblique orientations (Appelle, 1972; Pratte et al., 2017). The model assumed the extreme case, where orientations are clustered into one of three categories depending on their circular distance to vertical, horizontal, or oblique angles. This captures the relatively higher dissimilarity and distinctiveness of the cardinal axes (vertical and horizontal) compared with the oblique axes (−45 degrees and 45 degrees) and reflects neurophysiological findings of an increased number of neurons tuned to the cardinal axes (Shen et al., 2014). The model RDMs are shown for illustration in Figure 4D. The model RDMs were also z-scored and then both included into a multiple regression with the data RDM.

Statistical analysis. All statistical tests were the same between experiments. Sample sizes of all analyses were n = 30 and n = 28 in the auditory

and visual tasks, respectively. Sample size of the ERP analyses as a function of impulse modality and task was n = 16, as it only included participants who participated in both WM tasks. To determine whether the decoding values (see above) or model fits of the RSA are >0 or different between items, or whether the evoked potentials were different between tasks, we used a nonparametric sign-permutation test (Maris and Oostenveld, 2007). The sign of the decoding value, model fit value, or voltage difference of each participant were randomly flipped 100,000 times with a probability of 50%. The p value was derived from the resulting null distribution. The above procedure was repeated for each time point for time-series results. A cluster-based permutation test (100,000 permutations) was used to correct for multiple comparisons over time using a cluster forming and cluster significance threshold of p < 0.05. Complementary Bayes factors to test for decoding evidence for the cued and uncued items within each impulse epoch separately were also computed.

We were also interested whether there were differential effects on the decoding results between cueing (cued/uncued) and impulse modality (auditory/visual) during WM maintenance. To test this, we computed the Bayes factors of models with and without each of these predictors versus the null model that only included subjects as a predictor (Bayesian equivalent of repeated-measures ANOVA). The freely available software package JASP (JASP Team, 2018) was used to compute Bayes factors.

Differences in behavioral performance between tasks were tested with the partially overlapping samples t test (Derrick et al., 2017), since only some participants took part in both tasks. No violations of normality or equality of variances were detected.

Error bars for visualization are 95% confidence intervals (CI), that were computed by bootstrapping from the data in question 100,000 times.

Code and data availability. All data and custom MATLAB scripts used to generate the results and figures of this manuscript are available from the OSF database (osf.io/u7k3q).

Results
Behavioral results
Behavioral task performance was (mean ± SD) 82.322 ± 8.841% in the auditory WM task (Fig. 1B), and 87.908 ± 6.374% in the
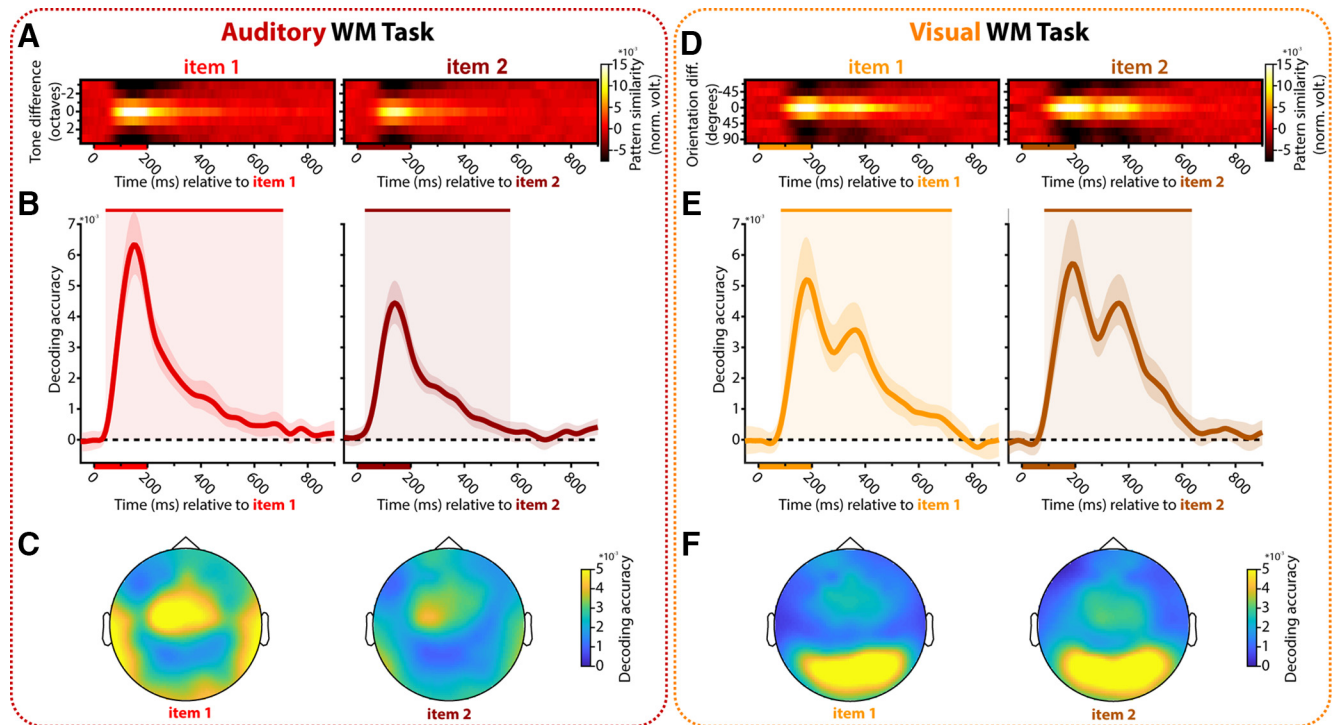
Figure 2. Decoding during item encoding. *A*–*C*, Auditory WM task. *D*–*F*, Visual WM task. *A*, *D*, Normalized average pattern similarity (mean-centered, sign-reversed Mahalanobis distance) of the neural dynamics for each time point between trials as a function of tone similarity in *A* and orientation similarity in *D*, separately for item 1 and item 2, in item 1 and item 2 epochs, respectively. Bars on the horizontal axes represent item presentations. *B*, *E*, Beta values in *B* and cosine vector means in *E* of pattern similarities for items 1 and 2. Upper bars and corresponding shading represent significant values. Error shading represents 95% CI of the mean. *C*, *F*, Topographies of each item of channelwise decoding (100 – 400 ms relative to item onset).

visual WM task (Fig. 1*D*). Performance was significantly higher in the visual than in the auditory task, $t(33.379) = 2.776$, $p = 0.009$, two-sided. Despite this difference, it is clear that participants performed well above chance in both tasks, suggesting that the relevant sensory features were reliably remembered and recalled.

