Delta phase resetting mediates non-rhythmic temporal prediction

Jonathan Daume[a*], Peng Wang[a], Alexander Maye[a], Dan Zhang[b] and Andreas K. Engel[a]

[a] Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Hamburg, 20246, Germany

[b] Department of Psychology, School of Social Sciences, Tsinghua University, Beijing, 100084, China

[*] Correspondence: j.daume@uke.de; twitter: @jonathan_daume

Abstract

Neural oscillations adjust their phase towards the predicted onset of rhythmic stimulation to optimize the processing of relevant information. Whether such phase alignments can be observed in non-rhythmic contexts, however, remains unclear. Here, we recorded the magnetoencephalogram while healthy participants were engaged in a temporal prediction task judging the visual or crossmodal (tactile) reappearance of a uniformly moving visual stimulus after it disappeared behind an occluder. The temporal prediction conditions were contrasted with a working memory control condition to dissociate phase adjustments of endogenous neural oscillations from stimulus-driven activity. During temporal predictions, we observed stronger delta band inter-trial phase consistency (ITPC) in a network of sensory, parietal and frontal brain areas. Delta ITPC further correlated with individual prediction performance in parts of the cerebellum and in visual cortex. Our results provide evidence that phase alignments of low-frequency neural oscillations underlie temporal predictions in non-rhythmic unimodal and crossmodal contexts.

Introduction

In recent years, increasing evidence has been gathered that the processing of information in the brain occurs in a rhythmic, oscillatory fashion. Extracellular local field potentials as well as the macroscopically observable magneto- and electroencephalogram (M/EEG) reflect alternating transmembrane current flows from large ensembles of synchronized neurons, oscillating with frequencies ranging from very slow (<0.1 Hz) to high gamma frequencies (>200 Hz). It has been proposed that these neural oscillations reflect alternating states of higher or lower neural excitability which can modulate the efficiency with which coupled neurons engage in mutual interactions (Buzsáki, 2006; Fries, 2005). As a result, neural communication and information processing has been shown to occur in a phase-dependent manner (Engel et al., 2001, 2013; Fries, 2005), reflected for example by fluctuations in perception thresholds that correlate with the phase of ongoing oscillations (VanRullen, 2016).

Since they represent fluctuations in neural excitability, oscillations were also linked to temporal predictions of upcoming relevant information (Arnal and Giraud, 2012; Engel et al., 2001; Rimmele et al., 2018). Environmental stimuli often entail temporal regularities that make aspects of the stimulation, such as the temporal onsets of upcoming changes, highly predictable. Studies have shown that animals can utilize these predictive aspects to optimize behavior, leading to faster reaction times (Gould et al., 2011; Lakatos et al., 2008; Rohenkohl and Nobre, 2011; Stefanics et al., 2010) or enhanced stimulus processing (Cravo et al., 2013; Wilsch et al., 2015). By means of top-down induced phase resets of neural oscillations, phases of high excitability might be adjusted towards the expected onset of relevant upcoming in order to optimize behavior. When presented with a 2 Hz isochronous rhythm, for instance, the phase of oscillations within the same frequency range (and their harmonics) might align to the stimulation rhythm in such a way that phases of high excitability optimally coincide with each recurring stimulus onset. In other words, the phase of the aligned oscillation codes for the predicted time point of each upcoming stimulus onset in the rhythm. It has been suggested

that such phase alignments could form the basis of neural mechanisms that underlie temporal predictions (Arnal and Giraud, 2012; Rimmele et al., 2018; Schroeder and Lakatos, 2009).

Due to the rhythmic and therefore temporally highly predictable nature of many auditory stimuli such as speech or music, particularly in the auditory domain, many studies gathered evidence that oscillations reset and thereby adjust their phase towards rhythmic stimuli of various frequencies (Doelling and Poeppel, 2015; Giraud and Poeppel, 2012). Also in the visual domain, studies showed that neural oscillations align to rhythmic visual input (Besle et al., 2011; Cravo et al., 2013; Gomez-Ramirez et al., 2011; Herrmann, 2001; Lakatos et al., 2008; Saleh et al., 2010). However, other studies investigating temporal predictions in the visual domain observed exclusive involvements of oscillations in the alpha range (8–12 Hz) rather than phase alignments of oscillations matching the low-frequency temporal structure of the stimulation (Rohenkohl and Nobre, 2011; Samaha et al., 2015), leaving the involvement of interval-matching phase alignments in visual predictions unsettled.

Moreover, despite their ecological relevance, using rhythms for the investigation of an involvement of oscillations in temporal predictions entails methodological and conceptual challenges. Rhythms have the obvious methodological advantage that the temporal structure of the stimulation and therefore the frequencies of the oscillations that should align to the rhythmic stimulation are well-defined. However, rhythmic input also leads to a continuous stream of regularly bottom-up evoked potentials, which are – at least – difficult to distinguish from top-down phase adjusted endogenous neural oscillations within the same frequency (Doelling et al., 2019; Zoefel et al., 2018). In addition, using only rhythmic stimulation lacks the opportunity to link phase adjustments to a more general neural mechanism that predicts the temporal structure of any external input. If phase adjustments form the basis of tracking the temporal regularities of any relevant information, neural oscillations should align also to predictable temporal regularities that are inferred from input that does not itself comprise rhythmic components, such as, for instance, monotonic motion. Nevertheless, the vast majority of studies investigating phase adjustments in the context of temporal predictions presented participants with streams of (quasi-)rhythmic stimulation. Disentangling phase alignments of

neural oscillations from a continuous stream of event-related potentials in a non-rhythmic predictive context therefore constitute important aspects for examining the involvement of endogenous neural oscillations in temporal prediction processes.

In the current study, we set out to investigate whether phase adjustments of neural oscillations can be observed for non-rhythmic, but predictable visual motion stimuli. We measured MEG while healthy participants watched a visual stimulus continuously moving across the screen until it disappeared behind an occluder. We manipulated the time for the stimulus to reappear on the other side of the occluder (on average 1.5 s). The task was to judge whether the stimulus reappeared too early or too late based on the speed of the stimulus earlier to disappearance. Hence, from the time point of disappearance behind the occluder, participants were required to temporally predict the correct time point of reappearance to be able to accomplish the task. We contrasted this condition to a control task, in which participants judged the luminance of the reappearing stimulus instead of its timing. Importantly, since physical appearance of both conditions was exactly the same in all aspects of the stimulation, any purely stimulus-related, bottom-up activity should level out between the two conditions. Moreover, since it has been shown that sensory stimulation can lead to crossmodal phase adjustments also in relevant but unstimulated other modalities (Lakatos et al., 2007; Mercier et al., 2013; ten Oever and Sack, 2015), we further introduced a third condition, in which a tactile instead of a visual stimulus was presented at reappearance. By contrasting it to the working memory control condition, we sought to determine whether phase adjustments can be observed in regions associated with tactile stimulus processing, when sensory information was in fact only provided to the visual system.

We hypothesized that in the two temporal prediction tasks, as compared to working memory, we would observe stronger inter-trial phase consistency (ITPC) within time windows between disappearance and expected reappearance. Enhanced ITPC specifically in these time windows would reflect phase resets of ongoing oscillations at disappearance of the stimulus, where temporal prediction processes might be initialized. Moreover, if the phase of these oscillations indeed codes for the time point of the expected reappearance in each

participant, two further hypotheses can be formulated: (a) participants showing a more consistent judgment of reappearance timing, as represented by a steep psychometric function, should have stronger ITPC during temporal predictions than participants who performed less accurately, since a consistent timing judgment across trials should also involve a consistent phase across trials. And (b), a clustering of a specific phase within the oscillation showing the strongest ITPC in each participant should be observable at the individual subjective time points of predicted reappearance. That is, not at the actual reappearance time points of the stimulus itself but at each individual's subjective "right on time" impression, i.e., the time point at which the reappearance is expected, a clustering of a specific ("right-on-time") phase should be observable across participants. This time point can also be inferred from the psychometric function as the point of subjective equality.

Taken together, observing enhanced ITPC during temporal prediction as well as a phase clustering at the time point of expected reappearance would provide evidence that endogenous neural oscillations align to the temporally predictive structure of external stimulation in a non-rhythmic visual context. This would strongly support the hypothesis that such phase alignments form the basis of the neural mechanisms that underlie temporal prediction processes in a unimodal visual as well as crossmodal visuotactile context.

Results

Behavioral results

Participants did not receive feedback about the correctness of their response. This ensured that participants relied on their individual and subjective "right on time" (ROT) impression in the temporal prediction and "point of subjective equivalence" (PSE) in the working memory condition. Across participants, there was no statistically significant bias towards "too early/darker" or "too late/brighter" responses in the visual temporal prediction (Δt (ROT_V) = 13.15 ± 155.20 ms; $t(22) = .41$; $p = .69$) or in the working memory task (ΔRGB (PSE) = -1.29 ± 4.54 RGB; $t(22) = -1.36$; $p = .19$), respectively (Figure 1B). In the tactile temporal prediction task, participants showed a significant bias towards "too early" responses (Δt (ROT_T) = 99.80 ± 150.00 ms; $t(22) = 3.19$; $p = .004$).

