Spatial Attention and Temporal Expectation Exert Differential Effects on Visual and Auditory Discrimination

Anna Wilsch[1], **Manuel R. Mercier**[2,3], **Jonas Obleser**[4,5], **Charles E. Schroeder**[6,7], **Saskia Haegens**[6,8]

[1]New York University

[2]University of Toulouse Paul Sabatier

[3]Aix Marseille University, Inserm, INS, Institut de Neurosciences des, Systèmes, Marseille, France

[4]University of Lübeck

[5]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[6]Columbia University College of Physicians and Surgeons

[7]Nathan Kline Institute, Orangeburg, SC

[8]Radboud University Nijmegen

Abstract

Anticipation of an impending stimulus shapes the state of the sensory systems, optimizing neural and behavioral responses. Here, we studied the role of brain oscillations in mediating spatial and temporal anticipations. Because spatial attention and temporal expectation are often associated with visual and auditory processing, respectively, we directly contrasted the visual and auditory modalities and asked whether these anticipatory mechanisms are similar in both domains. We recorded the magnetoencephalogram in healthy human participants performing an auditory and visual target discrimination task, in which cross-modal cues provided both temporal and spatial information with regard to upcoming stimulus presentation. Motivated by prior findings, we were specifically interested in delta (1–3 Hz) and alpha (8–13 Hz) band oscillatory state in anticipation of target presentation and their impact on task performance. Our findings support the view that spatial attention has a stronger effect in the visual domain, whereas temporal expectation effects are more prominent in the auditory domain. For the spatial attention manipulation, we found a typical pattern of alpha lateralization in the visual system, which correlated with response speed. Providing a rhythmic temporal cue led to increased postcue synchronization of low-frequency rhythms, although this effect was more broadband in nature, suggesting a general phase reset rather than frequency-specific neural entrainment. In addition, we observed delta-band synchronization with a frontal topography, which correlated with performance, especially in the auditory task. Combined, these findings suggest that spatial and temporal anticipations operate via a top-down modulation of the power and phase of low-frequency oscillations, respectively.

INTRODUCTION

Effective processing of the myriad of sensory input we receive in daily life requires attention to select and amplify relevant input and to attenuate irrelevant information. A critical component of attention is the ability to predict where and when important events will occur, such that processing resources can be directed in advance of anticipated stimuli. Here, we study the role of brain oscillations in anticipatory spatial attention (Posner, 1980) and temporal expectation (Nobre, Correa, & Coull, 2007). Our working hypothesis is that specific oscillatory mechanisms provide specific dedicated low-level operations, which can be employed for a range of higher-level functions, including attention and expectation. We focus on (1) the functional inhibitory role of the alpha rhythm (Haegens, Nácher, Luna, Romo, & Jensen, 2011; Jensen & Mazaheri, 2010; Klimesch, Sauseng, & Hanslmayr, 2007) and (2) selective sampling/rhythmic facilitation by low-frequency delta rhythms (Haegens & Zion Golumbic, 2018; Schroeder & Lakatos, 2009). Specifically, we ask how task demands influence the brain's oscillatory dynamics: Which oscillatory tools are employed, and how are they controlled?

Prior studies using (spatial) attention paradigms have shown that alpha activity in sensory cortex can be modulated in anticipation of input, with clear consequences for task performance. In this view, alpha oscillations reflect the focus of attention, with alpha decreasing for the attended (to-be-facilitated) location and increasing for the unattended (to-be-suppressed) location. This has been extensively reported for both the visual (Gould, Rushworth, & Nobre, 2011; Thut, Nietzel, Brandt, & Pascual-Leone, 2006; Worden, Foxe, Wang, & Simpson, 2000) and somatosensory (Anderson & Ding, 2011; Haegens, Händel, & Jensen, 2011; Jones et al., 2010) systems, with a few studies reporting similar findings in the auditory domain (Wöstmann, Herrmann, Maess, & Obleser, 2016; Frey et al., 2014).

Furthermore, it has been shown that, in sensory areas, the phase of intrinsic low-frequency oscillations (especially in the delta/theta range, approximately 1–7 Hz) predicts the magnitude of the neural response to input, as well as perceptual performance in response to this input, especially the probability of near-threshold stimulus detection (Henry, Herrmann, & Obleser, 2016; Ai & Ro, 2014; Fiebelkorn et al., 2011; VanRullen, Busch, Drewes, & Dubois, 2011; Busch & VanRullen, 2010; Busch, Dubois, & VanRullen, 2009; Mathewson, Gratton, Fabiani, Beck, & Ro, 2009; Lakatos, Chen, O'Connell, Mills, & Schroeder, 2007). Alignment of sensory cortical rhythms to a rhythmic input stream reportedly amplifies responses to stimuli occurring in- versus out-of-phase with the synchronized rhythm (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008; for reviews, see Haegens & Zion Golumbic, 2018; VanRullen, 2016). Phase entrainment has been suggested as a mechanism of this amplification at the level of primary sensory cortex (Lakatos et al., 2008) and potentially beyond (Besle et al., 2011); however, this mechanism and its generality is currently an actively debated topic (e.g., Meyer, Sun, & Martin, 2020; Helfrich, Breska, & Knight, 2019; Haegens & Zion Golumbic, 2018; Breska & Deouell, 2017). Furthermore, the extent to which the brain can use such a strategy in a purely anticipatory, top-down-controlled manner is an open question (Obleser & Kayser, 2019; Haegens & Zion Golumbic, 2018; Zoefel, ten Oever, & Sack, 2018).

Although attention and expectation are, by definition, supramodal processes, they may be constrained by distinctive features of each sensory modality. The different modalities have differing sensitivities for different types of information; for instance, vision provides information at a higher spatial resolution than audition, whereas audition provides better temporal resolution (Stein & Meredith, 1993; Freides, 1974). Accordingly, the sensory modalities might differ in their ability to form expectations and direct attention in a given dimension (i.e., time, space; see, for instance, Michalka, Kong, Rosen, Shinn-Cunningham, & Somers, 2015; Pasinski, McAuley, & Snyder, 2015; Talsma & Kok, 2002). Here, we use magnetoencephalography (MEG) recordings in healthy human participants performing a visual-auditory discrimination task, in which we manipulate both spatial attention and temporal expectation. Using a cross-modal cueing paradigm, we assess the modality specifics of the neuronal strategies brought into play in directing anticipatory resources. Specifically, we study the role of alpha- and delta-band dynamics in setting the state of the sensory systems just before stimulus presentation and assess their relative impact on behavioral outcomes.

METHODS

Participants

Twenty healthy right-handed participants (10 women; age range: 22–34 years, *Mdn* = 26 years) took part in the experiment. All participants had self-reported normal hearing. The local ethics committee (University of Leipzig) approved the study in accordance with the Declaration of Helsinki. Participants were fully debriefed about the nature and goals of the study and provided written informed consent before testing. All participants received financial compensation of €7 per hour for their participation. One participant had to be excluded from the visual-target experiment because of technical issues.

Experimental Task and Stimuli

Each participant took part in two MEG sessions performing a target discrimination task. Both sessions had the same setup and task structure, with each trial consisting of a cue stimulus directing temporal expectation and spatial attention to a target stimulus to which the participants had to respond. The difference between the two sessions was that, in one session, the cue was a visual stimulus and the target was an auditory stimulus, and vice versa in the other session. We used cross-modal cueing with the purpose of reducing evoked responses in the sensory modality of interest right before target-stimulus presentation (i.e., the window of interest for our analysis). Order of the sessions was counterbalanced across participants. In the auditory-target session, participants had to indicate whether a brief chirp went up or down in pitch. In the visual-target session, they had to discriminate whether the horizontal line of a cross was higher or lower than the center of the vertical line (i.e., whether the T was upright or upside-down). See below for more detailed information on cue and target stimuli.

Visual stimuli were presented via a back-projection screen (Panasonic PT-D7700E; 1400 × 1050 resolution, 60-Hz refresh rate). For both modality conditions, participants were

instructed to fixate on a fixation cross in the center of the screen throughout all trials. Distance to the screen was 90 cm.

Characteristics of the Spatial and Temporal Cues

The cue contained information regarding both the time and location of stimulus occurrence, to investigate whether temporal expectation and spatial attention affect auditory and visual target discrimination differently. We used a 2×2 design to independently modulate spatial attention and temporal expectation, in which each manipulation could be either on or off. Spatial attention was modulated with valid and neutral cue conditions, directing the participant's attention to the left or right side (valid spatial condition) or providing no location information (neutral spatial condition). In the visual-cueing (i.e., auditory target) case, the cue was an arrow displayed at the center of the screen, pointing either left or right providing information about which ear the target sound would be presented to (valid condition) or pointing to both sides containing no information about where the target would be presented (neutral condition). In the auditory-cueing (i.e., visual target) case, a tone to the left or right ear indicated the hemifield of upcoming target presentation (valid condition), whereas a tone to both ears provided no spatial information (neutral condition).

In a similar vein, temporal expectation was modulated with rhythmic and random cue conditions. In the rhythmic condition, the cue was presented four times at a ~ 1.6-Hz rate (i.e., with 600-msec stimulus onset intervals). For the visual-cueing condition, this meant that the arrow flashed at a 1.6-Hz rate; for the auditory-cueing condition, the tone was presented at a 1.6-Hz rhythm. Each single cue stimulus was presented for 300 msec corresponding to half a cycle of the 1.6-Hz rhythm. The target then occurred either in-phase (80% of trials) or antiphase (20% of trials) with this cued rhythm, within a window of maximum four cycles after the cue. That is, for in-phase trials, the target occurred at either 0.6, 1.2, 1.8, or 2.4 sec after the onset of the last cue (see Figure 1B for trial outline), whereas for antiphase trials, the target occurred at 0.9, 1.5, or 2.1 s after the last cue onset. This can be summed up as a 2.1-sec time window during which the rhythmic cue was presented and a 0.3- to 2.1-sec window after cue offset during which the stimulus could occur. In the random condition, the cue was presented continuously for 2.1 sec, mirroring the same period during which the rhythmic cue was presented, but not providing any specific temporal information about target occurrence. The target was then presented at a random time point in a window of 0.3–2.1 sec after cue offset (with flat probability distribution), again mirroring the rhythmic condition in terms of trial duration. Thus, in the rhythmic condition, participants could form a precise prediction about when in time the target stimulus would likely occur (i.e., in-phase with the cue rhythm and within one to four cycles after cue offset), whereas in the random condition, participants did not have precise information about target timing (other than sometime within a 0.3- to 2.1-sec window after cue offset). After the participant's response, there was a ~ 1-sec (jittered between 750 and 1200 msec) stimulus-free intertrial interval before the onset of the next trial.