Decoding visual and auditory stimuli
Auditory WM task
The neural dynamics of auditory stimulus processing suggest a parametric effect, with a positive relationship between tone and pattern similarity (Fig. 2*A*) for both memory items. The neural dynamics showed significant item-specific decoding clusters during, and shortly after, corresponding item presentation for item 1 (44–708 ms relative to item 1 onset, $p < 0.001$, one-sided, corrected) and item 2 (28–572 ms relative to item 2 onset, $p < 0.001$, one-sided, corrected; Fig. 2*B*). The topographies of channelwise item decoding for each item using the neural data from 100 to 400 ms after item onset, revealed strong decoding for frontal-central and lateral electrodes (Fig. 2*C*), suggesting that the tone-specific neural activity is most likely generated by the auditory cortex (Chang et al., 2016). These results provide evidence that stimulus-evoked neural activity fluctuations contain information about presented tones that can be decoded from EEG.

Visual WM task
Processing of visual orientations also showed a parametric effect (Fig. 2*D*), replicating previous findings (Saproo and Serences, 2010). The item-specific decoding time courses of the dynamic activity showed significant decoding clusters during and shortly after item presentation (item 1: 84–724 ms, $p < 0.001$; item 2: 84–636 ms, $p < 0.001$, one-sided, corrected; Fig. 2*E*). As expected, the topographies of channelwise item-decoding showed

strong effects in posterior channels that are associated with the visual cortex (Fig. 2*F*).

Content-specific impulse responses
Auditory WM task
In the auditory impulse epoch, the neural dynamics time course revealed significant cued-item decoding (180–308 ms, $p = 0.004$, one-sided, corrected), while no clusters were present for the uncued item (Fig. 3*A*,*B*, left). Similarly, the cued item was decodable in the visual impulse epoch (204–372 ms, $p = 0.009$, one-sided, corrected), while the uncued item was not (Fig. 3*A*,*B*, right).

The time-of-interest (100–400 ms relative to impulse onset) analysis provided similar results. The cued item showed strong decoding in both impulse epochs (auditory impulse: Bayes factor = 11,462.607, $p < 0.001$; visual impulse: Bayes factor = 85.843, $p < 0.001$, one-sided), but the uncued item did not (auditory impulse: Bayes factor = 0.968, $p = 0.075$; visual impulse: Bayes factor = 0.204, $p = 0.476$, one-sided; Fig. 3*C*). A model only including the cueing predictor yielded the highest Bayes factor of 8.123 (± 0.996%) compared with the null model. A model including impulse modality as a predictor resulted in a Bayes factor of 0.848 (± 1.075%). Including both predictors (impulse modality and cueing) in the model resulted in a Bayes factor of 7.553 (± 0.991%) that was slightly lower than only including cueing.

Together, these results provided strong evidence that both impulse stimuli elicit neural responses that contain information about the cued item in auditory WM, but none about the uncued item.
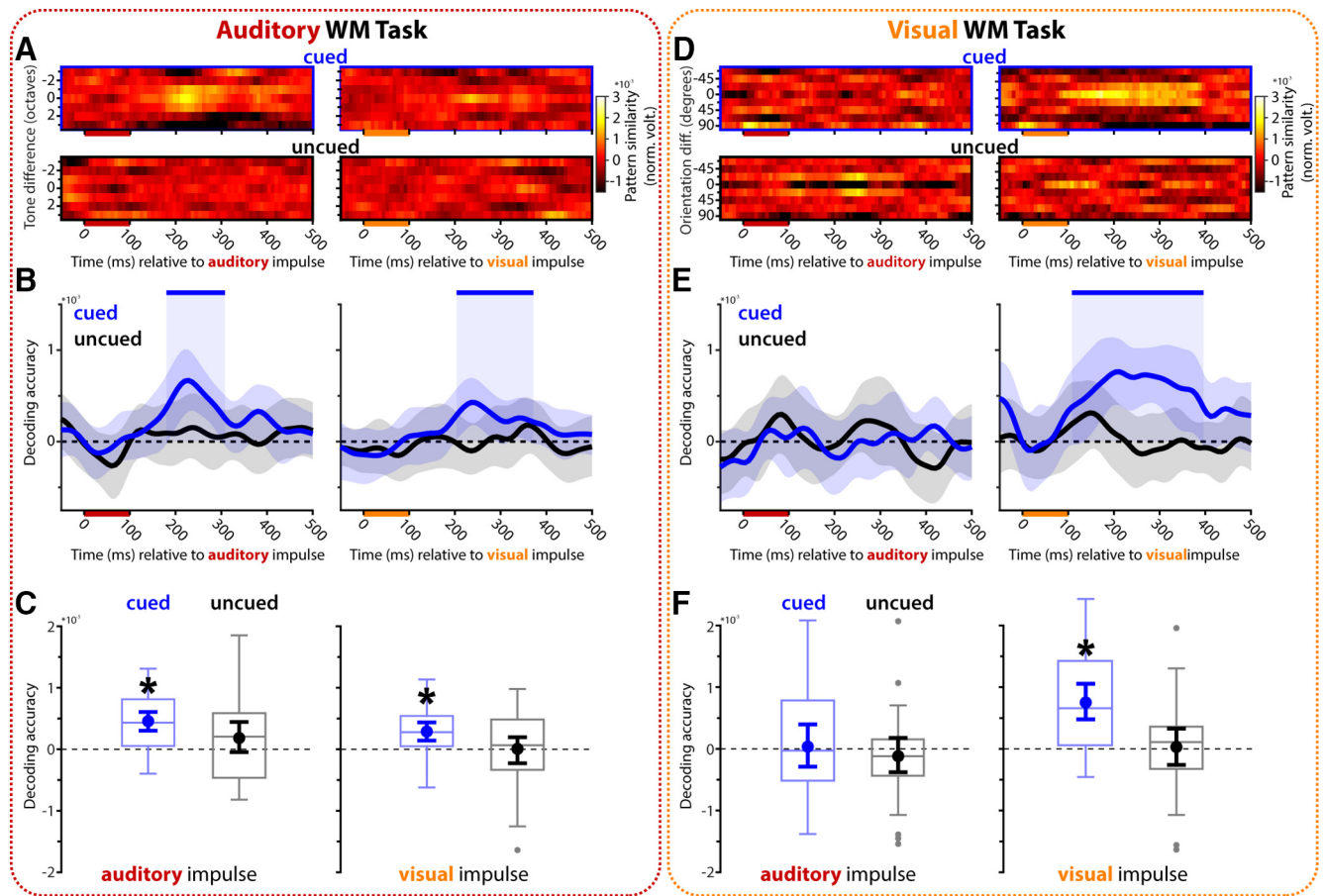
Figure 3. Decoding auditory and visual WM content from the impulse response. **A–C**, Auditory WM task. **D–F**, Visual WM task. **A, D**, Normalized average pattern similarity (mean-centered, sign-reversed Mahalanobis distance) of the neural dynamics for each time point between trials as a function of tone similarity in **A** and orientation similarity in **D**. Top row, Cued item; bottom row, uncued item; left column, auditory impulse; right column, visual impulse. **B, E**, Decoding accuracy time course: Beta values in **B** and cosine vector means in **E** of pattern similarities for cued (blue) and uncued item (black). Upper bars and shading represent significant values of the corresponding item. Error shading represents 95% CI of the mean. **C, F**, Boxplots represent the overall decoding accuracies for the cued (blue) and uncued (black) item, using the whole time window of interest (100–400 ms relative to onset) from the auditory (left) and visual (right) impulse epoch. Middle lines indicate the median. Box outlines indicate 25th and 75th percentiles. Whiskers indicate 1.5× the interquartile range. Extreme values are shown separately (dots). Superimposed circles represent mean. Error bars indicate 95% CI. *$p < 0.05$, significant decoding accuracies (one-sided).