Participants responded significantly faster in each of the temporal prediction tasks as compared to the working memory task (visual prediction: $t(22) = -2.55$; $p = .02$; temporal prediction: $t(22) = -4.29$; $p < .001$). To assess whether reaction times were dependent on the timing of the reappearing stimulus (Figure 1C), we averaged across all luminance differences and fitted a linear model to reaction time data in each condition. Since reaction times should be slowest for timing differences around 0 ms, where ambiguity for the timing of the reappearing stimulus was strongest, we expected reaction times to follow an inverted U-shape in the temporal prediction tasks. Therefore, we used a second-order polynomial regression with *timing difference* as predictor and tested each participant's coefficients against zero. Reaction times were significantly predicted by timing difference in all, the visual prediction (*first-order coefficient*: $-7.77 \times 10^{-4} \pm 5.27 \times 10^{-4}$, $t(22) = -7.08$, $p < .001$; *second-order coefficient*: $-1.42 \times 10^{-6} \pm 1.20 \times 10^{-6}$, $t(22) = -5.68$, $p < .001$), the tactile prediction (*first-order coefficient*: $-2.88 \times 10^{-4} \pm 4.43 \times 10^{-4}$, $t(22) = -3.12$, $p = .005$; *second-order coefficient*: $-1.26 \times 10^{-6} \pm 1.10 \times 10^{-6}$, $t(22) = -5.50$, $p < .001$) as well as in the working memory task (*first-order coefficient*: $-1.60 \times 10^{-4} \pm 1.44 \times 10^{-4}$, $t(22) = -5.31$, $p < .001$; *second-order coefficient*: $2.75 \times 10^{-7} \pm 3.51 \times 10^{-7}$, $t(22) = 3.76$, $p = .001$). Hence, although the timing of the stimulus was not

relevant in the working memory task, reaction times in that condition were (in part) also dependent on the timing of the reappearing stimulus and faster the later the stimulus reappeared.

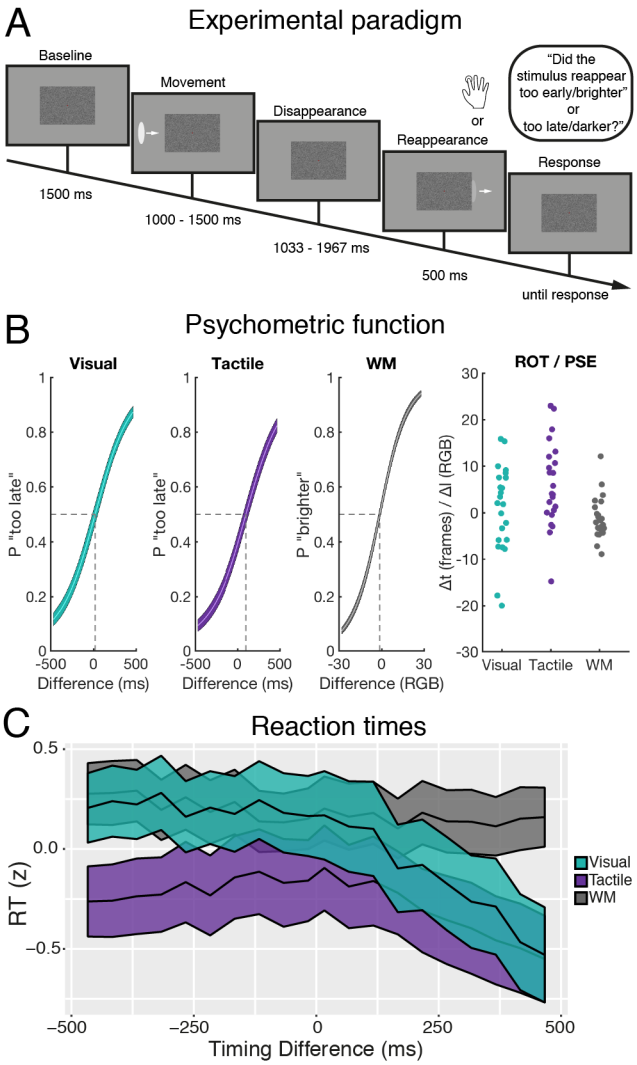

Figure 1. Experimental design and behavioral results. (A) Trial structure used in all three conditions. A stimulus was moving from the periphery towards the center of the screen until it disappeared behind an occluder. The time of the stimulus moving behind the occluder was manipulated. Task was to judge whether the stimulus reappeared *too early* or *too late*. In the working memory (control) condition, task was to judge whether the luminance became *brighter* or *darker* at reappearance of the visual stimulus. Importantly, physical stimulation was exactly the same as in the visual prediction task. In the tactile temporal prediction task, at reappearance a tactile stimulus was presented to the index finger contralateral to the disappearance of the visual stimulus. For clarity of presentation, the contrast of the occluder in this figure was increased. (B) Psychometric functions of the three conditions averaged across all participants and individual ROT/PSE estimates. A timing difference of 0 refers to the objectively correct reappearance of the stimulus after 1,500 ms. Analogously, a luminance difference of 0 refers to equal luminance after reappearance provided in RGB values (see Methods). For better comparability, individual ROT and PSE estimates are provided in frames (1 frame = 17 ms) and RGB. Colored areas around psychometric functions depict standard errors of the mean (SEM). (C) Log-transformed and standardized reaction times in all three conditions for all timing differences (mean ± SEM). Reaction times were averaged across all luminance differences. P = proportion; WM = working memory; t = time; l = luminance; RGB = red-green-blue.

Temporal prediction was associated with reduced beta power in sensory regions

Analyzing the neural data, we were first interested in investigating which frequency bands showed modulated spectral power during windows of temporal predictions. To determine frequency bands of interest, we tested an average of spectral power across all sensors and conditions against a pre-stimulus baseline window. Although we were most interested in time windows of temporal prediction, i.e., around disappearance of the stimulus, we first obtained a general overview of power modulations at each event in the experimental paradigm by computing cluster-based permutations statistics in three separate time windows centered on: (a) the onset of the moving stimulus ("Movement"), (b) disappearance of the stimulus behind the occluder ("Disappearance"), and (c) reappearance of the stimulus ("Reappearance"). All windows were normalized with a pre-movement baseline window (Figure 2A).

In time bins around movement onset as well as reappearance of the stimulus, clusters of frequencies in the theta and delta range showed a statistically significant increase of spectral power as compared to the baseline window. All time windows further depicted a significant decrease of spectral power in frequencies within the beta to lower gamma range, which extended into the higher gamma range at the end of the occluder window or the beginning of the reappearance window, respectively (all cluster p-values $< .008$). Importantly, even with using a liberal cluster alpha level of .05 (one-sided), we did not find a statistically significant modulation of delta power during the disappearance window. This was also not the case when reducing the test to sensors from occipital regions only (see Figure S1).

Since we were most interested in examining power modulations associated with temporal predictions, i.e., during the disappearance window, we compared spectral power estimates between the temporal prediction tasks and the working memory task in all sensors within the disappearance window. We restricted our analysis to the classical beta band ranging from 13 to 30 Hz, showing significant decreases as compared to baseline. Cluster-based permutation statistics revealed reduced beta power during visual temporal prediction in comparison to working memory in occipital sensors during all time-bins of the disappearance window (cluster-

186 p = .01). Source level statistics on the average across all time bins showing significant
187 differences on sensor level revealed a statistically significant decrease of beta power in a
188 cluster of bilateral occipital voxels (cluster-p = .01). A comparison between the tactile prediction
189 task and the working memory task showed that beta power was reduced during tactile
190 prediction in a cluster of occipital as well as left lateralized frontocentral sensors (cluster-p =
191 .002). In occipital sensors, beta power was reduced during all time bins, whereas the more
192 anterior power reduction evolved first in left frontal sensors during stimulus disappearance and
193 shifted towards more left-lateralized central sensors with ongoing disappearance time. At
194 source level, a significant power reduction in the beta band was most strongly apparent in parts
195 of bilateral visual as well as left-lateralized somatosensory cortex in an average across the
196 whole time window (cluster-p = .01; see Figure S2 for condition-specific beta modulations as
197 compared to baseline).