Note that, whereas our spatial attention manipulation was explicit and could be directly utilized (e.g., in the visual cueing case, an arrow pointing to the cued side would be instructive upon first presentation), our temporal expectation condition was more implicit

and expectations could be formed once one was familiar with the trial structure (i.e., this required implicit learning of the temporal statistics of the task design). Furthermore, for the spatial attention manipulation, the 100% valid versus neutral cues provided a contrast to assess both the behavioral benefit and neural correlate of spatial attention. Although the switch costs involved in invalid spatial cueing (and their relation to alpha lateralization) are an interesting topic of study (see, e.g., Haegens, Händel, et al., 2011), to restrict length of the experiment, we did not include an invalid spatial cueing condition here. In the temporal expectation manipulation, we did include invalid (i.e., antiphase) trials, as here we aimed to test two aspects: first, the behavioral benefit and neural correlate of a temporal prediction (rhythmic vs. random condition) and, second, the phasic aspect of rhythm-based predictions (benefit of in-phase vs. antiphase stimuli).

Characteristics of the Target Stimuli

The target stimuli were complex sounds in the auditory condition and visually presented characters "T" in the visual condition. Both types of target stimuli were presented for 40 msec and were presented either in an "up" or "down" version. The participants' task was then to discriminate the target and report via button press whether the target was "up" or "down."

The auditory target stimuli were mono sounds and consisted of 30 different frequencies randomly drawn from within 500 to 1500 Hz. The sounds were so called frequency-modulated sweeps created with the Matlab function *chirp* and were either increasing (i.e., "up") or decreasing ("down") in pitch. Sounds had a 10-msec cosine ramp fading in and fading out to avoid onset and offset click perception. All sounds were normalized (using peak normalization) to the same sound pressure level. The visual target stimulus was a light gray cross on dark gray background, presented 300 pixels to the left or right of the center (approximately 3°), respectively. The vertical line was 20 pixels long, and the horizontal line was 16 pixels wide, both with a line width of 0.8 pixels. The horizontal line was offset from the center of the vertical line, creating either an upright (i.e., "up") or upside-down ("down") "T."

Target stimuli were individually adjusted to the participants' discrimination threshold, using a custom adaptive-tracking procedure aiming for a discrimination performance between 65% and 85% correct responses. For the auditory targets, the threshold was the slope of the pitch increase and decrease, respectively, measured as the range from lowest to highest frequency (starting point to end). Each participant was presented with a pair of sounds ("up" and "down") consisting of the a priori randomly generated frequencies, modulated depending on their individual threshold; that is, the 30 base frequencies were the same for each participant, with the participant's individual threshold changing the start and end frequencies of the sounds. For the visual target, the shift of the horizontal bar from the center of the cross was individually adjusted. These thresholds were determined per participant before the experimental sessions.

Procedure

Participants were familiarized with the stimuli and task and performed two blocks of 36 practice trials (one per temporal cueing condition). Next, individual thresholds were estimated before the MEG measurement using two consecutive blocks of an adaptive staircase procedure (each block consisting of 36 trials per cueing condition, i.e., 72 trials per block), in which thresholds were incrementally adjusted to yield performance in the range of 65–85% correct responses.

In both experimental sessions (auditory target and visual target), brain activity was recorded with MEG during the performance of 480 trials completed in 10 blocks of 48 trials each. The temporal expectation manipulation (rhythmic vs. random cue condition; with 240 trials per condition, of which 180 trials rhythmic in-phase) was kept constant within a block, whereas the spatial attention manipulation (valid vs. neutral cue condition; with 160 trials each for valid cue left, valid cue right, and neutral cue) was equally distributed across trials within blocks. The order of trials within a block and the order of blocks were randomized for each participant. Each MEG session took place on a different day, with the order counterbalanced across participants. One entire session, including practice blocks and preparation of the MEG setup, took about 4 hr per participant, with 2.5 hr for the actual testing phase.

Data Recording and Analysis

Participants were seated in an electromagnetically shielded room (Vacuumschmelze). Magnetic fields were recorded using a 306-sensor Neuromag Vectorview MEG system (Elekta) with 204 orthogonal planar gradiometers and 102 magnetometers at 102 locations. Signals were sampled at a 1000-Hz rate. The participant's head position was monitored during the measurement by five head-position indicator coils. The signal space separation method was applied offline to suppress external interferences in the data, to interpolate bad channels, and to transform individual data to a default head position that allows statistical analyses across participants in sensor space (Taulu, Kajola, & Simola, 2004).

Subsequent data analyses were carried out with MATLAB (The MathWorks Inc.) and the FieldTrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). Analyses were conducted using only gradiometers, as they are most sensitive to magnetic fields originating directly underneath them (Hämäläinen, Hari, Ilmoniemi, Knuutila, & Lounasmaa, 1993), and only including trials to which correct responses were provided ("correct trials"). Trial epochs ranging from −2.0 to 6.0 sec time-locked to the onset of the cue were defined covering the longest possible duration of a trial. Epochs were low-pass filtered at 150 Hz and subsequently down-sampled to 200 Hz.

Epochs for which the signal range at any gradiometer exceeded 800 pT/m were rejected. Independent component analysis was applied to the remaining epochs to remove artifacts because of eye blinks and heartbeat. After independent component analysis, the remaining artifacts were removed by rejecting epochs for which the signal range exceeded 200 pT/m (gradiometer) and for which variance across sensors was deemed high relative to all others (per participant, per condition) based on visual inspection. On average, 375 trials per

participant were included in the analysis for auditory-target trials; and 390, for visual-target trials.

Spectral Analysis

Power spectra were calculated for each gradiometer. From each trial, a stimulus-free segment of data from 100 msec after cue offset to 50 msec before target onset was extracted. Because of the variable target onset times, the length of the segments varied from 0.45 to 2.25 sec. This segment was multiplied with a Hanning taper to compute the power of 0.5–30 Hz in 0.1-Hz steps using a fast Fourier transform (FFT) approach applying 10 sec of zero padding. The output of the analysis was complex Fourier data. For further analyses, power (squared magnitude of the complex-valued time-frequency representation [TFR] estimates) was averaged across single trials of each experimental condition (i.e., left and right cue, neutral cue, rhythmic cue, and random cue). Intertrial phase coherence (ITPC) was computed based on the complex Fourier data (Makeig et al., 2002; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996). ITPC is the magnitude of the amplitude-normalized complex values averaged across trials for each time-frequency bin per channel and experimental condition (Thorne, De Vos, Viola, & Debener, 2011). ITPC was computed across all trials in the rhythmic condition and all trials in the random condition.

To inspect the time course of the frequency effects, we also computed TFRs of the power spectra across the entire trial (i.e., −2.0 to 6.0 sec time-locked to cue onset). To this end, we used an adaptive sliding time window of four cycles length ($t = 4/f$) for each frequency from 1 to 30 Hz and applied a Hanning taper before estimating the power using an FFT approach.

Next, FFT power spectra, ITPC, and TFRs were averaged across gradiometers in each pair. This procedure resulted in one value for each time point (TFRs only), frequency bin, and sensor position. While TFRs were computed to display the oscillatory response throughout the time course of a trial, all further statistical analyses were performed on the computed FFTs.

Source Localization

To estimate the origin of sensor-level alpha-power and delta ITPC, source localizations were computed based on individual T1-weighted MRI images (3T Magnetom Trio, Siemens AG). Topographical representations of the cortical surface of each hemisphere were constructed with Freesurfer (www.surfer.nmr.mgh.harvard.edu), and the MRI coordinate system was coregistered with the MEG coordinate system using the head-position indicators and about 100 additional digitized points on the head surface (Polhemus FASTRAK 3-D digitizer). The brain volume of each participant was normalized to a common standardized source model and divided into a grid of 4-mm resolution. Leadfields were calculated for all grid points (Nolte, 2003). We applied the Dynamic Imaging of Coherent Sources beamformer approach (Gross et al., 2001) implemented in the Fieldtrip toolbox. For the source analysis of overall alpha power (Figure 2D), we calculated the cross-spectral density using FFT centered at 10.5 Hz (±2.5-Hz smoothing with four Slepian tapers; Percival & Walden, 1993). We calculated the FFTs on 1-sec data segments taken from the stimulus-free period between

cue offset and target onset. As described above, we selected segments from 0.1 sec after cue offset to 0.05 sec before target onset (with an overlap of 0.5 sec). All trials for which this window was shorter than 1 sec because of early target onsets were discarded from the analysis. On average, 253 trials per participant were included in the analysis for auditory-target trials; and 256, for visual-target trials.

Using each participant's leadfield and the cross-spectral density of all conditions, a common spatial filter was calculated for each of the 37,163 source locations inside the brain arranged in a regular grid of 4-mm steps. This filter was then applied to Fourier-transformed data of left and right valid spatially cued conditions separately to estimate alpha power for each grid point.

For a source analysis of ITPC, we derived a spatially adaptive filter based on each participant's leadfield and the cross-spectral density centered at 5 Hz with ±5-Hz spectral smoothing on the same 1-sec segments as described above. This filter was multiplied with single-trial FFTs (1.1–2.1 Hz, in steps of 0.1 Hz). ITPC at each grid point was calculated and averaged across frequencies. The purpose of source reconstruction was to identify the putative location of sensor-level effects; no further (statistical) analysis was done on this level.

Statistical Analysis

Target Discrimination Performance—We analyzed target discrimination performance by means of RT and the sensitivity measure d'. Here, d' describes the distance between the two decision models, that is, the decision a stimulus belongs to the category "up" versus the decision a stimulus belongs to the category "down" (e.g., Henry & McAuley, 2013). For each participant, d' was computed across all trials of each condition (spatially valid and neutral conditions as well as rhythmic and random conditions). RTs were measured time-locked to the offset of the target and averaged across all trials of each condition for each participant after removal of outlier trials deviating ±2 SDs from the participant's average RT. For statistical analyses, all RT values were speed-transformed (i.e., $1/RT$) to better approximate the normality assumption.

Effects of the experimental conditions on target discrimination performance were analyzed by means of repeated-measures ANOVAs. Both ANOVAs on RT and d' contained the within-participants factors Time (rhythmic vs. random), Space (valid vs. neutral), and Modality (auditory target vs. visual target). Trials with targets presented antiphase in the rhythmic condition were removed before the ANOVAs. Interactions yielding a p value $< .05$ were resolved with post hoc t tests.