Visual WM task

No significant time clusters were present in the auditory impulse epoch of the visual WM experiment for either the cued or the uncued item task (Fig. 3 D,E, left). The decoding time course of the visual impulse epoch revealed a significant decoding cluster of the cued item (108–396 ms, $p < 0.001$, one-sided, corrected) but not for the uncued item (Fig. 3 D,E, right), replicating previous findings (Wolff et al., 2017).

The analysis on the time-of-interest interval (100–400 ms) showed the same pattern of results; neither the cued nor uncued item in the auditory impulse epoch showed >0 decoding (cued: Bayes factor = 0.236, $p = 0.417$; uncued: Bayes factor = 0.119, $p = 0.787$, one-sided). In the visual impulse epoch, the cued item showed strong decodability (Bayes factor = 1695.823, $p < 0.001$, one-sided), but the uncued item did not (Bayes factor = 0.236, $p = 0.421$, one-sided; Fig. 3F). A model including both predictors (cueing and impulse modality) as well as their interaction resulted in the highest Bayes factor compared with the null model (Bayes factor = 56.284 ± 1.557%). Models with each predictor alone resulted in notably smaller Bayes factors (cueing: Bayes factor = 6.26 ± 0.398%; impulse modality: Bayes factor = 5.877 ± 0.686%). The Bayes factor of the model including both predictors without interaction (46.728 ± 0.886%) was only 1.205 times smaller than the model that also included the interaction,

highlighting that, while there was strong evidence in favor of both impulse modality and cueing, there was only weak evidence in favor of an interaction.

Overall, these results provided evidence that while a visual impulse clearly evokes a neural response that contains information about the cued visual WM item, replicating previous findings (Wolff et al., 2017), an auditory impulse does not.

Parametric encoding and maintenance of auditory pitch and visual orientation

As indicated, RSA was performed to explicitly test and explore for specific stimulus coding relationships in both experiments (Fig. 4 A,D).