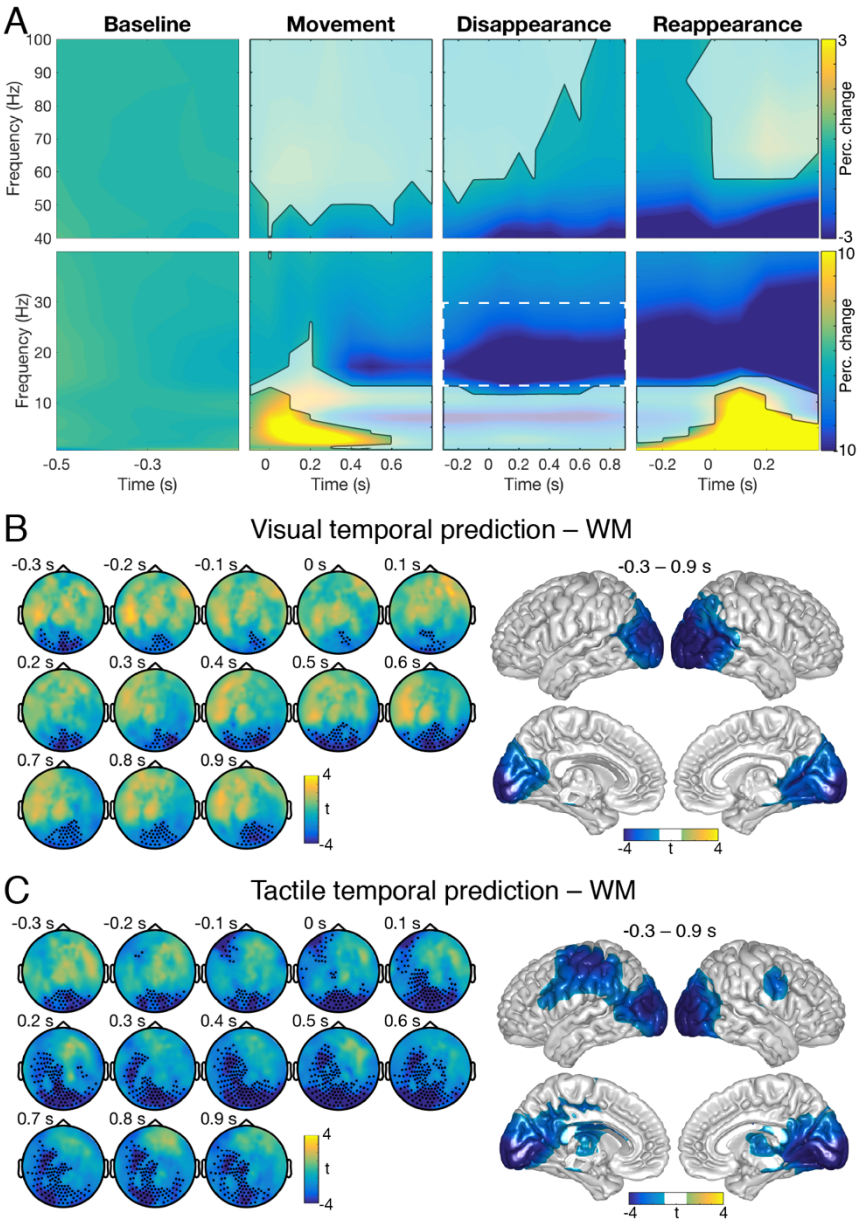

Figure 2. Power modulations during temporal prediction. (A) Spectral power for all time-frequency bins averaged across sensors, conditions, and participants. Each window was centered on the different events within the paradigm and normalized with pre-stimulus baseline. Time 0 refers to the onset of each event. Cluster-based permutation statistics were utilized to reveal significant power modulations as compared to baseline (unmasked colors). See also Figure S1. (B,C) Difference between the visual or tactile prediction and the working memory task, respectively, within the beta band (13 – 30 Hz) in time bins around stimulus disappearance. Marked sensors refer to the cluster of sensors that revealed statistically significant differences between the conditions. At source level, time bins were averaged across all bins showing significant differences at sensor level. Cluster-based permutation statistics were then employed to reveal cluster of voxels showing significant differences between the conditions (colored voxels). See also Figures S2 and S3. WM = working memory.

Inter-trial phase consistency in the delta band was stronger during temporal prediction than working memory

For the analysis of ITPC, we followed a similar approach as for the analysis of spectral power. First, we tested ITPC differences as compared to baseline in the three time windows for an average across all sensors and conditions by cluster-based permutation statistics. ITPC was significantly increased across a range of different frequencies in time bins around movement onset, disappearance and reappearance of the stimulus as compared to baseline (all cluster-$p < .001$; Figure 3A). Significant ITPC changes were broadband for the time windows centered on movement onset as well as reappearance of the stimulus with strongest increases in the delta to alpha range. At disappearance of the stimulus, significant ITPC differences to baseline were observed up to the low beta range with strongest increases in the delta and theta band.

The delta band showed strongest increases in ITPC but no increase in power as compared to baseline for an average across all conditions (see Figures 2A, 3A, and S1). Therefore, we restricted our further statistical analyses to frequencies between 0.5 to 3 Hz and time bins around disappearance of the stimulus. Differences between the two temporal prediction tasks and the working memory task were examined by computing cluster-based permutation statistics across all sensors for an average in this frequency band. For a better estimation of when differences in ITPC between the conditions became apparent, we enlarged the analysis of ITPC to time bins ranging from -1,900 ms to 1,900 ms centered on the disappearance of the stimulus. Note that in this enlarged analysis window the timing of the movement onset as well as the reappearance of the stimulus strongly jittered across trials. The effect of these events on ITPC estimates were thus strongly reduced (in comparison to the time windows that were centered on these events; see Figure S4).

By comparing the visual temporal prediction task to the working memory task, we found two clusters that showed significantly stronger ITPC during temporal predictions (Figure 3B; for clarity of presentation, only every second time bin, i.e. 200 ms, of the cluster was plotted).

One cluster included sensors from right temporal, frontal and occipital regions in time bins from -400 to 1,900 ms (cluster $p < .001$). The second cluster included left frontotemporal sensors in time bins ranging from 0 to 1,900 ms (cluster $p = .01$) Source level analysis revealed that for an average of the time window from -400 to 1,900 ms, ITPC differences between the two conditions were strongest in right-lateralized central and inferior frontal voxels (cluster $p < .001$).

For the contrast of tactile temporal prediction to working memory, we found a similar pattern of significant ITPC differences as for the contrast of visual prediction to working memory (Figure 3C). ITPC was also significantly enhanced in bilateral temporal sensors, evolving around -400 ms in right temporal sensors and shifting towards left hemisphere with ongoing disappearance time (cluster $p < .001$). In this contrast, however, differences in ITPC were more strongly apparent also in frontal and central sensors. Besides strongest differences in ITPC again in right superior parietal and inferior frontal voxels, source level analysis also revealed strong differences in bilateral somatosensory voxels for the contrast of tactile prediction to working memory (cluster $p < .001$).

Figure 3D depicts absolute ITPC estimates for all three conditions in the enlarged disappearance time window. ITPC was averaged across participants and all the sensors that exhibited the top 20% of t values in the ITPC contrast between visual temporal prediction and working memory between 0 and 1,500 ms (see Figure 3B; similar results were obtained for sensors showing the top 10% or 5% of t values, see Figure S4D). ITPC also increased in the working memory condition around disappearance of the stimulus, but dropped down to stimulus movement level shortly afterwards. ITPC in the visual as well as tactile temporal prediction tasks also decreased after an initial overshoot, but stayed elevated throughout the entire disappearance window. Importantly, enhanced ITPC estimates during temporal predictions became apparent roughly at around disappearance of the stimulus and did not exist during baseline or early movement time windows.