For d' only, we tested whether targets presented in-phase with the cued rhythm were discriminated better than targets presented antiphase. Here, d' was computed across all in-phase targets and all antiphase targets separately. d' increased the later the target occurred, because of increasing temporal expectation as time elapsed. Therefore, we detrended the d' values to remove this effect of temporal expectation on task performance. Then, we computed a repeated-measures ANOVA on the within-participant factors In-phase versus Antiphase and Modality (auditory target vs. visual target). As effect size measures, we report

partial η^2 for repeated-measures ANOVAs and $r_{\text{equivalent}}$ for dependent-sample t tests (Rosenthal & Rubin, 2003).

Sensor Level Analyses—Statistical analyses were performed (1) to test for lateralization of alpha power (8–13 Hz) according to the modulation of spatial attention (collapsing over temporal expectation conditions) as well as (2) to test for differences in delta ITPC (1.1–2.1 Hz; centered at the cued rhythm of 1.6 Hz) between rhythmic and random temporal expectation conditions (collapsing over spatial attention conditions). We used cluster-based permutation tests when contrasting data from all sensors, whereas t tests and ANOVAs were used for data extracted from ROIs.

To test for alpha lateralization, each participant's trials were averaged for attention left and right separately within each modality condition and then normalized in the following way: (attention-left − attention-right) / (attention-left + attention-right). This approach gives negative values for a stronger decrease in the attention-left condition and positive values when there is a stronger decrease in the attention-right condition. This procedure also reduces interparticipant variability in the power estimates, thus providing a convenient normalization. To establish whether the difference between attention left and right was significantly different from 0, a cluster-based nonparametric randomization test (Maris & Oostenveld, 2007) was applied using 1000 randomizations of the data (effectively swapping the labels of the conditions of interest) for auditory and visual modalities separately. By clustering neighboring sensors that show the same effect, this test deals with the multiple-comparisons problem and, at the same time, considers the dependency of the data points (see also Haegens, Luther, & Jensen, 2012; Haegens, Händel, et al., 2011).

To capture the relative prestimulus alpha distribution over both hemispheres in one measure, a lateralization index of alpha power was calculated for each participant, using sensors within predefined ROIs ("auditory" sensors for auditory-target trials and "visual" sensors for visual-target trials; based on standard Neuromag ROIs): alpha-lateralization index = (alpha-ipsilateral-ROI — alpha-contralateral-ROI) / (alpha-ipsilateral-ROI + alpha-contralateral-ROI). Here, contralateral and ipsilateral refer to the hemispheres with respect to the spatial cue. Before computing the lateralization index, alpha power of the spatially valid cued condition was normalized by subtracting alpha power of the spatially neutral cued condition at each sensor of the selected ROIs. This index is comparable to the one applied by Haegens, Händel, et al. (2011) and gives a positive value when alpha power is higher over the ipsilateral hemisphere and/or lower over the contralateral hemisphere (ipsilateral > contralateral) and a negative value for the opposite scenario (contralateral > ipsilateral).

To test for differences in the alpha-lateralization index between visual-target and auditory-target trials, we used a dependent-sample t test. We then performed a post hoc analysis to test whether alpha lateralization was related to target discrimination performance for visual-target trials. We applied a median split across trials within each participant based on RT, computed the alpha lateralization index for low versus high RT trials, and compared these across participants with a dependent-sample t test.

To test for differences in ITPC between rhythmic and random conditions, we used a similar cluster-based nonparametric randomization approach as described above, for auditory- and visual-target trials separately. Next, we used a repeated-measures ANOVA with factors Time (rhythmic vs. random condition) and Modality (auditory vs. visual targets) to assess differences in delta-band ITPC (1.1–2.1 Hz) between modalities. We here averaged the delta ITPC across sensors of the significant clusters (see topographies in Figure 4A) in each participant, for rhythmic and random conditions separately.

Correlation of Neural Markers and Discrimination Performance—Finally, we tested the relationship between the neural measures and participants' behavioral performance. We correlated pretarget delta ITPC with participants' RTs using a cluster-based nonparametric randomization test. First, we performed this correlation across participants. That is, we used each participant's RT and used this as a regressor to correlate with delta ITPC. This was done for rhythmic and random conditions separately, as well as for auditory- and visual-target trials. Note that, for the rhythmic condition, we only used trials where the target was presented in-phase with the cue rhythm, to avoid variance in RT because of decreased discrimination performance in the antiphase trials. Because we found significant correlations for the rhythmic as well as for the random condition in the auditory-target trials (see Results), we subsequently performed a within-participant across-trial correlation of RT and delta ITPC for rhythmic and random conditions separately, focusing on auditory-target trials. Here, we sorted all trials within one participant according to his or her RT. Then, we split the trials in 10 equally sized bins and computed average RT and delta ITPC for each bin (note that ITPC cannot be computed on the single-trial level; this binning procedure allowed us to perform an "across-trial" correlation analysis). Within each participant, we performed a correlation across the binned RT and delta ITPC at each sensor. The resulting correlation coefficients of each sensor were tested against zero across participants by means of a cluster-based nonparametric randomization dependent-sample t test. Finally, we performed another analysis correlating d' of the rhythmic antiphase trials with delta ITPC, again using a cluster-based nonparametric randomization test. This analysis was computed for visual-target trials and auditory-target trials separately.

RESULTS

Here, we studied spatial attention and temporal expectation in the context of auditory and visual target discrimination using MEG. We were specifically interested in delta- and alpha-band oscillatory state in anticipation of target presentation and the impact on task performance.

Effects of Spatial Attention and Temporal Expectation on Auditory and Visual Target Discrimination

First, we assessed the effects of the temporal and spatial manipulations on task performance, in terms of sensitivity (d') and RT (Figure 2A and B), using a repeated-measures ANOVA with factors Time (rhythmic vs. random condition; note that, in the rhythmic condition, only trials with in-phase targets were included in the analysis), Space (valid vs. neutral condition), and Modality (visual vs. auditory target condition). Overall, participants were

faster on the visual-target trials than on the auditory-target trials, $F(1, 19) = 76.11, p < .0001$, $\eta^2 = .8$, with no significant difference in sensitivity, $F(1, 19) = 1.16, p = .296, \eta^2 = .057$.

We found no significant effect of the temporal expectation manipulation on either d', $F(1, 19) = 0.039$, $p = .845$, $\eta^2 = .0021$, or RT, $F(1, 19) = 2.60$, $p = .123$, $\eta^2 = .12$, when contrasting rhythmic versus random conditions (for auditory and visual targets combined). In contrast, there was a significant effect for the spatial attention manipulation on d' and RT, with higher d' for the valid than for the neutral cue condition, $F(1, 19) = 7.91$, $p = .011$, $\eta^2 = .29$, as well as faster RTs for valid cues, $F(1, 19) = 15.09$, $p = .0009$, $\eta^2 = .44$. Furthermore, there was a significant interaction between the factors Time and Space, $F(1, 19) = 6.69, p = .018$, $\eta^2 = .26$ (Figure 2C, left), with a stronger effect of the spatial manipulation on RT for the random condition, $t(19) = 4.75$, $p = .0001$, $r = .75$, than for the rhythmic condition, $t(19) = 2.45, p = .024$, $r = .5$. The interaction of the factors Modality and Time was also significant, $F(1, 19) = 5.91, p = .025$, $\eta^2 = .24$ (Figure 2C, right), with faster RTs for random than rhythmic trials in the visual modality, $t(19) = 2.67, p = .015$, $r = .53$. There was no significant effect in the auditory modality, $t(19) = -0.43$, $p = .672$, $r = .1$. (For criterion, no significant effects or trends were found across all tests; all ps > .15.)

When zooming in on the rhythmic condition, contrasting in-phase with antiphase trials with the factor Modality (Figure 2D), we found improved d' for in-phase trials, $F(1, 19) = 14.37$, $p = .001$, $\eta^2 = .43$. Note that this held true across modalities (main effects of Modality and interaction of Modality and Phase, both Fs < 1). This analysis demonstrates improved sensitivity for targets that could be anticipated using the temporal cue.

Furthermore, we asked whether there was an effect of target onset time (early vs. late; median split) on performance measures. For d', the main effect of Onset time was significant, $F(1, 19) = 4.48$, $p = .048$, $\eta^2 = .19$, with later onset times yielding better task performance than early onset times. This was true for both rhythmic and random cueing conditions (the interaction of factors Time and Onset time was not significant, $F(1, 19) = 0.17$, $p = .69$, $\eta^2 = .009$). For RT, both the main effect of Onset time, $F(1, 19) = 14.19$, $p = .001$, $\eta^2 = .43$, and the interaction of the factors Time and Onset time was significant, $F(1, 19) = 7.05$, $p = .016$, $\eta^2 = .27$. The resolution of this interaction by the factor Time showed that, in the rhythmic condition, the factor Onset time was significant, $F(1, 19) = 16.49$, $p = .0007$, $\eta^2 = .47$, with faster responses for later onsets, whereas in the random condition, the main effect of Onset time was not significant, $F(1, 19) = 2.21$, $p = .15$, $\eta^2 = .10$.

Temporal Expectation Modulates ITPC in the Delta Range

Next, we turn to the anticipatory brain state in response to the temporal cueing. TFRs of power show a low-frequency response to the cue that is sustained after cue offset (Figure 3). For the rhythmic visual-cue condition (i.e., auditory target), there was a clear rhythmic response in the TFR during cue presentation, not apparent in the auditory-cue condition (i.e., visual target). The postcue, pretarget responses to auditory cues seem somewhat more limited to the 5- to 10-Hz range (especially in the rhythmic case), whereas the visual-cue responses cover a broader low-frequency range.

We further assessed these anticipatory (postcue, pretarget) low-frequency modulations using ITPC measures. For both the auditory-target and visual-target conditions, ITPC was significantly stronger for the rhythmic condition than for the random condition (one significant cluster for the auditory-target condition: $p = .002$, two significant clusters for the visual-target condition: $p = .002$ and $p = .008$). The effect in the auditory-target condition was not limited to the delta band but covered a 0.8- to 15.5-Hz range (Figure 4A), whereas the significant effect was limited to a 0.5- to 2.0-Hz range for the visual-target condition (Figure 4B). Focusing on the delta-band ITPC of interest (i.e., the cued 1.6-Hz rhythm), we found a stronger modulation for rhythmic than random conditions, $F(1, 18) = 11.28$, $p = .004$, $\eta^2 = .39$, as well as a stronger modulation for auditory-than visual-target conditions, $F(1, 18) = 58.97$ $p < .0001$, $\eta^2 = .77$ (Figure 4C). Sources of these modulations were localized to visual regions in the auditory-target condition and to auditory regions in the visual-target condition, both with a right-hemisphere dominance (Figure 4A and B, bottom; contrasting rhythmic and random delta ITPC by subtracting random from rhythmic and dividing this by rhythmic condition). Note that our analysis of anticipatory brain activity reveals an effect of the temporal cue on the phase of ongoing oscillations in the regions that correspond to the modality of the cue.