Auditory WM task

The RDMs of each epoch of interest are shown in Figure 4B. There was strong evidence in favor of the pitch height difference model during item encoding (item 1 and item 2 presentation epochs; Bayes factor > 100,000, $p < 0.001$, one-sided), whereas evidence against the pitch chroma model was evident (Bayes factor = 0.177, $p = 0.523$, one-sided; Fig. 4 B,C, left). Moderate evidence in favor of the pitch height model was also evident for the cued item in the auditory impulse epoch (Bayes factor = 4.016, $p = 0.0113$, one-sided), whereas there was weak evidence
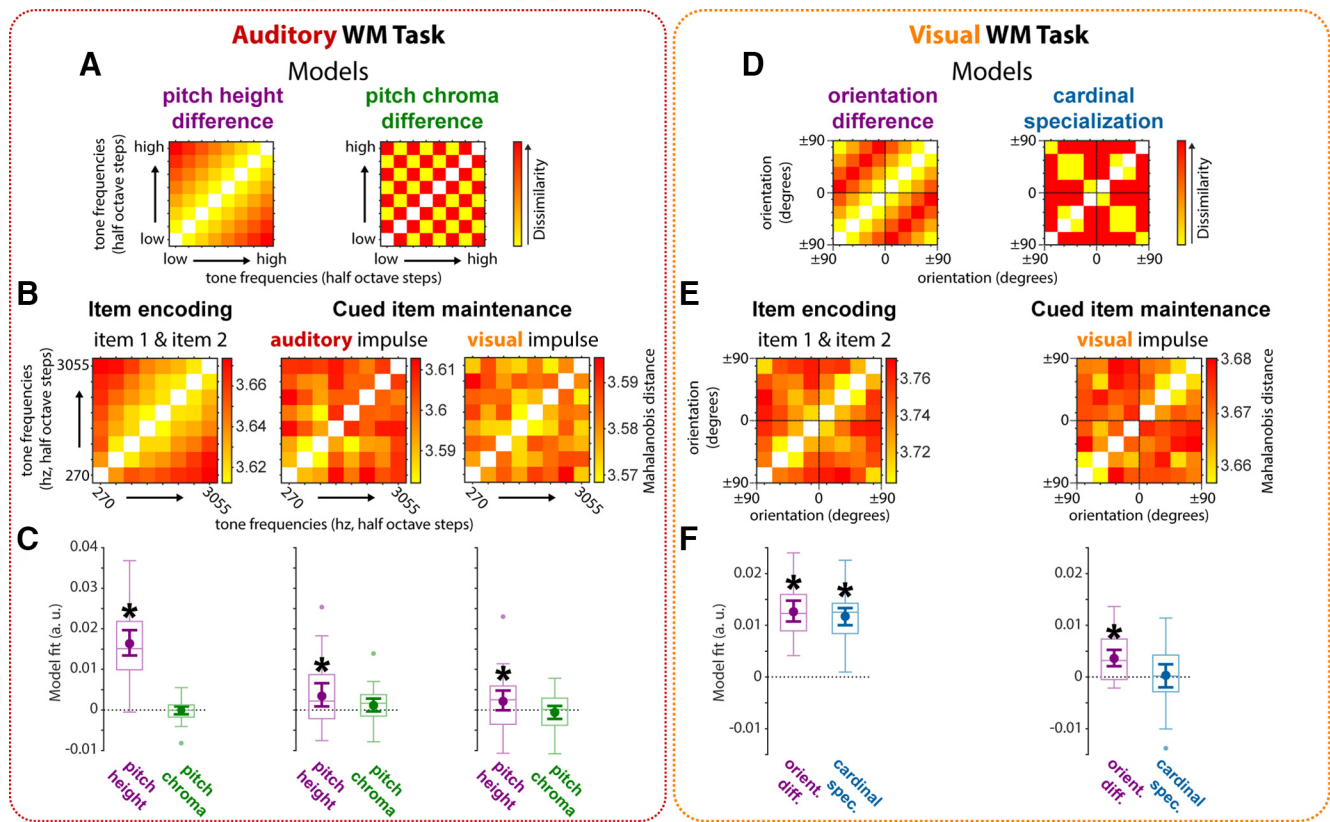
Figure 4. Stimulus coding relationship during encoding and maintenance. ***A–C***, Auditory WM task. ***D–F***, Visual WM task. ***A, D***, Model RDMs of pitch (***A***) and orientation (***D***). ***B, E***, Data RDMs. ***C, F***, Model fits of model RDMs on data RDMs. Middle lines indicate the median. Box outlines indicate 25th and 75th percentiles. Whiskers indicate 1.5× the interquartile range. Extreme values are shown separately (dots). Superimposed circles represent mean. Error bars indicate 95% CI. *$p < 0.05$, significant model fits (one-sided).

against the pitch chroma model (Bayes factor = 0.838, $p = 0.079$, one-sided; Fig. 4B,C, middle). The visual impulse epoch also suggested a pitch height coding model of the cued auditory item, although the evidence was weak (Bayes factor = 1.346, $p = 0.049$, one-sided), and there was again evidence against the pitch chroma model of the cued item (Bayes factor = 0.123, $p = 0.736$, one-sided; Fig. 4B,C, right).

Overall, these RSA results provide evidence that both the encoding and maintenance of pure tones are coded parametrically according to pitch height (Uluc et al., 2018), but not pitch chroma.

Visual WM task
The RDMs of the averaged encoding epochs (item 1 and item 2) and the visual impulse epoch are shown in Figure 4E. There was strong evidence in favor for a circular orientation difference code (Bayes factor > 100,000, $p < 0.001$, one-sided), as well as an additional "cardinal specialization" code (Bayes factor > 100,000, $p < 0.001$, one-sided) during item encoding (Fig. 4E,F, left). The evoked neural response by the visual impulse also provided strong evidence for a circular orientation difference code for the maintenance of the cued item (Bayes factor = 362.672, $p < 0.001$, one-sided). No evidence in favor of an additional "cardinal specialization" code during maintenance was found, however (Bayes factor = 0.252, $p = 0.318$, one-sided; Fig. 4E,F, right).

These results provide evidence that orientations are encoded and maintained in a parametric, orientation selective code (e.g., Ringach et al., 2002; Saproo and Serences, 2010). We additionally considered the "cardinal specialization" coding model, which

captures the expected increased neural distinctiveness of horizontal and vertical orientations compared with tilted orientations, based on the superior visual discrimination of cardinal orientations (Appelle, 1972) as well as previous neurophysiological reports of cardinal specialization (Li et al., 2003; Shen et al., 2014). Evidence for this model was only found during orientation encoding, but not maintenance.

No WM-specific cross-generalization between impulse and WM-item presentation
It has been shown previously that the visual WM-dependent impulse response does not cross-generalize with visual item processing (Wolff et al., 2015). Here we tested whether this is also the case for auditory WM, and additionally explored the cross-generalizability between impulses.

Auditory WM task
The representation of the cued item did neither cross-generalize between item presentation and either of the impulse epochs (auditory impulse: Bayes factor = 0.225, $p = 0.58$; visual impulse: Bayes factor = 0.356, $p = 0.26$, two-sided), nor between impulse epochs (Bayes factor = 0.267, $p = 0.417$, two-sided; Fig. 5A).

Visual WM task
Replicating previous reports (Wolff et al., 2015, 2017), the visual impulse response of the cued visual item did not cross-generalize with item processing during item presentation (Bayes factor = 0.491, $p = 0.168$, two-sided; Fig. 5B).
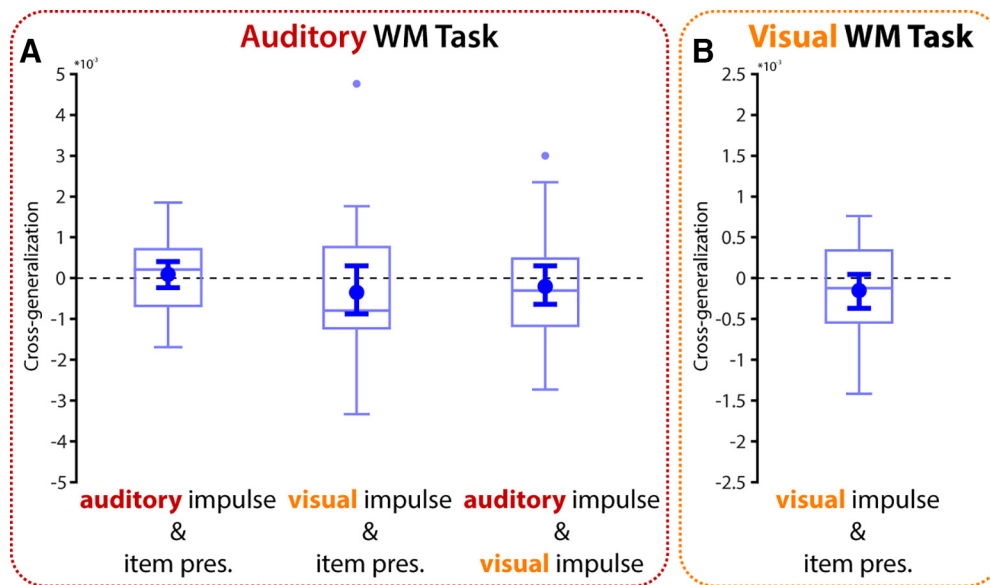
Figure 5. Cross-generalization between epochs. **A**, Cross-generalization of the cued item between the memory item epoch and impulse epochs in the auditory WM task. **B**, Cross-generalization between visual impulse and memory item in the visual WM task. Middle lines indicate the median. Box outlines indicate 25th and 75th percentiles. Whiskers indicate 1.5× the interquartile range. Extreme values are shown separately (dots). Superimposed circles represent mean. Error bars indicate 95% CI.