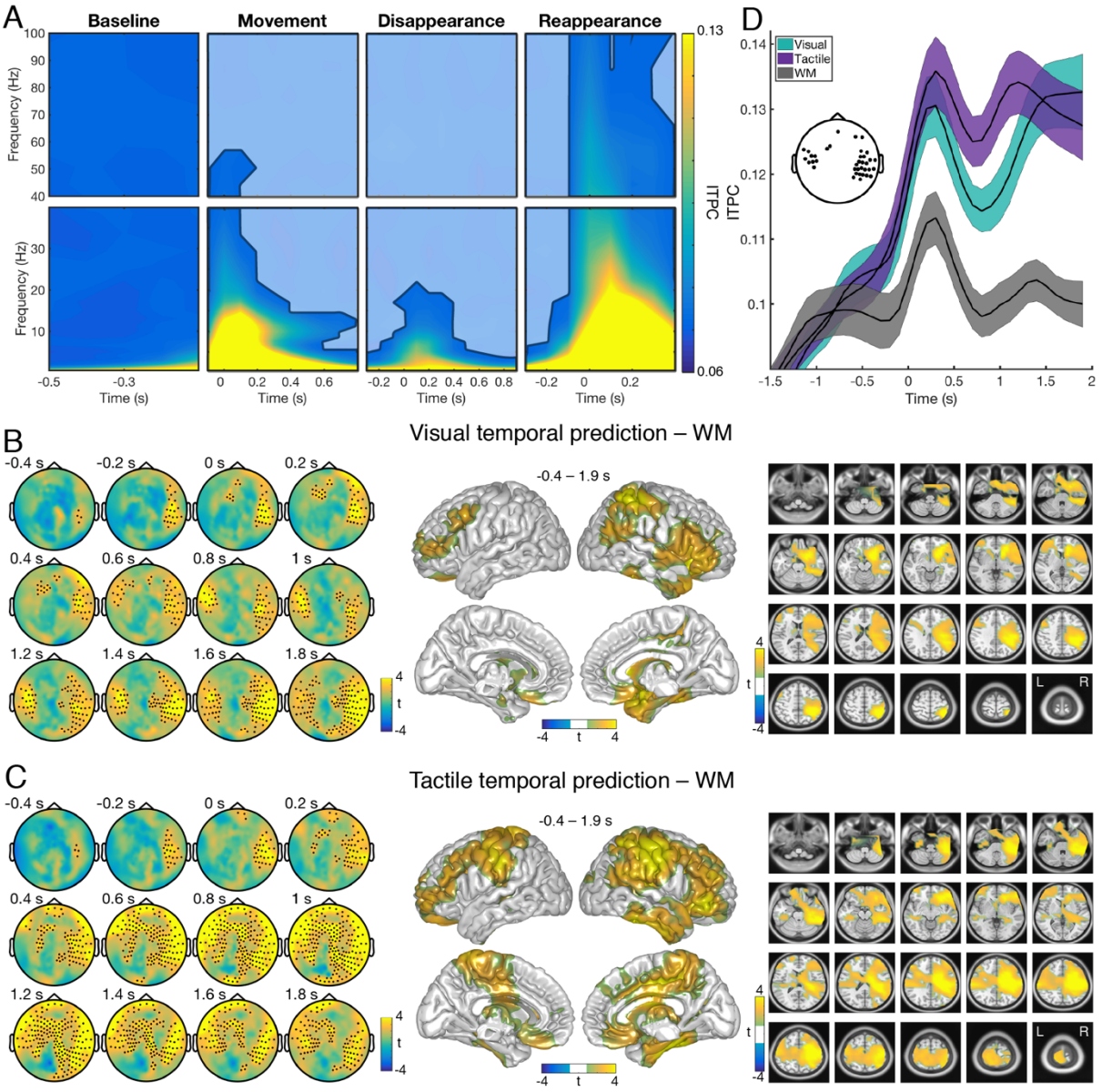

Figure 3. Delta ITPC was enhanced as compared to working memory during temporal prediction. (A) ITPC estimates for all time-frequency bins averaged across sensors, conditions, and participants. Masked colors refer to non-significant ITPC modulations as compared to pre-stimulus baseline revealed by cluster-based permutation statistics. (B,C) Difference in ITPC between the visual or tactile prediction and the working memory task, respectively, within the delta band. Marked sensors refer to the cluster of sensors that revealed significant differences between the conditions in an enlarged analysis window around stimulus disappearance (-1,900 to 1,900 ms; only every second time bin was plotted). On source level, time bins were averaged across all bins showing significant differences on sensor level and cluster-based permutation statistics were employed to reveal clusters of voxels showing significant differences between the conditions (colored voxels). See also Figures S4 and S5. (D) Time course of absolute ITPC estimates within each condition for time bins centered around disappearance of the stimulus (time 0; mean ± SEM). ITPC estimates were averaged across channels that showed the top 20% of t-values for the comparison of the visual prediction with the working memory task (see topography). WM = working memory.

Inter-trial phase consistency predicts individual behavioral performance

If the phase of neural oscillations was indeed associated with temporal predictions, a participant who judged the reappearance of the stimulus within her individual subjectively correct ROT framework in a consistent manner should exhibit stronger ITPC during temporal predictions than a participant who performed less consistently, as a consistent timing judgement across trials should involve a similar phase across trials. The consistency of judgements can be inferred from the steepness of the psychometric function. That is, the steeper the psychometric function, the more consistent the answers of the participant. To examine the relationship between individual ITPC estimates and the steepness of the psychometric functions, we computed Pearson correlations of source level delta ITPC with the steepness of the psychometric function in all voxels of the 5,003 voxels grid across participants. Using cluster-based permutation statistics, we found statistically significant positive correlations in the visual (cluster $p = .003$) as well as in the tactile temporal prediction task (cluster $p = .002$; Figure 4). Strongest correlations were found in the cerebellum and right lateralized early visual areas in both tasks. No clusters showing significant positive or negative correlations were observed in the working memory task (all cluster $p > .1$).

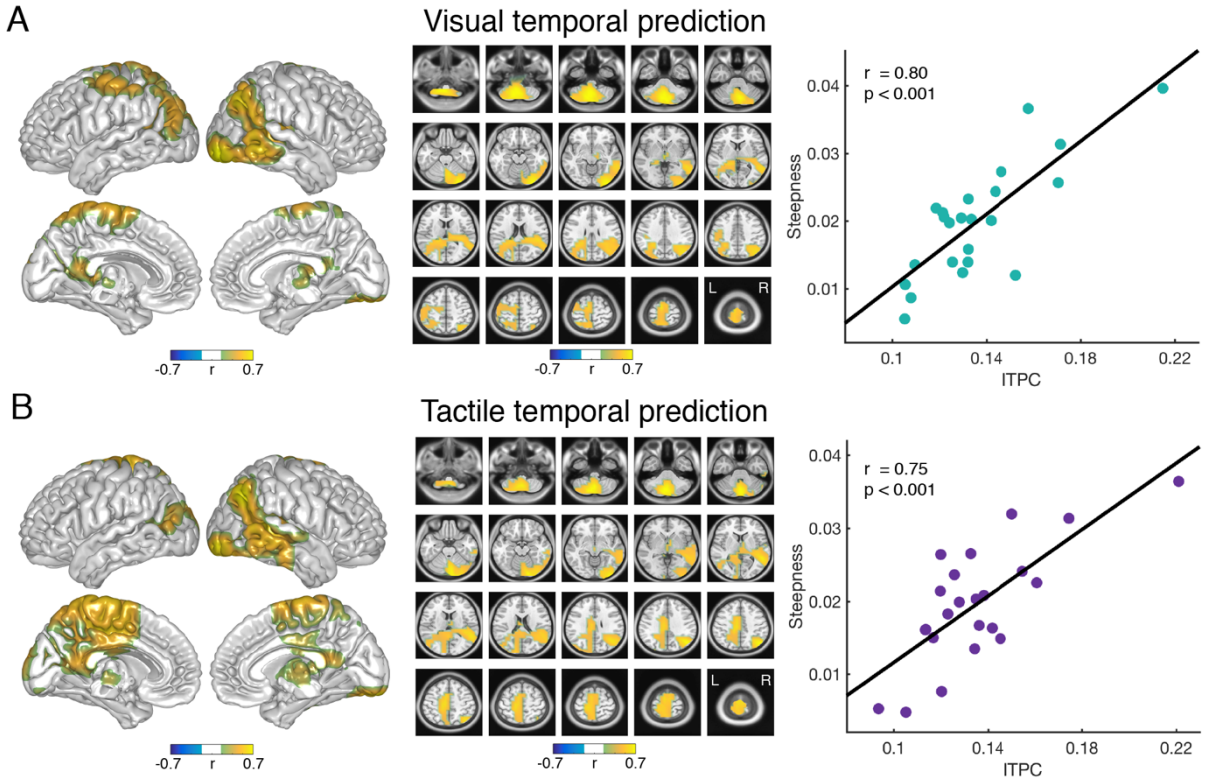

Figure 4. Correlation analysis between delta ITPC and the steepness of the psychometric function. (A,B) Correlation of individual ITPC estimates with the individual steepness of the psychometric function within all voxels, shown in (A) for the visual prediction, and in (B) for the tactile prediction condition. ITPC estimates were averaged within the delta band and time windows of 0 to 1,000 ms centered on the disappearance of the stimulus. Only the clusters of voxels showing significant correlations are colored. Each dot in the scatter plots represents one participant. ITPC estimates were averaged across all voxels within the clusters of significant correlations. There was no significant correlation observed for the working memory condition.

Delta phase clusters at individually predicted reappearance time points

One of our main interests in this study was to examine whether the phase of slow oscillations codes for the predicted time point of reappearance of the stimulus, i.e., whether a clustering of a specific low-frequency phase can be observed at each individual's ROT. In order to test that, we extracted the mean phase of that delta frequency that showed the strongest ITPC within each temporal prediction task as compared to the working memory task at ROT in each participant. In case there was no relationship between delta phase and individual ROTs, all phases extracted at ROT should be randomly distributed across the unit circle. That is, even if delta oscillations showed a phase reset at disappearance but this phase reset was not relevant for temporal predictions, the phases extracted at ROT should strongly

differ across participants, since individual ROTs strongly differed across participants as well (see Figure 1B). Moreover, if the frequency showing strongest ITPC within the frequency band of 0.5 to 3 Hz was further not related to temporal predictions, the phase extracted at ROT for these various frequencies should also vary strongly across participants.

For this analysis, we again used the sensors that showed the strongest statistical differences in ITPC for the contrasts of each prediction task to the working memory (see Figure 3B and C). Moreover, only trials in which the stimulus actually reappeared later than each individuals ROT were considered, so that stimulus onset related brain activity would not distort phase estimates at ROT. Mean phases extracted at ROT from each channel and all participants were then plotted into a histogram for each condition (Figure 5, upper row). We quantified the distance of the observed distribution to a uniform distribution by means of the modulation index (MI; Tort et al., 2010). In the working memory condition, we used individual ROTs from the visual prediction task (which employed identical stimulation) and extracted the phase from the frequency that showed the strongest ITPC as compared to the visual prediction task.