We then asked how these delta-band modulations correlated with behavioral outcomes. First, we focused on the relationship between participants' overall delta ITPC and their task performance; that is, are participants with stronger synchronization better/faster at the task? We found that, for the auditory-target condition, participants with stronger delta ITPC were faster at the task, for both rhythmic (one significant cluster: $p = .048$) and random (one cluster: $p = .028$) conditions (Figure 4D). Considering that we found this correlation regardless of temporal condition, we seem to be tapping into a measure here that distinguishes fast from slow participants across the board (i.e., better synchronizers are faster at the task in general). There was no significant effect on d' (no clusters were detected, neither for the rhythmic nor random condition). For the visual-target condition, there was no significant relationship between participants' delta ITPC and performance measures (all $ps > .15$).

Next, we followed up the across-participants correlation for the auditory modality by asking how delta ITPC affected performance on a trial-by-trial basis: If delta synchronization is stronger on an individual trial, does that lead to faster responses? Furthermore, importantly, would this be specific to the rhythmic condition, for which we had predicted delta synchronization to be an effective mechanism for selective sampling of the subsequent target? Indeed, we find that trials with higher delta ITPC had faster RTs, only for the rhythmic (one significant cluster: $p = .026$) and not for the random (no clusters) condition (Figure 4E).

We did not find any positive correlation of delta ITPC and task performance for the visual-target condition when analyzing valid trials from the rhythmic condition (i.e., in-phase trials). However, when analyzing the rhythmic invalid trials (i.e., antiphase trials; Figure 4F), we found a significant negative correlation between d' and delta ITPC (one negative cluster: $p = .034$). This demonstrates that, with stronger delta ITPC, discrimination performance for

targets at (unattended) antiphasic positions gets worse. For the auditory-target condition, this analysis did not yield any significant effects.

All effects demonstrating a relationship between pretarget delta ITPC and target discrimination performance were located at right frontal sensors, for both the visual and auditory modalities, tentatively suggesting a possible supramodal frontal mechanism regulating temporal expectation.

Spatial Attention Modulates Alpha Power Lateralization

Next, we turn to the spatial attention manipulation. We computed the postcue, pretarget alpha power for left versus right attention trials, for the two sensory modalities separately (Figure 5A and B). We found significant alpha lateralization for the visual-target condition (one left-lateralized positive cluster: $p = .03$) but not for the auditory-target condition (no significant clusters: $p > .05$). Source reconstruction of alpha power (and computation of the attention index on source level; Figure 5B) corroborates the findings from the sensor level: In the visual modality, we found a positive attention index in the left hemisphere and a negative attention index in the right hemisphere. The attention index was the strongest at bilateral occipito-parietal regions. For the auditory modality, the attention index was more evident in source space than in sensor space. Specifically, there was a negative attention index at right occipito-parietal regions.

To further zoom in on these effects, we then computed the alpha-lateralization index per participant for a preselected set of sensors per modality. Contrasting the visual and auditory conditions, we further confirm stronger alpha lateralization for the visual case, $t(18) = 3.04$, $p = .007$, $r = .59$ (Figure 5C, left).

Finally, we asked how alpha lateralization correlated with performance and found that alpha lateralization was higher on trials with low RT than trials with high RT, for the visual modality, $t(18) = 2.20$, $p = .041$, $r = .47$ (Figure 5C, right). Note that only trials with valid spatial cues (i.e., left or right cues) were included in this analysis.

DISCUSSION

In this MEG study, we used a visual-auditory discrimination task with cross-modal cues providing both temporal and spatial information with regard to upcoming stimulus presentation. We report a dissociation between the two sensory modalities in response to the different manipulations, with stronger temporal expectation effects in the auditory task and stronger spatial attention effects in the visual task. We found that providing a rhythmic temporal cue (as compared to a random/continuous temporal cue) led to increased postcue synchronization of low-frequency rhythms, which, especially in the auditory task, correlated with performance. For the spatial attention manipulation, we found a typical pattern of alpha lateralization in the visual (but not the auditory) system, which correlated negatively with RT.

Differences between the Auditory and Visual Systems

The observed differences in response to temporal expectation and spatial attention manipulations between the visual and auditory systems (compare Figures 4C and 5C) are in line with our knowledge about differences between these sensory modalities and their physiological basis (Thorne & Debener, 2014; Freides, 1974). We do not claim a complete dissociation between the two modalities and sensitivity to spatial versus temporal information, respectively, although it seems clear that there is a bias in that the visual system is more sensitive to the spatial dimension, whereas the auditory system is more sensitive to temporal aspects.

The lack of alpha lateralization in the auditory system in response to our spatial attention condition could have several causes. For one, we might simply be less sensitive to auditory alpha than to the prominent posterior alpha rhythm with EEG/MEG (Weisz, Hartmann, Müller, Lorenz, & Obleser, 2011; Pfurtscheller & Lopes da Silva, 1999), in which case our observation might partly be explained by recording method. Intracranial recordings could remedy this—and have shown alpha modulation in auditory selective attention tasks (de Pesters et al., 2016; Lakatos et al., 2016)—although signal-to-noise ratio is probably not the sole reason for the lack of an alpha lateralization effect here. The auditory cortex is also less spatially lateralized than the visual and tactile systems (Grothe, Pecka, & McAlpine, 2010; Langers, van Dijk, & Backes, 2005), and previous work shows that spatial attention predominantly affects alpha in the right auditory cortex (ElShafei, Bouet, Bertrand, & Bidet-Caulet, 2018; Müller & Weisz, 2012). Most importantly, previous work has shown that alpha lateralization effects (especially the ipsilateral increase of alpha) are boosted with the introduction of distracters on the unattended side (i.e., providing an input that needs to be actively suppressed), for both visual (Sauseng et al., 2009) and somatosensory (Haegens et al., 2012) systems. For the auditory system, the few studies assessing alpha-power lateralization all presented a distracting sound to the not-to-be attended ear (Wöstmann et al., 2016; Frey et al., 2014; Kerlin, Shahin, & Miller, 2010). Thus, we suggest that the absence of clear alpha lateralization in our auditory task could also be because of the lack of to-be-suppressed distracters in our paradigm.

Temporal Dynamics

One of the questions we aimed to address was whether a (cross-modal) rhythmic input could be actively utilized to predict onset of relevant input (Haegens & Zion Golumbic, 2018; Zoefel et al., 2018) and thereby improve task performance.

The effects of our temporal expectation manipulation were localized (on both sensor and source levels) to the modality of the cross-modal cue, whereas the spatial attention effect appeared in the modality of the upcoming target. That is, when presented with a rhythmic auditory cue in anticipation of a visual target, the temporal modulation, as evident in increased delta-band ITPC, was localized to auditory/temporal regions, whereas the spatial modulation in terms of alpha lateralization had a visual/posterior cortical origin. Thus, it appears that our spatial attention manipulation exerted a cross-modal influence, modulating excitability in early sensory areas (here, visual cortex) with consequences for behavior, whereas our temporal expectation manipulation exerted an effect that was limited to the

modality of cue presentation and that itself did not correlate with task performance. We did, however, find that frontal delta-band synchronization correlated with performance, pointing to a potential supramodal mechanism.

The question is whether using a unimodal rather than cross-modal paradigm would change the picture. Although we had good reasons to avoid unimodal cueing here (as to separate a series of evoked responses from actual preparatory activity—Figure 3B illustrates this concern), of course, it is possible that this changed or limited potential temporal anticipation effects. Previous work has suggested that unimodal temporal expectations form in a bottom-up, automatic fashion, also when actually irrelevant or even detrimental to task performance (Breska & Deouell, 2014, 2016). These same authors further showed that such exogenous effects can be modulated by top-down, endogenous factors (Breska & Deouell, 2016). Although at this stage we can only speculate, it is possible that we are picking up on a combination of these exogenous and endogenous factors here. Future studies could address this issue by using paradigms with explicit endogenous expectations, for example, symbolic cues that the participant has to translate into a prediction, via learned associations (see, e.g., Auksztulewicz, Friston, & Nobre, 2017).

Surprisingly, RTs to visual targets were faster in the random than rhythmic condition, contrary to our hypothesis and to previous findings (e.g., Cravo, Rohenkohl, Wyart, & Nobre, 2013), although it should be noted that we found temporal expectation effects were generally weakest for the visual condition. In addition, although there was no overall rhythmic versus random condition benefit, when correcting for hazard rate effects (i.e., increasing temporal expectation over time in a situation in which an event is certain to occur, but the exact time of occurrence is not known), we did observe a small but consistent performance improvement for in-phase versus antiphase trials. It has been argued that behavioral entrainment effects can be observed specifically in difficult detection tasks using near-threshold targets embedded in noise and/or followed by a mask, making precise temporal expectations more critical (e.g., ten Oever et al., 2017; Henry & Obleser, 2012). The lack of a stronger rhythmic benefit for performance might be because of the use of a (cross-modal) discrimination task (without noise, masks, or distracters)—although when using the neural synchronization measure as a correlate, we did observe rhythmic condition-specific behavioral effects, pointing to a successful (although more subtle) temporal expectation manipulation.

Indeed, we found that right frontal delta ITPC correlated with performance measures (RT and *d'*). For the auditory task, the observed neuronal synchronization correlated negatively with RTs, such that participants with higher delta-band ITPC were faster at the task overall (including the random condition; i.e., a nonspecific effect), whereas on individual trials, higher delta ITPC boosted RTs for the rhythmic condition only (i.e., a specific effect). In the visual task, we found a tentative negative effect of synchronization on invalid cue trials— stronger synchronization hampers discrimination performance for antiphase target presentation.

Thus, it seems that, here, there are two phenomena at play: a unimodal, "bottom-up" sensory effect and a supramodal, "top-down" anticipatory effect (cf. Auksztulewicz et al., 2017). The

first effect is mainly captured by ITPC during the cue-target interval, in the modality of the cue, and does not correlate with performance, whereas the second effect has a right frontal topography (see also Wilsch, Henry, Herrmann, Maess, & Obleser, 2015) and shows up in the analysis linking the neural signal to behavior. Previous work studying the effects of spontaneous prestimulus phase on (visual) detection/discrimination performance found the phase of frontal theta/alpha band oscillations to predict performance (Dugué, Marque, & VanRullen, 2011; VanRullen et al., 2011; Busch et al., 2009). These findings have been suggested to reflect attentional cycles (VanRullen, Zoefel, & Ilhan, 2014), that is, rhythmic fluctuation of attentional sampling. Furthermore, several studies have shown that, in the absence of external rhythms, visual spatial attention fluctuates at a theta rate (4–8 Hz; Fiebelkorn, Pinsk, & Kastner, 2018; Helfrich et al., 2018; Landau, Schreyer, van Pelt, & Fries, 2015; Fiebelkorn, Saalmann, & Kastner, 2013; Landau & Fries, 2012), with lower periodicities when attention has to be divided over multiple locations (e.g., a ~7-Hz intrinsic attentional rhythm might be split over two locations in a bilateral attention task and show up as a 3- to 4-Hz rhythm in each location; VanRullen, 2016).