Evoked response magnitudes of impulse stimuli are comparable between tasks

Since the impulse stimuli were always the same across trials, presented at the same relative time within each trial, and were completely task irrelevant, we believe that the WM-specific impulse responses reported here and in previous work rely on low-level interactions of the impulse stimuli with the WM network, which do not depend on higher-order cognitive processing of the impulse.

Nevertheless, it could be argued that the impulse stimuli are differentially processed, even at an early stage between the WM tasks. Since the auditory impulse was the only auditory stimulus in the visual WM task, it may have been more easily filtered out and ignored compared with the other impulse stimuli. Indeed, it is possible that the neural response to the auditory impulse stimulus was just too "weak" to result in a measurable, WM-specific neural response in the visual WM task. However, given the uniqueness of the auditory impulse in the visual WM task, the opposite could be argued as well.

To test for potential differences of attentional filtering of impulse stimuli between tasks, we examined the ERPs to the impulse stimuli in both tasks from electrodes associated with sensory processing (Fz, FCz, and Cz for auditory impulse; O1, Oz, and O2 for visual impulse). If there is indeed a difference in early sensory processing, this should be visible in associated early evoked responses within 250 ms of stimulus presentation (Luck et al., 2000; Boutros et al., 2004). Because ERPs are subject to large individual differences, only participants who participated in both tasks (n = 16) were included in this analysis.

We also considered potential voltage differences between tasks from 250 to 500 ms postimpulse onsets to test. This is the expected time range of the P3 ERP component and its two subcomponents, the P3a and the P3b, which have been linked to the attentional processing of rare and unpredictable nontargets, and the processing (including memory consolidation) of target stimuli, respectively (Squires et al., 1975; Polich, 2007). The presence of these components would imply that higher-order cognitive processes may be involved in the processing of the impulses,

despite their regularity and task irrelevance. To explore whether the impulses elicited these endogenous components and test for potential differences between tasks, we considered the average voltages from channels Fz, FCz, and Cz for the P3a, and the average voltage from Pz for the P3b (Conroy and Polich, 2007).

Auditory ERPs

The early auditory ERP evoked from the auditory impulse stimulus within each task is shown in Figure 6A (left). The P50, N1, and P2 components, all of which have been shown to be reduced when irrelevant auditory stimuli are filtered out (sensory gating) (Kisley et al., 2004; e.g., Boutros et al., 2004; Cromwell et al., 2008), can clearly be identified in both tasks. One time cluster of the difference between tasks was significant within the time window of interest (148–184 ms, p = 0.048, two-sided, corrected). Visual inspection of the ERPs suggests that, while there is no difference in P50 and N1 amplitude between tasks, P2 amplitude is larger in the visual than in the auditory task. This difference goes in the opposite direction as would be expected if the auditory impulse stimulus was somehow more easily filtered out and ignored in the visual than in the auditory task.

The late ERP elicited by the auditory impulse stimuli in both tasks in shown in Figure 6A (right). Visual inspection of the voltage traces suggests that no clear P3a or P3b components are evident, although it could be argued that the upward inflection at 300 ms in the frontal/central electrodes hints at a small P3a component (Fig. 6A, bottom right). Nevertheless, no significant time clusters in the difference between the auditory and the visual WM task were found in the time window of interest in either voltage trace (p > 0.19, two-sided, corrected).

Visual ERPs

The early visual impulse ERP recorded from occipital electrodes is shown in Figure 6B (left). Early components of interest (C1, P1, N1), which have been shown to be modulated by attentional processes (Luck et al., 2000; e.g., Di Russo et al., 2003; Rauss et al., 2009), have been marked. Visual inspection suggests that there are no discernible differences in these visual components be-
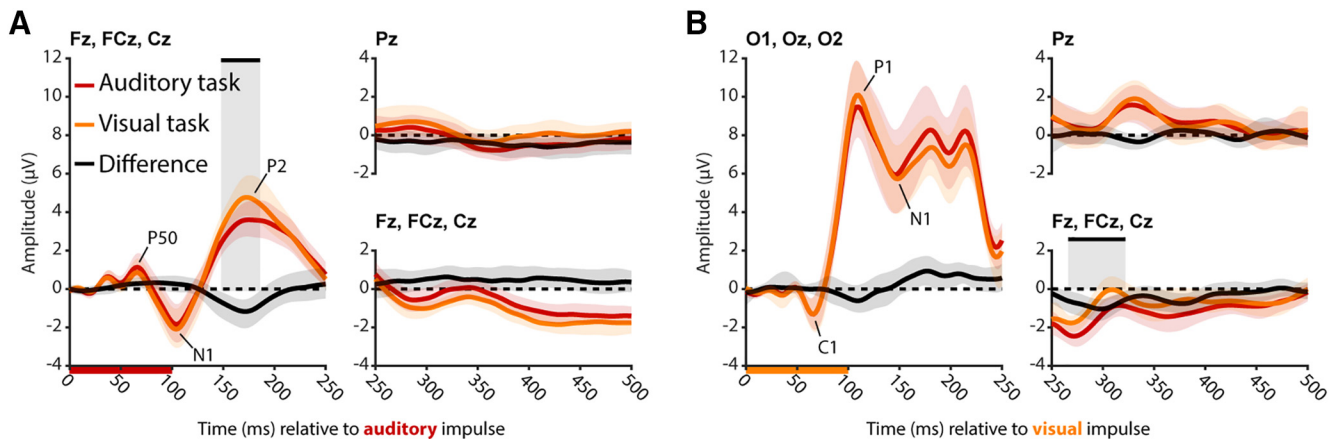
Figure 6. Evoked responses to impulse stimuli as a function of task for participants who participated in both tasks (*n* = 16). ***A***, Average voltages evoked by auditory impulse in the auditory task (red) and visual task (orange). Black represents difference voltage (auditory task − visual task). Individual ERP components of interest are labeled. Error shadings represent 95% CI of the mean. Black bar represents the significant time cluster of difference (*p* < 0.05, corrected, two-sided). ***B***, Average voltages evoked by the visual impulse. Same convention as in ***A***.