To test whether the observed MI was significantly stronger than a random distribution obtained from surrogate MIs, we repeated the analysis 10,000 times using a randomly chosen frequency from the same delta band for each participant in each repetition. We found that for both, the visual prediction (p = .03) as well as the tactile prediction task (p = 0), the observed MI was significantly stronger than the surrogate MIs. Phases at ROT from both tasks clustered roughly around ±90°. In the working memory task, no significant clustering at a specific phase was found (p = .96).

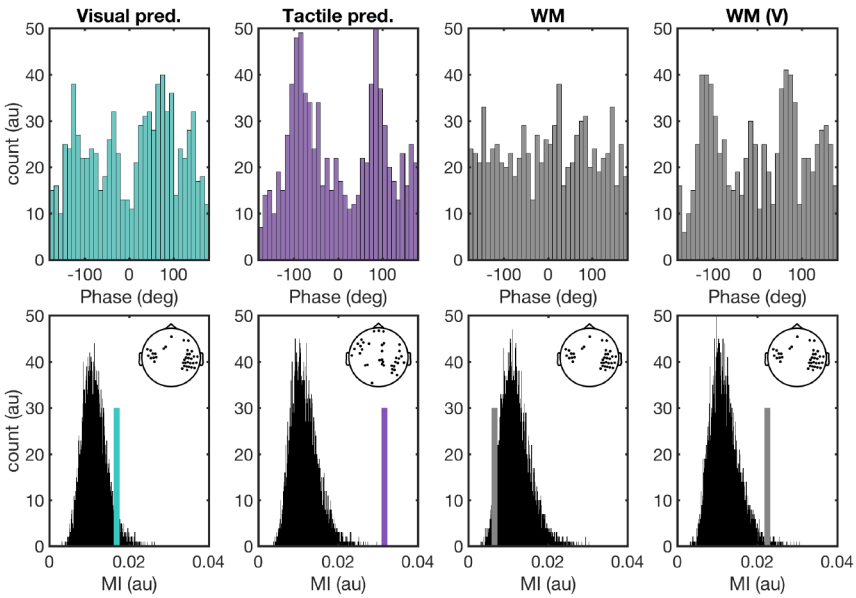

Figure 5. Delta band phase clusters at individual ROT. In each condition, the mean phase observed at individual ROT for each participant was extracted from the top 20% of channels (see Figure 2 and topographies) and from the delta frequency showing the strongest differences in ITPC to working memory. All phases were plotted into a histogram (upper panels) and the modulation index was computed from that distribution (colored line in lower panels). Permutations (n = 10,000) were generated by extracting the phase from random frequencies within the delta band (as opposed to the frequency with strongest ITPC) and computing the MI for each permutation (distribution in lower panels). WM = working memory; WM (V) = data from working memory condition but frequencies determined within the visual prediction condition (see main text); MI = modulation index.

Our reaction time analysis revealed that also in the working memory task, participants had a certain expectation about the temporal reappearance of the stimulus. Therefore, we hypothesized that the phase of the frequency that showed the strongest ITPC during the visual prediction task might also code for the timing of the reappearing stimulus in the working memory task, since physical stimulation was identical in both tasks. We repeated the above described analysis for the working memory condition, now using the same frequencies as obtained from the visual prediction condition and again tested the observed MI against 10,000 repetitions with randomly chosen frequencies (Figure 5, Panel 4: WM (V)). With frequencies obtained from the visual prediction task, the MI observed for the working memory task was significantly stronger than MIs obtained from the random repetitions (p = .02).

ITPC estimates during temporal predictions do not correlate with eye movements

One potential explanation for the observed effects in ITPC could be that participants tracked the moving stimulus with their eyes to be able to judge the correct time point of reappearance. Thus, consistent horizontal eye movements with the speed of the stimulus might lead to enhanced ITPC in the delta band. To make sure that differences in eye movements do not explain the observed differences in ITPC between the conditions, we analyzed horizontal eye movements recorded by an eye tracker (ET) during the MEG measurement. Figure 6A depicts condition-wise horizontal eye positions averaged across all participants and centered on the disappearance of the stimulus, showing no systematic differences between the conditions. Moreover, if horizontal eye movements would explain the effects in ITPC, we should observe the same effects between the conditions when we compute ITPC for the ET data. Differences in ITPC between the two temporal prediction conditions and the working memory condition are depicted Figure 6B and C. Using cluster-based permutation statistics, we did not observe any time-frequency cluster that revealed significant differences between the conditions (all cluster $p > .1$).

Further, we tested whether there are any significant correlations between individual ITPC values obtained from the MEG data and from the ET data. We averaged ITPC values from a time window of 0 to 1.500 ms and again used the top 20% of channels showing the strongest effect for ITPC for the MEG data (for channels see Figure 3D). Correlations between the data from the two measuring devices are depicted in Figure 6D. Again, we did not observe significant correlations between the ITPC values obtained from MEG and ET data. The strongest, albeit not significant correlation was found in the working memory condition, which confirms that the ITPC differences found in the MEG data cannot be explained by horizontal eye movements during temporal predictions.

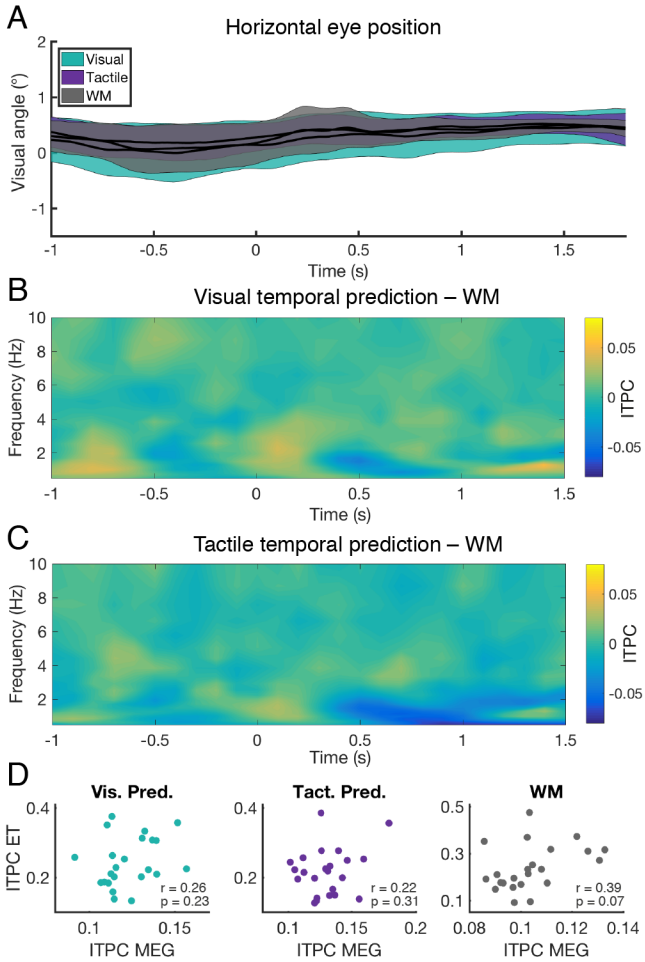

Figure 6. Analysis of horizontal eye movements. (A) Condition-wise eye positions centered on stimulus disappearance (time 0 s) and averaged across all participants. A visual angle of 0° refers to the fixation dot (1° visual angle roughly corresponds to 1 cm on the screen). Colored areas depict SEM. (B,C) Differences in ITPC between (B) the visual prediction and (C) the tactile prediction condition to the working memory condition in low frequencies and time bins around disappearance of the stimulus (time 0 s). Utilizing cluster-permutation statistics, no clusters of significant differences were observed between the conditions. (D) Condition-wise correlations between ITPC estimates obtained from the eye tracker data and the MEG sensors across all participants. ET = eye tracker.

Discussion

Our results support the idea that phase adjustments of ongoing neural oscillations could form the neuronal basis of temporal predictions and suggest that this framework can be extended to temporal predictions inferred from stimulation that does not itself comprise rhythmic components. Our task design enabled us to disentangle the phase reset of ongoing neural oscillations from evoked event related potentials and showed that phase adjustments are stronger in the context of temporal predictions than in tasks where temporal structure is less relevant. The strength of the observed phase adjustments correlated with the ability to consistently judge the temporal reappearance of the stimulus across participants. Moreover, the phase of individual delta oscillations clustered at around 90° at each participant's predicted time point of reappearance, possibly indicating an optimal phase of neural oscillations in the context of temporal prediction.