In the current study, we found an interaction effect of the spatial and temporal manipulations on RT. Furthermore, the observed synchronization effects were exclusively present in the rhythmic condition, and the behavioral modulations also specifically emerged in rhythmic trials. These effects linking neuronal activity to behavior had a right frontal topography (across modality conditions), in line with higher-order, supramodal mechanisms rather than a purely low-level, sensory-specific phenomenon. Here, we might be tapping into a complex interplay (potentially mitigated through cross-frequency coupling) of spatial attention effects (and their intrinsic rhythmic fluctuations) on the one hand and temporal expectation on the other.

Is This Evidence of Entrainment?

On the basis of previous work, we hypothesized that the brain would pick up on the temporal information in the rhythmic cue and form a (rhythmic) prediction of when the target would most likely be presented. There is a considerable literature tentatively showing that attended or otherwise salient rhythmic input leads to (bottom-up/ automatic) entrainment or synchronization of neuronal populations in early sensory areas, which is then maintained after cue offset (for a number of cycles), rhythmically biasing subsequent stimulus processing (for a review and critical discussion, see Haegens & Zion Golumbic, 2018). This has been shown with both neural (e.g., Lakatos et al., 2008, 2013; Henry & Obleser, 2012) and behavioral (Fiebelkorn et al., 2011; Rohenkohl, Coull, & Nobre, 2011; Jones, Moynihan, MacKenzie, & Puente, 2002; Barnes & Jones, 2000) data and has been suggested as a mechanism of selective attention and/or to implement temporal predictions (Schroeder & Lakatos, 2009; Nobre et al., 2007). Lakatos et al. (2008), for example, have shown that, when competing auditory and visual streams are present, Layer 4 of macaque V1 aligns its high excitability phase to the visual stream, regardless of which stream is attended, whereas a few hundred microns away, Layer 3 aligns its high excitability phase with the attended stream (regardless of whether it is auditory or visual).

However, several open questions remain, with one main issue being the nontrivial distinction of a rhythmic input causing a series of evoked responses versus a genuine intrinsic oscillator getting phase aligned (Helfrich et al., 2019; Haegens & Zion Golumbic, 2018; Lakatos et al., 2013; Capilla, Pazo-Alvarez, Darriba, Campo, & Gross, 2011). Here, we used a novel paradigm with MEG measures to provide further evidence for (1) the proposed oscillatory entrainment (i.e., increased frequency-specific synchronization after cue offset) and (2) the extent to which the brain can use such a strategy for top–down-controlled anticipation (here, cross-modal cueing effects).

The main question is then whether the here-reported phase-synchronization effects constitute evidence of entrainment in the narrow sense (cf. Obleser & Kayser, 2019), that is, entrainment of intrinsic, self-sustained oscillators in early sensory networks coupling with an external rhythm, resulting in frequency and phase alignment between the two. Indeed, we do find stronger neuronal synchronization in response to a rhythmic cue than to a continuous cue; however, this synchronization was not exclusively observed at the cued frequency of 1.6 Hz but rather extended to a substantial range of low-frequency rhythms. This lack of narrow-band synchronization could potentially be explained by a number of methodological limitations. A first set of concerns is related to the paradigm used: Perhaps, four cycles is not enough to properly drive the system (though see, e.g., Breska & Deouell, 2014; de Graaf et al., 2013). In addition, the task might not have been demanding enough to warrant a strong phase coupling (see also discussion of behavioral effects above). Further methodological explanations relate to the signal used: Perhaps, MEG in combination with our analysis techniques does not provide the right spectral and spatial resolution and/or signal-to-noise ratio to detect the focal, band-limited effect of interest. Arguably, here, we might mix signals from neuronal populations with different responses: ones that do phase-align to the 1.6-Hz rhythm (cf. Besle et al., 2011; Gomez-Ramirez et al., 2011; Lakatos et al., 2008) and other neuronal populations oscillating at different frequencies, which, when combined, lead to the observed broadband effects. Possibly, the hierarchical organization of neuronal oscillations through cross-frequency coupling could lead/contribute to the spread of narrow-band synchronization over multiple frequencies (Lakatos et al., 2005).

These caveats notwithstanding, our observations do not provide conclusive evidence for frequency-specific entrainment of early sensory systems. We do find evidence for an effect of rhythmic input on synchronization in a wider range of low frequencies. Rather than entrainment, this might reflect a more general boost of intrinsic low-frequency oscillators, perhaps via (modulatory) phase reset (ten Oever, van Atteveldt, & Sack, 2015) or resonance response (see Lakatos et al., 2009). Such phase reset could in turn lead to an overall increase in low-frequency synchronization (as measured here with ITPC). An alternative explanation is that rhythmic cue presentation causes a series of (rhythmic) evoked responses, after which the system goes back to its natural state (i.e., superposition; Capilla et al., 2011). The evoked response and accompanying phase reset caused by these inputs then explain the temporary increase in synchronization, because now more cells are phase aligned. Note that these phase reset scenarios could have a cross-modal impact as well (via feedforward and/or feedback pathways; Mercier et al., 2013, 2015; Lakatos et al., 2007, 2008, 2009). Intracranial recordings, with a more fine-grained spatial resolution than the methods used here, might be able to distinguish between these possibilities.

In conclusion, we find (i) that the brain utilizes different oscillatory tools depending on task demands, (ii) that these mechanisms operate on different levels of the system, and (iii) that sensory modalities differ in their ability to direct anticipatory resources in a given dimension. Future intracranial work should allow for a more detailed understanding of the underlying neurophysiology of these mechanisms.

Acknowledgments

REFERENCES

Ai L & Ro T (2014). The phase of prestimulus alpha oscillations affects tactile perception. Journal of Neurophysiology, 111, 1300–1307. [PubMed: 24381028]

Anderson KL, & Ding M (2011). Attentional modulation of the somatosensory mu rhythm. Neuroscience, 180, 165–180. [PubMed: 21310216]

Auksztulewicz R, Friston KJ, & Nobre AC (2017). Task relevance modulates the behavioural and neural effects of sensory predictions. PLoS Biology, 15, e2003143. [PubMed: 29206225]

Barnes R, & Jones MR (2000). Expectancy, attention, and time. Cognitive Psychology, 41, 254–311. [PubMed: 11032658]

Besle J, Schevon CA, Mehta AD, Lakatos P, Goodman RR, McKhann GMet al, et al. (2011). Tuning of the human neocortex to the temporal dynamics of attended events. Journal of Neuroscience, 31, 3176–3185. [PubMed: 21368029]

Breska A, & Deouell LY (2014). Automatic bias of temporal expectations following temporally regular input independently of high-level temporal expectation. Journal of Cognitive Neuroscience, 26, 1555–1571. [PubMed: 24392898]

Breska A, & Deouell LY (2016). When synchronizing to rhythms is not a good thing: Modulations of preparatory and post-target neural activity when shifting attention away from on-beat times of a distracting rhythm. Journal of Neuroscience, 36, 7154–7166. [PubMed: 27383591]

Breska A, & Deouell LY (2017). Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. PLoS Biology, 15, e2001665. [PubMed: 28187128]

Busch NA, Dubois J, & VanRullen R (2009). The phase of ongoing EEG oscillations predicts visual perception. Journal of Neuroscience, 29, 7869–7876. [PubMed: 19535598]

Busch NA, & VanRullen R (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. Proceedings of the National Academy of Sciences, U.S.A, 107, 16048–16053.

Capilla A, Pazo-Alvarez P, Darriba A, Campo P, & Gross J (2011). Steady-state visual evoked potentials can be explained by temporal superposition of transient event-related responses. PLoS One, 6, e14543–e14515. [PubMed: 21267081]

Cravo AM, Rohenkohl G, Wyart V, & Nobre AC (2013). Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. Journal of Neuroscience, 33, 4002–4010. [PubMed: 23447609]

de Graaf TA, Gross J, Paterson G, Rusch T, Sack AT, & Thut G (2013). Alpha-band rhythms in visual task performance: Phase-locking by rhythmic sensory stimulation. PLoS One, 8, e60035. [PubMed: 23555873]

de Pesters A, Coon WG, Brunner P, Gunduz A, Ritaccio AL, Brunet NM, et al. (2016). Alpha power indexes task-related networks on large and small scales: A multimodal ECoG study in humans and a non-human primate. Neuroimage, 134, 122–131. [PubMed: 27057960]

Dugué L, Marque P, & VanRullen R (2011). The phase of ongoing oscillations mediates the causal relation between brain excitation and visual perception. Journal of Neuroscience, 31, 11889–11893. [PubMed: 21849549]
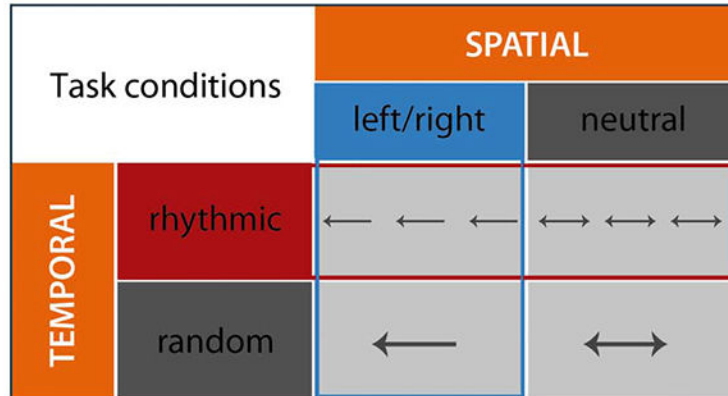
ElShafei HA, Bouet R, Bertrand O, & Bidet-Caulet A (2018). Two sides of the same coin: Distinct sub-bands in the alpha rhythm reflect facilitation and suppression mechanisms during auditory anticipatory attention. eNeuro, 5, 1–14.