tween tasks. Indeed, no significant time clusters were found (*p* > 0.19, two-sided, corrected), suggesting that the visual impulse stimulus was processed similarly between tasks.

The late ERP in response to the visual impulse stimuli is shown in Figure 6*B* (right). One significant time cluster of the difference of the voltage traces between tasks was found in the frontal/central electrodes (266–322 ms, *p* = 0.023, two-sided, corrected; Fig. 6*B*, bottom right). Visual inspection suggests that this could be due to a higher P3a amplitude in the visual than in the auditory task, implying that the visual impulse elicited more attentional processes. However, due to the generally small amplitude, a clear conclusion on what caused this difference cannot be drawn. The visual impulse stimulus resulted in WM-specific responses in both tasks, so the observed voltage difference does not reconcile those findings. No time clusters were found in the voltage difference between tasks on the posterior electrode (Fig. 6*B*, top right).

Discussion

It has been shown that the bottom-up neural response to a visual impulse presented during the delay of a visual WM task contains information about relevant visual WM content (Wolff et al., 2015, 2017), which is consistent with WM theories that assume information is maintained in activity-silent brain states (Stokes, 2015). We used this approach to investigate whether sensory information is maintained within sensory-specific neural networks, shielded from other sensory processing areas. We show that the neural impulse response to sensory-specific stimulation is WM content-specific not only in visual WM, but also in auditory WM, demonstrating the feasibility and generalizability of the approach in the auditory domain. Furthermore, for auditory WM, a content-specific response was obtained not only during auditory, but also during visual stimulation, suggesting a sensory modality-unspecific path to access the auditory WM network. In contrast, only visual, but not auditory, stimulation evoked a neural response containing relevant visual WM content. This pattern of impulse responsivity supports the idea that visual pathways may be more dominant in WM maintenance.

Recent studies have shown that delay activity in the auditory cortex reflects the content of auditory WM (Huang et al., 2016; Kumar et al., 2016; Uluc et al., 2018). Thus, similar to visual WM maintenance, which has been found to result in content-specific delay activity in the visual cortex (Harrison and Tong, 2009),

auditory WM content is also maintained in a network that recruits the same brain area responsible for sensory processing. However, numerous visual WM studies have shown that content-specific delay activity may in fact reflect the focus of attention (Lewis-Peacock et al., 2012; Watanabe and Funahashi, 2014; Sprague et al., 2016). The memoranda themselves may instead be represented within connectivity patterns that generate a distinct neural response profile to internal or external neural stimulation (Lundqvist et al., 2016; Rose et al., 2016; Wolff et al., 2017). While previous research has focused on visual WM, we now provide evidence for a neural impulse response that reflects auditory WM content, suggesting a similar neural mechanism for auditory WM.

The neural response to a visual impulse stimulus also contained information about the behaviorally relevant pitch. It has been shown that visual stimulation can result in neural activity in the auditory cortex (Martuzzi et al., 2007; Morrill and Hasenstaub, 2018). Thus, direct connectivity between visual and auditory areas (Eckert et al., 2008) might be such that visual stimulation activates auditory WM representations in auditory cortex, providing an alternate access pathway. Alternatively, visual cortex itself might retain auditory information. It has been shown that natural sounds can be decoded from the activity in the visual cortex during processing and imagination (Vetter et al., 2014). Even though pure tones were used in the present study, it is nevertheless possible that they have been visualized, for example, by imagining the pitch as a location in space. Tones may have also resulted in semantic representations, by categorizing them into arbitrary sets of low, medium, and high tones. The decodable signal from the impulse response might thus not necessarily originate from the sensory processing areas, but rather from higher brain regions, such as the prefrontal cortex (Stokes et al., 2013). Future studies that use imaging tools with high spatial resolution might be able to arbitrate the neural origin of the cross-modal impulse response in WM.

While the neural impulse response to visual stimulus contained information about the relevant visual WM item, replicating previous results (Wolff et al., 2017), the neural response to external auditory stimulation did not. This suggests that, in contrast to auditory information, visual information is maintained in a sensory-specific neural network with no evidence of content-specific connectivity with the auditory system, possibly reflecting

the visual dominance of the human brain (Posner et al., 1976). Indeed, while it has been found that auditory stimulation results in neural activity in the visual cortex, it is notably weaker than the other way around (Martuzzi et al., 2007), which corresponds with our asymmetric findings of sensory specific and sensory nonspecific impulse responses of visual and auditory WM.

One might argue that the asymmetric findings reported here could result from the asymmetry between experiments; whereas the auditory impulse was the only nonvisual stimulus in the visual task, the auditory task contained several nonauditory stimuli (cue, fixation cross, visual impulse). The auditory impulse may have thus been more easily filtered out in the visual task, causing the neural response to be too "weak" to perturb the neural WM network. However, we found no evidence for this alternative explanation. None of the early sensory auditory ERPs was smaller in amplitude in the visual task compared with the auditory task. Indeed, the auditory P2 was larger in the visual task, the opposite direction, as would be expected if the auditory impulse was more easily ignored. There were furthermore no reliable differences in the early visual ERPs between tasks. In the later time window, there was no difference in the auditory ERPs either. The visual ERP at frontal electrodes did show elevated amplitude from 266 to 322 ms in the visual task, but the posterior electrode showed no difference. Perhaps most obvious was the lack of a clear P3 component in general, suggesting that the impulses did not elicit higher-level cognitive processing (for review on P3, see Polich, 2007). This is not unexpected, given their predictability and task-irrelevance in both tasks and modalities. Collectively, the ERPs do not support the idea that there might be systematic differences in impulse processing that could explain the differences in WM-specific impulse responses between tasks.