Cross-modal temporal predictions are reflected by a beta power reduction in both sensory systems

It has been suggested that temporal predictions of upcoming events might be mediated by neuronal oscillations in the delta and beta frequency range (Arnal and Giraud, 2012). The enhanced phase consistency of delta oscillations as well as the power modulations in the beta band observed in the current study are in line with this hypothesis. However, earlier reports on beta power modulations during temporal predictions are inconsistent. In a study by Fujioka *et al.* (2012), after an initial reduction in power, beta oscillations transiently re-synchronized to reach a maximum at the time point of the expected subsequent stimulus in a rhythm, increasing differentially as a function of the utilized frequency of the rhythmic stimulation. Other studies found that beta power was even increased shortly before the onset of the expected stimulus in auditory (Arnal et al., 2015; Gulberti et al., 2015) and visual rhythmic stimulation (Saleh et al., 2010). On the other hand, van Ede et al. (2011) found that predicting the onset of a tactile stimulus was specifically associated with a reduction of beta power in contralateral tactile areas

and accompanied by faster reaction times. The authors suggest that a reduction in beta power might signal preparatory processes in the sensory system that expects the upcoming event.

The observed decrease in beta power in task-relevant sensory regions in the current study largely match the results reported by van Ede et al. (2011). During visual temporal predictions, beta band power was reduced in visual sensory regions as compared to the visual control condition during the entire disappearance time. During crossmodal predictions, in which temporal information was provided to the visual system, but reappearance was expected in the tactile domain, beta band power was decreased in both, visual as well as tactile regions.

Since also in the working memory condition participants expected to perceive a visual stimulus, preparatory processes alone cannot explain this reduction in beta power. This is especially the case in the crossmodal condition, in which no visual stimulus was expected, but stronger decreases in beta were also observed in visual areas. Moreover, since we observed beta decreases also in tactile regions at the time of visual stimulus disappearance, the decrease could not solely be an effect of external stimulation.

One could argue that potential working memory maintenance processes associated with beta power increases in visual regions (Daume et al., 2017b, 2017a) could in fact explain the stronger decrease in beta power during temporal predictions. The lack of a beta modulation relative to the pre-stimulus baseline in the working memory condition would be in line with this explanation, since working memory maintenance (reflected by a beta power increase) combined with processes related to preparation (reflected by a beta power decrease) would level out beta modulations in early visual areas (see Figure S2). This could, however, not explain the observed beta power reduction in visual areas during the tactile temporal prediction condition, since no visual stimulation was expected here.

Beta decreases observed during temporal predictions might therefore relate to more than only to preparatory processes to an upcoming stimulus. Cross-modal decreases in beta band activity in both the temporal information providing visual as well as the stimulation expecting tactile system might reflect that both sensory modalities are continuously involved in temporal

prediction processes, not only in processes preparing for the upcoming stimulation. We found no significant increases in beta power during temporal predictions, even if the time window was centered on the time point of predicted reappearance (ROT) in each participant in either of the two prediction conditions (see Figure S3). Whether decreases in beta power are associated with non-rhythmic temporal predictions while increases might reflect temporal predictions during rhythmic stimulation, remains subject to future research.

Neural oscillations at low frequencies adapt to the temporal structure of visual moving stimuli

Seminal work by Lakatos et al. (2008) showed that the phase of delta oscillations in visual area V1 of monkeys followed the attended regular stream of either visual or auditory rhythmic stimulation. Moreover, reaction times towards a target stimulus varied as a function of the phase of delta oscillations. Such phase alignments could form the basis for temporal prediction processes to facilitate processing of predictive upcoming stimulation (Arnal and Giraud, 2012; Schroeder and Lakatos, 2009). Similar results have also been found in studies of the human brain (Besle et al., 2011; Gomez-Ramirez et al., 2011).

Phase entrainment of neural oscillations does not only occur in the delta band but can flexibly adapt to the frequency of the external input also at higher frequencies such as the theta or the alpha band. Doelling and Poeppel (2015) presented participants with rhythmic, musical stimuli composed of peak note rates varying from 0.5 to 8 notes per second and found neural oscillations with frequencies matching the different peak note rates to show increased phase entrainment to the stimuli. Moreover, multiple rhythmic streams presented at different frequencies can be simultaneously tracked by neural oscillations in different frequency bands, and behavior is especially enhanced when the phases from both rhythms coincide (Henry et al., 2014). This is specifically important in speech processing, where neural oscillations can track the complex spectrum of spoken language (Giraud and Poeppel, 2012). Accordingly, it

457 has been shown that the frequency of the entrained rhythm can modulate the comprehension
458 of spoken words (Kösem et al., 2018).

459 To serve as a potential mechanism to temporally predict the onset of relevant upcoming
460 information, the entrainment of neural oscillations from different frequencies is crucial in order
461 to flexibly adapt to the naturally occurring temporal regularities. However, in the visual system,
462 evidence for the tracking of temporally predictive input by neural oscillations from different
463 frequency bands is not as clear. On the one hand, studies showed that the phase of delta
464 oscillations is involved in temporal predictions of visual input (Cravo et al., 2013; Saleh et al.,
465 2010; Wilsch et al., 2015). On the other hand, studies suggested that temporal predictions in
466 the visual system were specific to the alpha band. Rohenkohl and Nobre (2011), for instance,
467 used rhythmically presented visual stimuli at 2.5 and 1.25 Hz moving across the screen until it
468 disappeared behind an occluder. Neural oscillations exclusively from the alpha band showed
469 modulated activity associated with temporal predictions during the disappearance time. They
470 found no phase locking of oscillations in lower frequencies. Moreover, results reported by
471 Samaha et al. (2015) suggest that specifically the phase of alpha oscillations is modulated to
472 predict the onset of a visual stimulus.

473 In the current study, we provide further evidence that neural oscillations from the delta
474 band show enhanced phase alignment during a window of explicit visual temporal predictions
475 across trials. In order to adapt to the temporal regularity of the presented visual stimulus, delta
476 frequencies in a wide network of parietal and frontal brain areas exerted more consistent phase
477 resets at around the time point of disappearance of a visual stimulus as compared to a working
478 memory control condition. The strength of this phase adjustment in each participant correlated
479 with the consistency in judging a reappearance of the visual stimulus as too early or too late.
480 This was the case only in the temporal prediction tasks, which underlines the behavioral
481 relevance of the observed phase adjustments for temporal predictions.

482 Moreover, by providing no feedback about the correctness of their response, we made
483 sure that participants used individual time points at which they subjectively expected the
484 stimulus to reappear. Within each participant's neural oscillation that showed the strongest

485 ITPC during temporal predictions, we found a clustering of phases roughly around ±90° at
486 each participant's ROT. This was not the case when using the frequencies showing the
487 strongest ITPC in the working memory condition, where timing was not as important. That is,
488 within each individual's subjective temporal framework, neural oscillations adjusted their phase
489 to the external stimulation such that a phase of 90° eventually coincided with each individual's
490 predicted time point of reappearance. This provides strong support for the notion that in the
491 context of temporal predictions the phase of delta oscillations adjusts to the temporal structure
492 of the stimulation to code for the timing of the predicted reappearance. Our results are in line
493 with results reported by Cravo *et al.* (2013), who showed that contrast sensitivity was a function
494 of the phase of entrained delta oscillations. In their study, the strongest contrast sensitivity for
495 visual stimuli was also observed at a delta phase around 90°. This phase range might therefore
496 indicate an optimal phase for processes related to temporal prediction.

497 Only when using the frequencies that exposed the strongest ITPC in the visual prediction
498 task also for the analysis of phase clustering in the working memory task, we observed a
499 significant phase clustering across all participants again roughly around ±90° as well. The
500 results of the behavioral data suggest that the temporal structure of the stimulation was not
501 totally irrelevant to the control task (see Figure 1C). Therefore, phase adjustments in neural
502 oscillations related to the temporal structure might have also occurred within this task (see
503 Figure S5 for condition specific ITPC), but they were less consistent than the phase
504 adjustments observed in the temporal prediction tasks.

505 Importantly, our study suggests that the mechanism of phase adjustments for temporal
506 predictions can be extended to external stimulation that does not as such involve rhythms. We
507 found that low-frequency oscillations can adjust their phase also to the temporal structure of
508 external stimulation that had to be inferred from motion. Many natural stimuli are composed of
509 highly predictable regularities, but not all of them are intrinsically rhythmic. Our results
510 therefore indicate that the framework of phase adjustments during temporal predictions might
511 be generalized to all forms of predictive stimulation.

Enhanced ITPC during temporal predictions is associated with phase adjustments in endogenous neural oscillations

In earlier investigations of phase adjustments to external stimulation participants were mostly presented with streams of rhythmic input. An undeniable advantage of using rhythms is the well-defined frequency range of interest for analyzing the recorded brain activity. However, rhythmic input also causes evoked brain activity within the same frequency range, which makes it difficult to disentangle streams of evoked activity from entrained endogenous neural oscillations (Doelling et al., 2019; Zoefel et al., 2018).