Fiebelkorn IC, Foxe JJ, Butler JS, Mercier MR, Snyder AC, & Molholm S (2011). Ready, set, reset: Stimulus-locked periodicity in behavioral performance demonstrates the consequences of cross-sensory phase reset. Journal of Neuroscience, 31, 9971–9981. [PubMed: 21734288]

Fiebelkorn IC, Pinsk MA, & Kastner S (2018). A dynamic interplay within the frontoparietal network underlies rhythmic spatial attention. Neuron, 99, 842–853. [PubMed: 30138590]

Fiebelkorn IC, Saalmann YB, & Kastner S (2013). Rhythmic sampling within and between objects despite sustained attention at a cued location. Current Biology, 23, 2553–2558. [PubMed: 24316204]

Freides D (1974). Human information processing and sensory modality: Cross-modal functions, information complexity, memory, and deficit. Psychological Bulletin, 81, 284–310. [PubMed: 4608609]

Frey JN, Mainy N, Lachaux JP, Muller N, Bertrand O, & Weisz N (2014). Selective modulation of auditory cortical alpha activity in an audiovisual spatial attention task. Journal of Neuroscience, 34, 6634–6639. [PubMed: 24806688]

Gomez-Ramirez M, Kelly SP, Molholm S, Sehatpour P, Schwartz TH, & Foxe JJ (2011). Oscillatory sensory selection mechanisms during intersensory attention to rhythmic auditory and visual inputs: A human electrocorticographic investigation. Journal of Neuroscience, 31, 18556–18567. [PubMed: 22171054]

Gould IC, Rushworth MF, & Nobre AC (2011). Indexing the graded allocation of visuospatial attention using anticipatory alpha oscillations. Journal of Neurophysiology, 105, 1318–1326. [PubMed: 21228304]

Gross J, Kujala J, Hämäläinen M, Timmermann L, Schnitzler A, & Salmelin R (2001). Dynamic imaging of coherent sources: Studying neural interactions in the human brain. Proceedings of the National Academy of Sciences, U.S.A, 98, 694–699.

Grothe B, Pecka M, & McAlpine D (2010). Mechanisms of sound localization in mammals. Physiological Review, 90, 983–1012.

Haegens S, Händel BF, & Jensen O (2011). Top-down controlled alpha band activity in somatosensory areas determines behavioral performance in a discrimination task. Journal of Neuroscience, 31, 5197–5204. [PubMed: 21471354]

Haegens S, Luther L, & Jensen O (2012). Somatosensory anticipatory alpha activity increases to suppress distracting input. Journal of Cognitive Neuroscience, 24, 677–685. [PubMed: 22066587]

Haegens S, Nácher V, Luna R, Romo R, & Jensen O (2011). α-Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. Proceedings of the National Academy of Sciences, U. S. A , 108, 19377–19382.

Haegens S, & Zion Golumbic E (2018). Rhythmic facilitation of sensory processing: A critical review. Neuroscience & Biobehavioral Reviews, 86, 150–165. [PubMed: 29223770]

Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, & Lounasmaa OV (1993). Magnetoencephalography: Theory, instrumentation, and applications to noninvasive studies of the working human brain. Review of Modern Physics, 65, 413–497.

Helfrich RF, Breska A, & Knight RT (2019). Neural entrainment and network resonance in support of top-down guided attention. Current Opinion in Psychology, 29, 82–89. [PubMed: 30690228]

Helfrich RF, Fiebelkorn IC, Szczepanski SM, Lin JJ, Parvizi J, Knight RT, et al. (2018). Neural mechanisms of sustained attention are rhythmic. Neuron, 99, 854–865. [PubMed: 30138591]

Henry MJ, Herrmann B, & Obleser J (2016). Neural microstates govern perception of auditory input without rhythmic structure. Journal of Neuroscience, 36, 860–871. [PubMed: 26791216]

Henry MJ, & McAuley JD (2013). Failure to apply signal detection theory to the Montreal battery of evaluation of amusia may misdiagnose amusia. Music Perception, 30, 480–496.

Henry MJ, & Obleser J (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. Proceedings of the National Academy of Sciences, U.S.A, 109, 20095–20100.

Jensen O, & Mazaheri A (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. Frontiers in Human Neuroscience, 4, 186. [PubMed: 21119777]

Jones SR, Kerr CE, Wan Q, Pritchett DL, Hämäläinen M, & Moore CI (2010). Cued spatial attention drives functionally relevant modulation of the mu rhythm in primary somatosensory cortex. Journal of Neuroscience, 30, 13760–13765. [PubMed: 20943916]

Jones MR, Moynihan H, MacKenzie N, & Puente J (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. Psychological Science, 13, 1313–1319.

Kerlin JR, Shahin AJ, & Miller LM (2010). Attentional gain control of ongoing cortical speech representations in a "cocktail party.". Journal of Neuroscience, 30, 620–628. [PubMed: 20071526]

Klimesch W, Sauseng P, & Hanslmayr S (2007). EEG alpha oscillations: The inhibition-timing hypothesis. Brain Research Reviews, 53, 63–88. [PubMed: 16887192]

Lakatos P, Barczak A, Neymotin SA, McGinnis T, Ross D, Javitt DC, et al. (2016). Global dynamics of selective attention and its lapses in primary auditory cortex. Nature Neuroscience,19, 1707–1717. [PubMed: 27618311]

Lakatos P, Chen CM, O'Connell MN, Mills A, & Schroeder CE (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. Neuron, 53, 279–292. [PubMed: 17224408]

Lakatos P, Karmos G, Mehta AD, Ulbert I, & Schroeder CE (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. Science, 320, 110–113. [PubMed: 18388295]

Lakatos P, Musacchia G, O'Connel MN, Falchier AY, Javitt DC, & Schroeder CE (2013). The spectrotemporal filter mechanism of auditory selective attention. Neuron, 77, 750–761. [PubMed: 23439126]

Lakatos P, O'Connell MN, Barczak A, Mills A, Javitt DC, & Schroeder CE (2009). The leading sense: Supramodal control of neurophysiological context by attention. Neuron, 64, 419–430. [PubMed: 19914189]

Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, & Schroeder CE (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. Journal of Neurophysiology, 94, 1904–1911. [PubMed: 15901760]

Landau AN, & Fries P (2012). Attention samples stimuli rhythmically. Current Biology, 22, 1000–1004. [PubMed: 22633805]

Landau AN, Schreyer HM, van Pelt S, & Fries P (2015). Distributed attention is implemented through theta-rhythmic gamma modulation. Current Biology, 25, 2332–2337. [PubMed: 26279231]

Langers DR, van Dijk P, & Backes WH (2005). Lateralization, connectivity and plasticity in the human central auditory system. Neuroimage, 28, 490–499. [PubMed: 16051500]

Makeig S, Westerfield M, Jung T-P, Enghoff S, Townsend J, Courchesne E, et al. (2002). Dynamic brain sources of visual evoked responses. Science, 295, 690–694. [PubMed: 11809976]

Maris E, & Oostenveld R (2007). Nonparametric statistical testing of EEG- and MEG-data. Journal of Neuroscience Methods, 164, 177–190. [PubMed: 17517438]

Mathewson KE, Gratton G, Fabiani M, Beck DM, & Ro T (2009). To see or not to see: Prestimulus alpha phase predicts visual awareness. Journal of Neuroscience, 29, 2725–2732. [PubMed: 19261866]

Mercier MR, Foxe JJ, Fiebelkorn IC, Butler JS, Schwartz TH, & Molholm S (2013). Auditory-driven phase reset in visual cortex: Human electrocorticography reveals mechanisms of early multisensory integration. Neuroimage, 79, 19–29. [PubMed: 23624493]

Mercier MR, Molholm S, Fiebelkorn IC, Butler JS, Schwartz TH, & Foxe JJ (2015). Neuro-oscillatory phase alignment drives speeded multisensory response times: An electro-corticographic investigation. Journal of Neuroscience, 35, 8546–8557. [PubMed: 26041921]

Meyer L, Sun Y, & Martin AE (2020). Synchronous, but not entrained: Exogenous and endogenous cortical rhythms of speech and language processing. Language, Cognition and Neuroscience. 10.1080/23273798.2019.1693050.

Michalka SW, Kong L, Rosen ML, Shinn-Cunningham BG, & Somers DC (2015). Short-term memory for space and time flexibly recruit complementary sensory-biased frontal lobe attention networks. Neuron, 87, 882–892. [PubMed: 26291168]

Müller N, & Weisz N (2012). Lateralized auditory cortical alpha band activity and interregional connectivity pattern reflect anticipation of target sounds. Cerebral Cortex, 22, 1604–1613. [PubMed: 21893682]

Nobre AC, Correa A, & Coull JT (2007). The hazards of time. Current Opinion in Neurobiology, 17, 465–470. [PubMed: 17709239]

Nolte G (2003). The magnetic lead field theorem in the quasistatic approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. Physics in Medicine and Biology, 48, 3637–3652. [PubMed: 14680264]

Obleser J, & Kayser C (2019). Neural entrainment and attentional selection in the listening brain. Trends in Cognitive Sciences, 23, 913–926. [PubMed: 31606386]

Oostenveld R, Fries P, Maris E, & Schoffelen J-M (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Computational Intelligence Neuroscience, 2011, 156869. [PubMed: 21253357]

Pasinski AC, McAuley JD, & Snyder JS (2015). How modality specific is processing of auditory and visual rhythms? Psychophysiology, 53, 198–208. [PubMed: 26459153]

Percival DB, & Walden AT (1993). Spectral analysis for physical applications: Multitaper and conventional univariate techniques. Cambridge, UK: Cambridge University Press.

Pfurtscheller G, & Lopes da Silva FH (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. Clinical Neurophysiology, 110, 1842–1857. [PubMed: 10576479]

Posner MI (1980). Orienting of attention. Quarterly Journal of Experimental Psychology, 32, 3–25.

Rohenkohl G, Coull JT, & Nobre AC (2011). Behavioural dissociation between exogenous and endogenous temporal orienting of attention. PLoS One, 6, e14620. [PubMed: 21297968]

Rosenthal R, & Rubin DB (2003). r equivalent: A simple effect size indicator. Psychological Methods, 8, 492–496. [PubMed: 14664684]

Sauseng P, Klimesch W, Heise KF, Gruber WR, Holz E, Karim AA, et al. (2009). Brain oscillatory substrates of visual short-term memory capacity. Current Biology, 19, 1846–1852. [PubMed: 19913428]

Schroeder CE, & Lakatos P (2009). Low-frequency neuronal oscillations as instruments of sensory selection. Trends in Neurosciences, 32, 9–18. [PubMed: 19012975]

Stein BE, & Meredith MA (1993). The merging of the senses. Cambridge, MA: MIT Press.