We found that both the processing and maintenance of pure tones were coded parametrically according to the height of the pitch, similar to previous reports of parametric auditory WM (Spitzer and Blankenburg, 2012; Uluc et al., 2018). On the other hand, a neural code for pitch chroma, the cyclical similarity of the same notes across different octaves, was not found during either perception or maintenance. It has previously been found that complex tones may be more likely to result in a neural representation of pitch chroma than pure tones (as were used in this study) during perception (Briley et al., 2013).

Visual orientations were clearly coded parametrically during encoding and maintenance, replicating previous findings (e.g., Saproo and Serences, 2010). Interestingly, we also found evidence for a neural coding scheme that reflects the specialization of orientations close to the cardinal axes (horizontal and vertical) compared with the oblique orientations during the encoding of orientations. This coding scheme is related to the previously reported "oblique effect" (higher discrimination and report accuracy of cardinal compared with oblique orientations) (Appelle, 1972), and neural evidence for specialized neural structures in cat and macaque visual cortices for cardinal orientations (Li et al., 2003; Shen et al., 2014). The visual impulse response did not reveal such a coding scheme during maintenance, however, which could reflect a genuinely different coding scheme, but could also be due to the generally weaker orientation code during maintenance.

It has been reported that the WM-related neural pattern evoked by the impulse response does not cross-generalize with the neural activity evoked by the memory stimulus itself (Wolff et al., 2015), suggesting that the neural activation patterns are qualitatively different. In the present study, we also found no cross-generalization between item processing and the impulse response

in either the visual or in the auditory WM task. The neural representation of WM content may thus not be an exact copy of stimulation history, literally reflecting the activity pattern during information processing and encoding, but rather a reconfigured code that is optimized for future behavioral demands (Myers et al., 2017). Similarly, no generalizability was found between auditory and visual impulse responses in the auditory task. This could suggest that distinct neural networks are perturbed by the different impulse modalities, or, as alluded to above, that it reflects the unique interaction between impulses and the perturbed neural network. Future research should use neural imaging tools with high spatial resolution to investigate the neural populations involved in the WM-dependent impulse response.

The present results provide a novel approach to the ongoing debate on the extent to which sensory processing areas are essential for the maintenance of information in WM (Gayet et al., 2018; Scimeca et al., 2018; Xu, 2018). This is usually investigated by measuring WM-specific delay activity in the visual cortex in visual WM tasks (Harrison and Tong, 2009; Bettencourt and Xu, 2016), where null results are interpreted as evidence against the involvement of specific brain regions, which is inherently problematic (Ester et al., 2016), and by which nonactive WM states are not considered. In the present study, we found that sensory-specific stimulation, and both sensory specific and nonspecific stimulation, resulted in WM-specific neural responses during the maintenance of visual and auditory information, respectively. Sensory cortices were thus linked to WM maintenance not by relying on ambient delay activity, but rather by perturbing the underlying, connectivity-dependent, representational WM network via a bottom-up neural response.

References

Appelle S (1972) Perception and discrimination as a function of stimulus orientation: the "oblique effect" in man and animals. Psychol Bull 78: 266–278.

Bettencourt KC, Xu Y (2016) Decoding the content of visual short-term memory under distraction in occipital and parietal areas. Nat Neurosci 19:150–157.

Boutros NN, Korzyukov O, Jansen B, Feingold A, Bell M (2004) Sensory gating deficits during the mid-latency phase of information processing in medicated schizophrenia patients. Psychiatry Res 126:203–215.

Briley PM, Breakey C, Krumbholz K (2013) Evidence for pitch chroma mapping in human auditory cortex. Cereb Cortex 23:2601–2610.

Buonomano DV, Maass W (2009) State-dependent computations: spatio-temporal processing in cortical networks. Nat Rev Neurosci 10:113–125.

Chang A, Bosnyak DJ, Trainor LJ (2016) Unpredicted pitch modulates beta oscillatory power during rhythmic entrainment to a tone sequence. Front Psychol 7:327.

Conroy MA, Polich J (2007) Normative variation of P3a and P3b from a large sample: gender, topography, and response time. J Psychophysiol 21:22–32.

Cromwell HC, Mears RP, Wan L, Boutros NN (2008) Sensory gating: a translational effort from basic to clinical science. Clin EEG Neurosci 39: 69–72.

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

De Maesschalck R, Jouan-Rimbaud D, Massart DL (2000) The Mahalanobis distance. Chemometr Intell Lab Syst 50:1–18.

Derrick B, Toher D, White P (2017) How to compare the means of two samples that include paired observations and independent observations: a companion to Derrick, Russ, Toher and White (2017). Quant Methods Psychol 13:120–126.

Di Russo F, Martínez A, Hillyard SA (2003) Source analysis of event-related cortical activity during visuo-spatial attention. Cereb Cortex 13:486–499.

Driver J, Spence C (1998) Attention and the crossmodal construction of space. Trends Cogn Sci 2:254–262.

Eckert MA, Kamdar NV, Chang CE, Beckmann CF, Greicius MD, Menon V

(2008) A cross-modal system linking primary auditory and visual cortices. Hum Brain Mapp 29:848–857.

Ester EF, Rademaker RL, Sprague TC (2016) How do visual and parietal cortex contribute to visual short-term memory? ENeuro 3:ENEURO. 0041–16.2016.

Gayet S, Paffen CL, Van der Stigchel S (2018) Visual working memory storage recruits sensory processing areas. Trends Cogn Sci 22:189–190.

Grootswagers T, Wardle SG, Carlson TA (2017) Decoding dynamic brain patterns from evoked responses: a tutorial on multivariate pattern analysis applied to time series neuroimaging data. J Cogn Neurosci 29: 677–697.

Harrison SA, Tong F (2009) Decoding reveals the contents of visual working memory in early visual areas. Nature 458:632–635.