Our results provide further evidence that phase resets of low-frequency oscillations observed during temporal predictions cannot solely be explained by stimulus-evoked, bottom-up brain activity (see also, Doelling et al., 2019; Kösem et al., 2018; ten Oever et al., 2017). In the current study, we aimed at reducing such brain responses to a minimum by presenting participants with a continuously moving stimulus instead of several discrete stimuli. We were particularly interested in the time point at which the stimulus transiently disappeared behind an occluder (as opposed to sharp onsets and offsets in rhythms). At disappearance, we did not observe an increase in low-frequency power as compared to pre-stimulus baseline in any of the conditions studies, which could have been associated with stimulus-evoked brain activity. Moreover, by introducing a control condition in which physical stimulation was exactly the same as during temporal predictions, we further aimed at controlling for brain responses that were not specific to temporal predictions. Importantly, delta power was not stronger during temporal predictions as compared to the working memory task (see Figure S1).

We observed enhanced ITPC during temporal predictions as compared to the working memory condition as well as to the pre-stimulus baseline window at the time point of disappearance. This strongly favors the notion that ongoing, endogenous neural oscillations underwent a phase reset around the time point of disappearance, which was more consistent during temporal predictions than during working memory processes. These phase resets can

538 therefore not be solely related to brain responses evoked by the offset of the visual movement,
539 since we did not observe power differences at low frequencies.

540 Furthermore, we introduced a crossmodal temporal prediction task, in which a visual
541 stimulus disappeared behind the occluder, while participants had to judge the timing of an
542 upcoming tactile stimulus. In this condition, we observed enhanced ITPC in somatosensory
543 areas at around the time point of disappearance of the visual stimulus. Since stimulation was
544 purely visual at that time, evoked brain activity could not explain the observed effects in ITPC.
545 However, we did not observe enhanced ITPC in early visual areas at any time point during
546 temporal predictions as compared to the working memory control condition (discussed further
547 below).

548 Phase resets occurred in a network of frontoparietal and sensory brain areas

549 We observed enhanced ITPC values in a network of mostly frontal and parietal brain areas
550 during visual as well as crossmodal temporal predictions. Strongest ITPC values were
551 observed in superior parietal and inferior frontal cortex contralateral to stimulus disappearance
552 during both, visual as well as crossmodal temporal predictions. During tactile temporal
553 predictions, strong ITPC values were also observed in bilateral somatosensory as well as in
554 inferior parietal cortex contralateral to the predicted tactile stimulation. Similarly, Besle et al.
555 (2011) observed significant phase entrainment to audiovisual stimulation in a wide network of
556 distributed areas including parietal and inferior frontal areas. These observations support the
557 notion that brain areas involved in temporal predictions may constitute a frontoparietal timing
558 network (Coull and Nobre, 2008; Rimmele et al., 2018; Wiener et al., 2010).

559 The fact that strong ITPC differences were not only observed in areas contralateral to the
560 disappearance, but also in areas contralateral to the predicted reappearance of the stimulus
561 in the opposite hemifield suggests that phase adjustments might not only reflect processes of
562 temporal, but also processes of spatial predictions. Similarly, we observed enhanced ITPC
563 values also in early somatosensory areas contralateral to the disappearance of the purely

visual stimulus during crossmodal temporal predictions, despite the fact that prediction-relevant information was provided only by a moving visual stimulus. This supports evidence reported earlier showing that stimulation within one modality can crossmodally reset the phase of ongoing low-frequency in other modalities, which might be an important mechanism for multisensory integration processes (Lakatos et al., 2007; Mercier et al., 2013; ten Oever and Sack, 2015). Similarly, during crossmodal predictions we observed enhanced ITPC also in somatosensory areas contralateral to the expected tactile stimulation. Phase adjustments in a distributed network of areas might therefore reflect temporal as well as spatial or crossmodal predictions in areas that are relevant for providing as well as receiving the predictable information.

Along the same line, we expected to find enhanced ITPC during temporal predictions in early visual areas, especially as the visual system was the only modality explicitly confronted with the external temporal information. In fact, strong ITPC estimates was also observed in occipital sensors as compared to baseline in each condition (see Figure S5), but they were not different between the conditions. One reason for this could be that processes not related to temporal prediction and therefore equal in all conditions could have overshadowed effects in the visual system. However, we found that voxels in early visual areas showed strong correlations between individual ITPC estimates and the steepness of the psychometric function in both temporal prediction tasks, but not in the working memory task. This suggests that consistent phase resets of delta oscillations within visual areas might have supported consistent timing judgments with the participants' subjective timing frameworks. This indicates a critical involvement of the visual system also in processes related to temporal prediction.

Moreover, we observed strong correlations between ITPC and behavior in the cerebellum, supporting earlier reports on a involvement of the cerebellum in temporal prediction processes (Breska and Ivry, 2016; Ivry and Keele, 1989). Roth and coworkers (2013), for instance, found that cerebellar patients were significantly impaired in recalibrating sensory temporal predictions of a reappearing visual stimulus. This finding is of particular interest as we adapted the authors' experimental paradigm for the use in the current study. Theirs and our results

592 therefore indicate that the cerebellum might be crucially involved in accurate and consistent
593 judgments of temporal regularities deployed in perceiving object motion.

594 Conclusions

595 We provide evidence that the phase of neural oscillations can adjust to the temporal
596 regularities of external stimulation. Such phase alignments could provide a key mechanism
597 that predicts the onset of upcoming events in order to optimize processing of relevant
598 information and thereby adapt behavior. Our results further reveal that this concept of phase
599 adjustments during temporal predictions can be extended to non-rhythmic, but predictable
600 visual motion stimuli, which suggests that phase adjustments could be a general mechanism
601 for temporal prediction processes. In a crossmodal setting, we show that temporal information
602 provided to one modality leads to phase adjustments in another modality when crossmodal
603 temporal predictions are necessary. Such crossmodal phase resets could be the neuronal
604 basis of multisensory integration processes. Moreover, by introducing a physically matched
605 control condition, our results support the notion that phase alignments observed during
606 temporal predictions are based on phase resets in ongoing neural oscillations and do not arise
607 as a byproduct of bottom-up stimulus processing. Importantly, we observed that these phase
608 adjustments occurred on an individual level to match each individual's subjective temporal
609 predictions time points. Additionally, the more consistent a participant was in estimating the
610 time of reappearance of the stimulus, the higher the phase alignment was across trials. Taken
611 together, our results provide important insights into the neural mechanisms that might be
612 utilized by the brain to predict the temporal and spatial onsets of upcoming events.

613 # Methods

614 ## Participants

615 Twenty-three healthy volunteers (mean age ± standard deviation (SD): 27.13 ± 4.30 years;
616 20 females; all right-handed) took part in the study. They gave informed written consent and
617 were monetarily compensated with 13 €/hour for participation. All volunteers had normal or
618 corrected-to-normal vision, normal touch, as well as no background of psychiatric or
619 neurological disorder. The ethics committee of the Medical Association Hamburg approved the
620 study protocol (PV5073), and the experiment was carried out in accordance with the approved
621 guidelines and regulations.

622 ## Experimental procedure

623 The experimental paradigm used in the current study was adopted from an earlier report
624 investigating visual temporal predictions in cerebellar patients (Roth et al., 2013). Our
625 experiment consisted of three conditions: a *visual* temporal prediction task, a crossmodal
626 (*tactile*) temporal prediction task, and a *working memory* (control) task. The trials of all
627 conditions started with the presentation of a randomly generated, white noise occluder (size:
628 7.5° x 11.3° (h x w)) that was smoothed with a Gaussian filter (imgaussfilt.m in MATLAB) and
629 presented in the middle of the screen against a grey background screen (luminance: 44 cd/m^2;
630 corresponds to 115 red-green-blue (RGB) values in our setting; see Figure 1A). At the center
631 of the occluder, a red fixation dot was presented. We instructed participants to fixate this dot
632 throughout the entire trial. After 1500 ms, an oval stimulus (size: 3.5° x 1.0°) set on in the
633 periphery of the screen, moving towards the occluder with a speed of 6.9 °/s. The luminance
634 of the stimulus differed in all trials between 120 to 161 cd/m^2 (6 steps, counterbalanced,
635 corresponds to 170 to 220 RGB). For half of the participants, the stimulus started on the left
636 side of the occluder and moved from left side towards the right side. For the other half, the
637 stimulus started on the right side and moved from right to left. The direction of movement was

kept constant for each participant throughout the entire experiment. In each trial, the starting point of the stimulus differed such that the stimulus took 1,000 to 1,500 ms to disappear completely behind the occluder from starting point, randomly jittered with 100 ms (counterbalanced). The size of the occluder and the speed of the stimulus were chosen so that the stimulus would need exactly 1,500 ms to reappear on the other side of the occluder. However, we manipulated the timing and the luminance of the reappearing stimulus. In each trial, the reappearance of the stimulus differed between ±17 to ±467 ms (randomly jittered, but counterbalanced in steps of 50 ms; corresponds to ±1 to ±28 frames with a jitter of 3 frames at 60 Hz) from the correct reappearance time of 1,500 ms. Hence, the stimulus was covered by the occluder for 1,033 to 1,967 ms and was reappearing at 20 different time points. In the visual prediction task as well as in the working memory task, we also manipulated the luminance of the reappearing stimulus relative the luminance the stimulus had before disappearance in each trial (jittered, but counterbalanced between ±1 to ±40 cd/m^2, also using 20 different values; corresponds to ±1 to ± 28 RGB in steps of 3 RGB to make it similar to the timing manipulation). After reappearance, the stimulus moved to the other side of the screen for 500 ms with the same speed until it set off the screen. The occluder was presented throughout the entire trial.