Tallon-Baudry C, Bertrand O, Delpuech C, & Pernier J (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. Journal of Neuroscience, 16, 4240–4249. [PubMed: 8753885]

Talsma D, & Kok A (2002). Intermodal spatial attention differs between vision and audition: An event-related potential analysis. Psychophysiology, 39, 689–706. [PubMed: 12462498]

Taulu S, Kajola M, & Simola J (2004). Suppression of interference and artifacts by the signal space separation method. Brain Topography, 16, 269–275. [PubMed: 15379226]

ten Oever S, Schroeder CE, Poeppel D, van Atteveldt N, Mehta AD, Mégevand P, et al. (2017). Low-frequency cortical oscillations entrain to subthreshold rhythmic auditory stimuli. Journal of Neuroscience, 37, 4903–4912. [PubMed: 28411273]

ten Oever S, van Atteveldt N, & Sack AT (2015). Increased stimulus expectancy triggers low-frequency phase reset during restricted vigilance. Journal of Cognitive Neuroscience, 27, 1811–1822. [PubMed: 25941871]

Thorne JD, & Debener S (2014). Look now and hear what's coming: On the functional role of cross-modal phase reset. Hearing Research, 307, 144–152. [PubMed: 23856236]

Thorne JD, De Vos M, Viola FC, & Debener S (2011). Cross-modal phase reset predicts auditory task performance in humans. Journal of Neuroscience, 31, 3853–3861. [PubMed: 21389240]

Thut G, Nietzel A, Brandt SA, & Pascual-Leone A (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. Journal of Neuroscience, 26, 9494–9502. [PubMed: 16971533]

VanRullen R (2016). Perceptual cycles. Trends in Cognitive Sciences, 20, 723–735. [PubMed: 27567317]

VanRullen R, Busch NA, Drewes J, & Dubois J (2011). Ongoing EEG phase as a trial-by-trial predictor of perceptual and attentional variability. Frontiers in Psychology, 2, 60. [PubMed: 21716580]

VanRullen R, Zoefel B, & Ilhan B (2014). On the cyclic nature of perception in vision versus audition. Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 369, 20130214. [PubMed: 24639585]

Weisz N, Hartmann T, Müller N, Lorenz I, & Obleser J (2011). Alpha rhythms in audition: Cognitive and clinical perspectives. Frontiers in Psychology, 2, 73. [PubMed: 21687444]

Wilsch A, Henry MJ, Herrmann B, Maess B, & Obleser J (2015). Slow-delta phase concentration marks improved temporal expectations based on the passage of time. Psychophysiology, 52, 910–918. [PubMed: 25684032]

Worden MS, Foxe JJ, Wang N, & Simpson GV (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. Journal of Neuroscience, 20, RC63. [PubMed: 10704517]

Wöstmann M, Herrmann B, Maess B, & Obleser J (2016). Spatiotemporal dynamics of auditory attention synchronize with speech. Proceedings of the National Academy of Sciences, U.S.A, 113, 3873–3878.

Zoefel B, ten Oever S, & Sack AT (2018). The involvement of endogenous neural oscillations in the processing of rhythmic input: More than a regular repetition of evoked neural responses. Frontiers in Neurosciences, 12, 95.
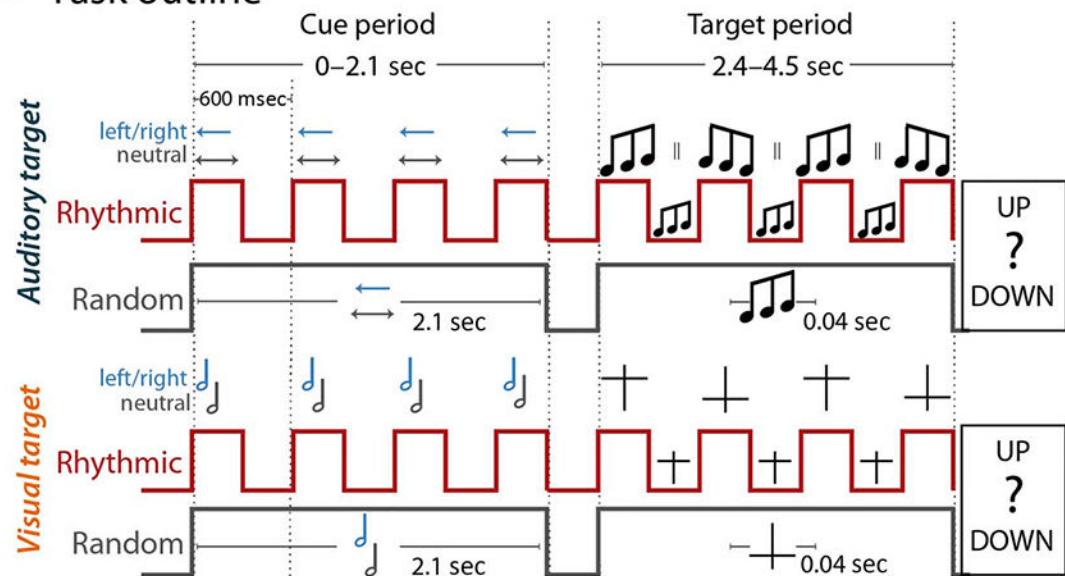
Figure 1.

Experimental design and task outline. (A) Schematic representation of the experimental manipulations: Spatial attention was modulated either with a valid cue focusing attention to the left or right (boxes with blue outline) or with a neutral cue (gray box) not providing information about the location of target occurrence. Temporal expectation was modulated either by presenting the cue rhythmically (boxes with red outline) with targets occurring in phase with the cue rhythm or by presenting the cue continuously with targets occurring at a random time point after cue presentation (gray box). (B) Trial outline separately for auditory (top) and visual (bottom) target trials and the rhythmic (red line) and random (gray line) conditions. The overall cue period ranged from 0 to 2.1 sec time-locked to cue onset. The target period started 2.4 sec after cue onset and ended 4.5 sec after cue onset. Vertical red lines display the onset and offset of the cue during the cue period illustrating the rhythmicity of cue presentation in contrast to the gray flat horizontal line in the random condition. The