Huang Y, Matysiak A, Heil P, König R, Brosch M (2016) Persistent neural activity in auditory cortex is related to auditory working memory in humans and nonhuman primates. Elife 5:e15441.

Iurilli G, Ghezzi D, Olcese U, Lassi G, Nazzaro C, Tonini R, Tucci V, Benfenati F, Medini P (2012) Sound-driven synaptic inhibition in primary visual cortex. Neuron 73:814–828.

Kisley MA, Noecker TL, Guinther PM (2004) Comparison of sensory gating to mismatch negativity and self-reported perceptual phenomena in healthy adults. Psychophysiology 41:604–612.

Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis: connecting the branches of systems neuroscience. Front Syst Neurosci 2:4.

Kumar S, Joseph S, Gander PE, Barascud N, Halpern AR, Griffiths TD (2016) A brain system for auditory working memory. J Neurosci 36:4492–4505.

Ledoit O, Wolf M (2004) Honey, I shrunk the sample covariance matrix. J Portfolio Manage 30:110–119.

Lewis-Peacock JA, Drysdale AT, Oberauer K, Postle BR (2012) Neural evidence for a distinction between short-term memory and the focus of attention. J Cogn Neurosci 24:61–79.

Li B, Peterson MR, Freeman RD (2003) Oblique effect: a neural basis in the visual cortex. J Neurophysiol 90:204–217.

Luck SJ, Woodman GF, Vogel EK (2000) Event-related potential studies of attention. Trends Cogn Sci 4:432–440.

Lundqvist M, Rose J, Herman P, Brincat SL, Buschman TJ, Miller EK (2016) Gamma and beta bursts underlie working memory. Neuron 90:152–164.

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164:177–190.

Martuzzi R, Murray MM, Michel CM, Thiran JP, Maeder PP, Clarke S, Meuli RA (2007) Multisensory interactions within human primary cortices revealed by BOLD dynamics. Cereb Cortex 17:1672–1679.

Mongillo G, Barak O, Tsodyks M (2008) Synaptic theory of working memory. Science 319:1543–1546.

Morrill RJ, Hasenstaub AR (2018) Visual information present in infragranular layers of mouse auditory cortex. J Neurosci 38:2854–2862.

Myers NE, Rohenkohl G, Wyart V, Woolrich MW, Nobre AC, Stokes MG (2015) Testing sensory evidence against mnemonic templates. Elife 4: e09000.

Myers NE, Stokes MG, Nobre AC (2017) Prioritizing information during working memory: beyond sustained internal attention. Trends Cogn Sci 21:449–461.

Nemrodov D, Niemeier M, Patel A, Nestor A (2018) The neural dynamics of facial identity processing: insights from EEG-based pattern analysis and image reconstruction. ENeuro 5:ENEURO.0358–17.2018.

Oostenveld R, Fries P, Maris E, Schoffelen JM, Oostenveld R, Fries P, Schoffelen JM (2010) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011:e156869.

Polich J (2007) Updating P300: an integrative theory of P3a and P3b. Clin Neurophysiol 118:2128–2148.

Posner MI, Nissen MJ, Klein RM (1976) Visual dominance: an information-processing account of its origins and significance. Psychol Rev 83:157–171.

Pratte MS, Park YE, Rademaker RL, Tong F (2017) Accounting for stimulus-specific variation in precision reveals a discrete capacity limit in visual working memory. J Exp Psychol Hum Percept Perform 43:6–17.

Rauss KS, Pourtois G, Vuilleumier P, Schwartz S (2009) Attentional load modifies early activity in human primary visual cortex. Hum Brain Mapp 30:1723–1733.

Ringach DL, Shapley RM, Hawken MJ (2002) Orientation selectivity in macaque V1: diversity and laminar dependence. J Neurosci 22:5639–5651.

Rose NS, LaRocque JJ, Riggall AC, Gosseries O, Starrett MJ, Meyering EE, Postle BR (2016) Reactivation of latent working memories with transcranial magnetic stimulation. Science 354:1136–1139.

Saproo S, Serences JT (2010) Spatial attention improves the quality of population codes in human visual cortex. J Neurophysiol 104:885–895.

Scimeca JM, Kiyonaga A, D'Esposito M (2018) Reaffirming the sensory recruitment account of working memory. Trends Cogn Sci 22:190–192.

Shen G, Tao X, Zhang B, Smith EL 3rd, Chino YM (2014) Oblique effect in visual area 2 of macaque monkeys. J Vis 14:3.

Spitzer B, Blankenburg F (2012) Supramodal parametric working memory processing in humans. J Neurosci 32:3287–3295.

Sprague TC, Ester EF, Serences JT (2016) Restoring latent visual working memory representations in human cortex. Neuron 91:694–707.

Squires NK, Squires KC, Hillyard SA (1975) Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. Electroencephalogr Clin Neurophysiol 38:387–401.

Stokes MG (2015) 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. Trends Cogn Sci 19:394–405.

Stokes MG, Kusunoki M, Sigala N, Nili H, Gaffan D, Duncan J (2013) Dynamic coding for cognitive control in prefrontal cortex. Neuron 78:364–375.

Uluc I, Schmidt TT, Wu YH, Blankenburg F (2018) Content-specific codes of parametric auditory working memory in humans. Neuroimage 183:254–262.

Vetter P, Smith FW, Muckli L (2014) Decoding sound and imagery content in early visual cortex. Curr Biol 24:1256–1262.

Watanabe K, Funahashi S (2014) Neural mechanisms of dual-task interference and cognitive capacity limitation in the prefrontal cortex. Nat Neurosci 17:601–611.

Wessel JR, Aron AR (2017) On the globality of motor suppression: unexpected events and their influence on behavior and cognition. Neuron 93:259–280.

Wolff MJ, Ding J, Myers NE, Stokes MG (2015) Revealing hidden states in visual working memory using electroencephalography. Front Syst Neurosci 9:123.

Wolff MJ, Jochim J, Akyürek EG, Stokes MG (2017) Dynamic hidden states underlying working-memory-guided behavior. Nat Neurosci 20:864–871.

Xu Y (2017) Reevaluating the sensory account of visual working memory storage. Trends Cogn Sci 21:794–815.

Xu Y (2018) Sensory cortex is nonessential in working memory storage. Trends Cogn Sci 22:192–193.