arrows in the auditory-target trial as well as the note signs in the visual-target trial indicate the left/right (blue) versus neutral cue information about the target location. During the target period, the red vertical lines display the possible onset of the target in-phase and antiphase with the cue rhythm. The flat gray line indicates the nonrhythmic random presentation of the target.
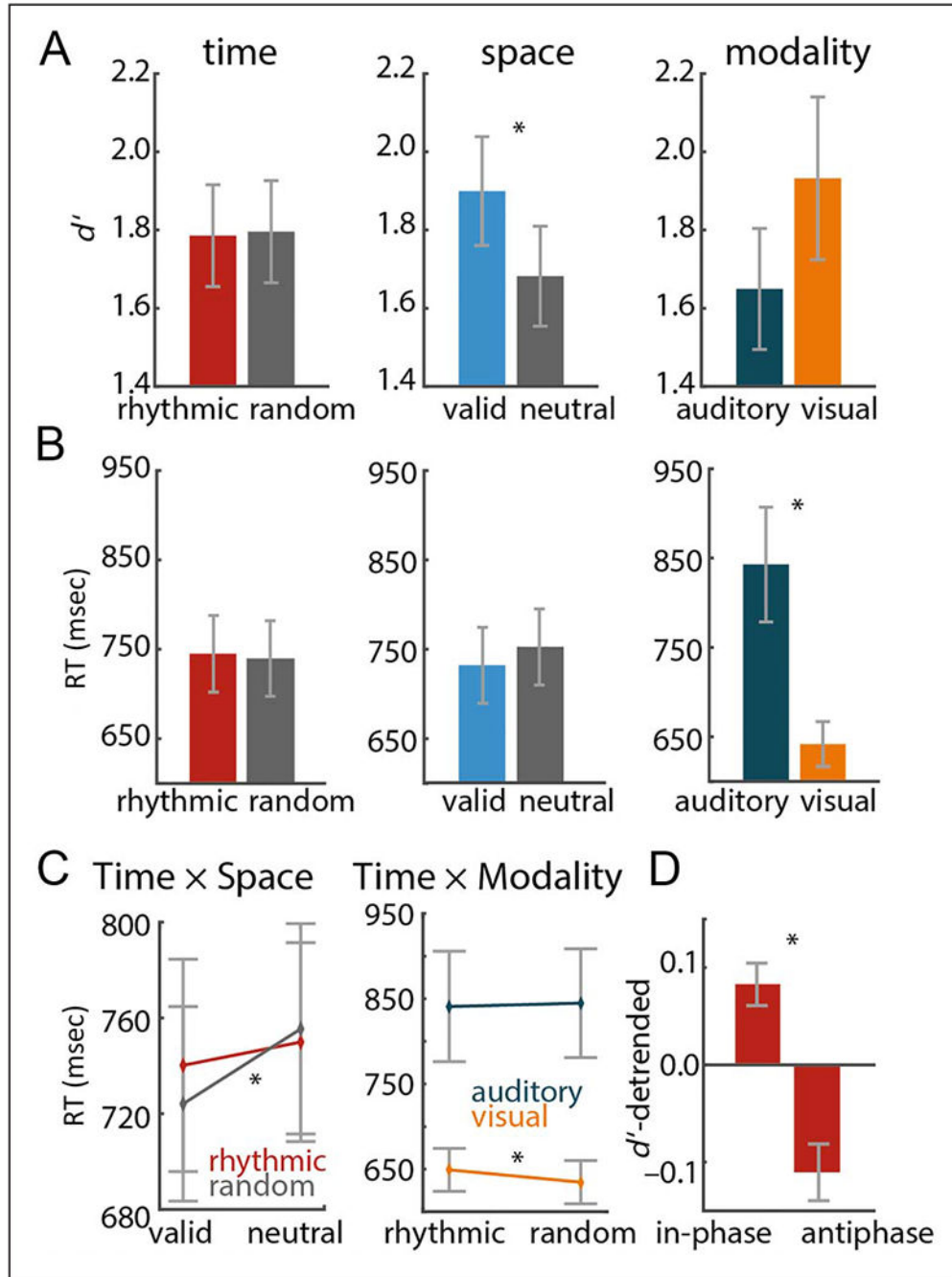
Figure 2.
Target discrimination performance. (A–C) Results of repeated-measures ANOVA (Time vs. Space vs. Modality). (A) Main effects of time, space, and modality on *d'*. (B) Main effects of RT. (C) Left: Interaction of time and space pooled across auditory and visual targets. Right: Interaction of time and modality pooled across left and right attended targets. (D) Contrast of *d'* on rhythmic in-phase and antiphase trials, pooled across auditory and visual targets. Error bars display between-participant *SEM*. Asterisks indicate significant results ($p < .05$).
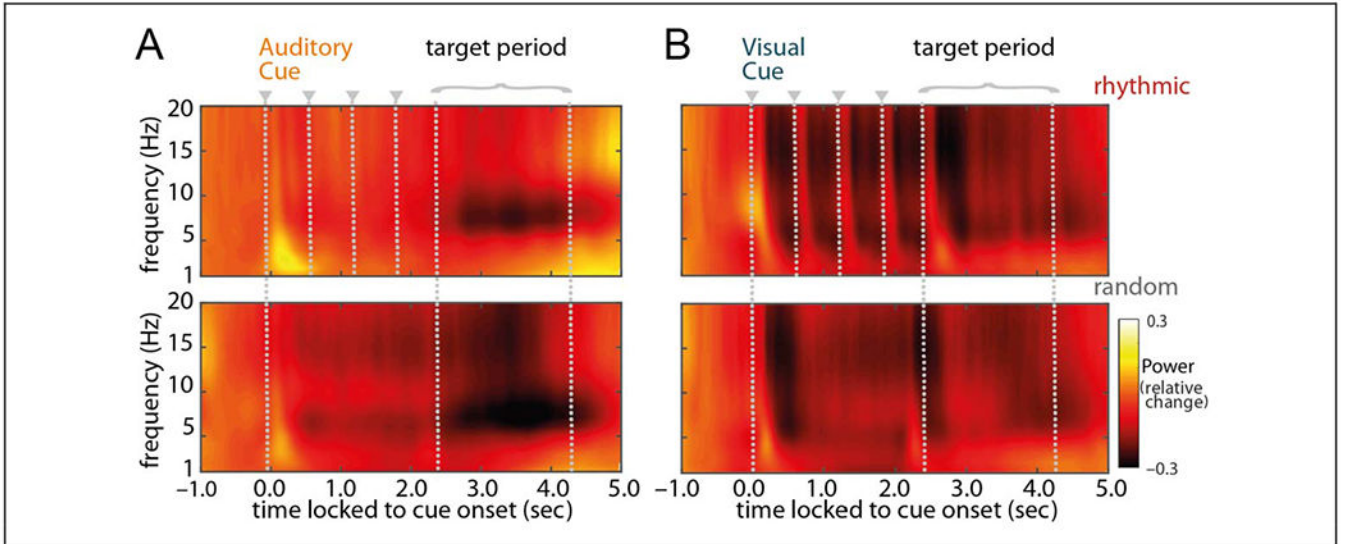
Figure 3.
Time course of grand-averaged power. TFRs from 1 to 20 Hz time-locked to cue onset, baseline-corrected with a 300-msec stimulus-free segment −0.8 to −0.5 sec before cue onset (i.e., relative change), and averaged across all sensors. Panels on the left display power averaged across all trials with an auditory cue followed by a visual target; panels on the right display power averaged across all trials with a visual cue followed by an auditory target. Top panels represent the rhythmic-cue condition; bottom panels represent the random-cue condition. Gray dotted lines mark the cue onsets (starting at $t = 0$ sec, plus following cues in the rhythmic condition) as well as the period during which the target was presented.
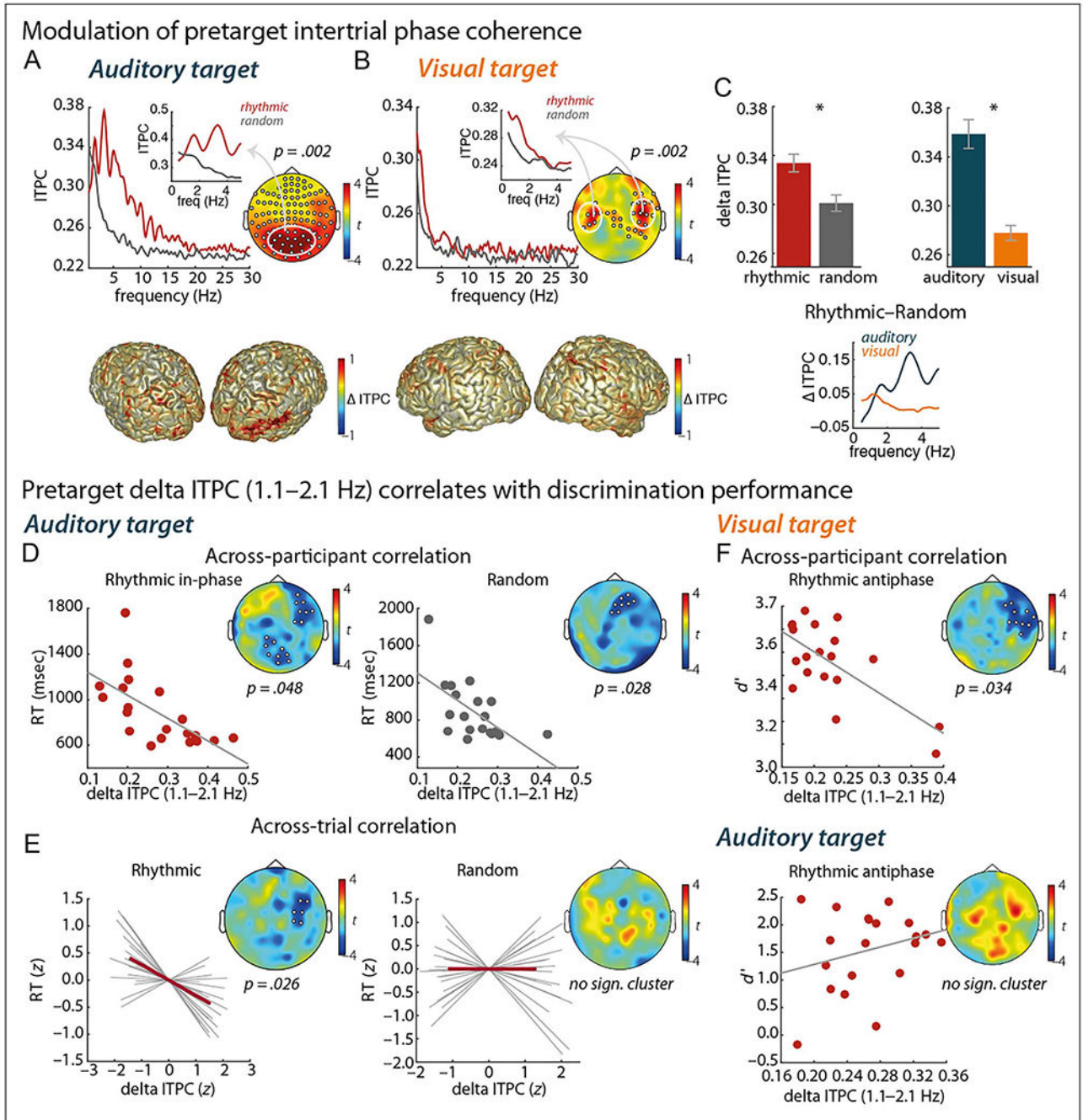
Figure 4.

Temporal expectation. (A–C) Modulation of pretarget ITPC. (A, B) Top panels display ITPC contrasted between rhythmic and random trials, for the auditory-target (A) and visual-target (B) conditions separately. Spectral plots show ITPC averaged across channels belonging to clusters with a significant difference between the rhythmic and random conditions (cluster-based permutation test; significant channels marked in the topographies). Insets show ITPC averaged across channels with significant difference centered at 1.6 Hz (i.e., 1.1–2.1 Hz), as marked by the white circle in the topographies. Bottom panels display source

reconstructions of the relative change in delta ITPC between the rhythmic and random conditions. Positive values indicate stronger ITPC for rhythmic trials. (C) Repeated-measures ANOVA (with factors Time and Modality) on delta ITPC (1.1–2.1 Hz). Error bars display between-participant *SEM*; asterisks indicate significant results ($p < .05$). Spectrum shows the difference in ITPC between the rhythmic and random trials, before auditory and visual targets separately. (D–F) Pretarget delta ITPC (1.1–2.1 Hz) correlates with discrimination performance. (D) Correlation of delta ITPC and RT across participants, for the rhythmic (left) and random (right) conditions, in the auditory-target condition. (E) Correlation of delta ITPC and RT across trials (binned per 10 trials sorted by RT), for the rhythmic (left) and random (right) conditions, in the auditory-target condition. Gray lines show single-participant correlations. (F) Correlation of delta ITPC and *d'* across participants, in rhythmic antiphase trials, for the visual-target (top) and auditory-target (bottom) conditions. Marked channels in topographies belong to a significant (sign.) cluster.
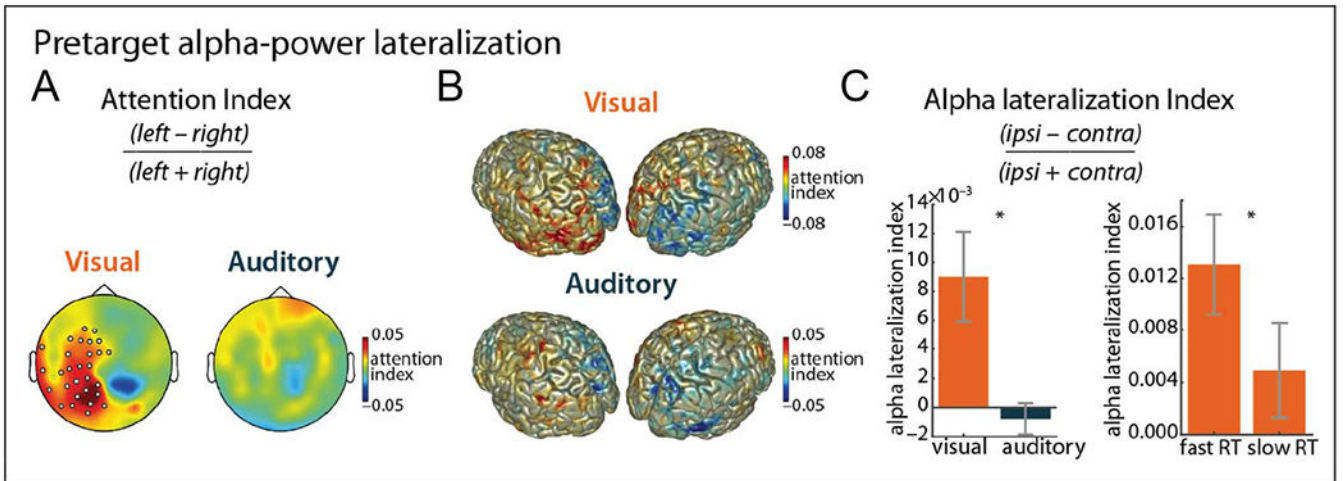
Figure 5.
Spatial attention: alpha lateralization. (A) Attention index for visual and auditory targets (8–13 Hz). Marked channels indicate a significant cluster ($p = .03$). (B) Source reconstruction of the attention index for visual and auditory targets (8–13 Hz). (C) Contrast of alpha lateralization index between auditory and visual targets (left) and between low and high RTs (median split; right) for visual-target trials. Error bars display between-participant *SEM*. Asterisks indicate significant results ($p < .05$).