By manipulating the timing as well as the luminance in both conditions, we made sure that both, the visual temporal prediction as well as the working memory task had the exact equal physical appearance throughout all trials. They only differed in their cognitive set. In the visual temporal prediction task, we asked participants to judge whether the stimulus was reappearing *too early* or *too late* based on the speed the stimulus had earlier to the occluder (which was kept constant throughout the entire experiment). In the working memory task, participants were asked to judge whether the luminance of the reappearing visual stimulus became *brighter* or *darker* as compared to the stimulus earlier to disappearance. Participants answered by pressing one of two buttons with their index or middle finger of the hand contralateral to the reappearing stimulus.

The tactile temporal prediction task was equal to the visual temporal prediction task, with the only difference that a tactile stimulus instead of a visual was presented at the time of reappearance to the right or left index finger (depending on which side the stimulus was expected to reappear behind the occluder). The tactile stimulus was presented by means of a Braille piezostimulator (QuaeroSys, Stuttgart, Germany; 2 x 4 pins, each 1 mm in diameter with a spacing of 2.5 mm), pushing up all eight pins for 200 ms. At that time, nothing happened on the screen. Participants gave their answer with the same hand as in the other two conditions (i.e., with the hand that was not stimulated by the Braille stimulator). Response mapping of the two buttons was counterbalanced across all participants. As soon as participants gave their answer, the fixation dot turned dark grey for 100 ms to indicate that the response was registered. However, participants did not receive trial-wise feedback about the correctness of their response. After a short delay of 200 ms, the white-noise occluder was randomly re-shuffled to signal the start of a new trial.

All three conditions were presented block-wise. At the beginning of each block, participants were informed about the current task. The order of presentation of the conditions was kept constant for each participant, but was randomized across participants (counterbalanced). At the end of each block, they were informed about the overall accuracy of their answers within the last block and were allowed to rest as long as they wanted. Each participant performed two sessions at two different recording days. The experimental procedure was kept constant across both sessions, i.e., movement direction, response mapping, as well as condition order did not change in the second session for individual participants. Each session comprised twelve blocks, i.e., four blocks per condition. Each block consisted of 60 trials, resulting in a total number of 480 trials per condition or 1,440 trials in total. Due to technical difficulties, for one participant we only acquired data from one session with a total number of 720 trials.

At the beginning of each recording day, participants performed a short training of all conditions to get familiar with the overall experimental procedure and the stimulus material. This training took part in the same environment as the subsequent recording session. At the

693 end of the second recording day, participants filled a questionnaire asking for any specific
694 strategy they might have used for the temporal prediction task.

695 We used MATLAB R2014b (MathWorks, Natick, USA; RRID: SCR_001622) and
696 Psychtoolbox (Brainard, 1997; RRID: SCR_002881) on a Dell Precision T5500 with Ubuntu
697 64-bit operating system (Version: 16.04.5 LTS) for stimulus presentation. The visual stimuli
698 were projected onto a matte backprojection screen at 60 Hz with a resolution of 1,920 × 1,080
699 pixels positioned 65 cm in front of participants. To mask the sound of the Braille stimulator
700 during tactile stimulation, we presented participants with auditory pink noise at sampling rate
701 of 48 kHz and volume of 85 dB using MEG-compatible in-ear headphones (SRM-252S, STAX
702 Limited, Fujimi, Japan) during all experimental blocks.

703 Data acquisition and pre-processing

704 MEG was recorded at a sampling rate of 1,200 Hz using a 275-channel whole-head
705 system (CTF MEG International Services LP, Coquitlam, Canada) situated in a dimly lit, sound
706 attenuated and magnetically shielded chamber. We additionally recorded electrical eye,
707 muscle and cardiac activity with Ag/AgCl-electrodes in order to have a better estimate for
708 endogenous artefacts. Online head localizations (Stolk et al., 2013) were used to navigate
709 participants back to their original head position prior to the onset of a new experimental block
710 if their movements exceeded five mm from their initial position. The initial head position from
711 the first recording day was saved so that participants could be navigated back to their initial
712 head position also during the second recording day. This assured comparable head positions
713 of each participant across sessions. Five malfunctioning channels were either not recorded or
714 excluded from further analysis for all participants. To further control for eye movement artifacts,
715 eye movements were tracked with an MEG-compatible EyeLink 1000 Long Range Mount
716 system (SR Research, Osgoode, Canada).

717 We analyzed reaction time data using R (R Core Team, 2014; RRID: SCR_001905) and
718 RStudio (RStudio Inc., Boston, USA; RRID: SCR_000432). Trials with reaction times longer

719 than three standard deviations were excluded from analysis. Due to the right-skewed nature
720 of reaction times, reaction time data were first log-transformed and then standardized across
721 all trials from each participant.

722 All other data were analyzed using MATLAB R2016b with FieldTrip (Oostenveld et al.,
723 2011; RRID: SCR_004849), the MEG and EEG Toolbox Hamburg (METH, Guido Nolte; RRID:
724 SCR_016104), or custom made scripts. The physiological continuous recording of each
725 session was first cut into epochs of variable length. Each trial was cut 1,250 ms earlier to
726 stimulus movement onset and 1,250 ms after offset of the reappeared stimulus. Trial length
727 therefore varied between 4,717 and 6,183 ms. To prevent that the timing in a given trial was
728 not exactly as intended, e.g., by short movement interruptions of the stimulus, we removed
729 trials which contained MEG marker timings that differed from the intended timing of the moving
730 stimulus in the trial by at least one frame (17 ms). Thus, we excluded on average 1.2 trials in
731 each participant and each session (range: 0 – 24 trials).

732 Moreover, trials containing strong muscle artifacts or jumps were detected by semi-
733 automatic procedures implemented in FieldTrip and excluded from analysis. The remaining
734 trials were filtered with a high-pass filter at 0.5 Hz, a low-pass filter at 170 Hz, and three band-
735 stop filters at 49.5–50.5 Hz, 99.5–100.5 Hz and 149.5–150.5 Hz and subsequently down-
736 sampled to 400 Hz.

737 We performed an independent component analysis (infomax algorithm) to remove
738 components containing eye-movements, muscle, and cardiac artefacts. Components were
739 identified by visual inspection of their time course, variance across samples, power spectrum,
740 and topography (Hipp and Siegel, 2013). On average, 25.7 ± 8.6 components were rejected in
741 each participant and each session. All trials were again visually inspected and trials containing
742 artefacts that were not detected by the previous steps were removed.

743 As a final step, using procedures described by Stolk *et al.* (2013) and online
744 (http://www.fieldtriptoolbox.org/example/how_to_incorporate_head_movements_in_MEG_an
745 alysis/) we identified trials in which the head position of the participant differed by 5 mm from

746 the mean circumcenter of the head position from the whole session (on average: 2.6 trials per
747 participant and session, range: 0 – 86 trials) and excluded them from further analysis. 670.2 ±
748 26.7 trials of the total of 720 trials remained from pre-processing on average per participant in
749 each session.

750 Quantification and statistical analysis

751 In the current experiment, we introduced a control condition that was physically identical
752 to our temporal prediction tasks (until reappearance in the tactile condition) in order to account
753 for processes that are not directly related temporal predictions. Hence, for most of our
754 statistical analyses, we were interested in comparing the two temporal prediction tasks with
755 the working memory control task, respectively, and not in comparing the two temporal
756 prediction tasks with each other. Therefore, instead of computing an analysis of variance
757 across all three conditions, we directly computed two separate *t*-tests for the comparison of
758 the visual or the tactile temporal prediction with the working memory task, respectively, and
759 accounted for multiple comparisons by adjusting the alpha level.

760 *Psychometric curve*

761 We did not provide participants with feedback about the correctness of their response.
762 Hence, participants responded within their individual framework of a “subjectively correct”
763 reappearance timing or a “subjectively equal” luminance of the stimulus, respectively. To
764 obtain these subjective points of “right-on-time” (ROT) in the temporal prediction tasks or the
765 “points of subjective equality” (PSE) in the working memory task, we fitted a psychometric
766 curve to the behavioral data of each participant from all trials in each condition. First, for each
767 timing difference or luminance difference, respectively, we computed the proportion of “too
768 late” or “brighter” answers for each participant. Then, we fitted a binomial logistic regression
769 (psychometric curve) using the glmfit.m and gmlval.m functions provided in MATLAB. The
770 fitted timing or luminance difference value at 50% proportion “too late” or “brighter” answers

771 was determined as ROT or PSE for each participant, respectively. To test for a significant bias
772 towards one of the answers, we tested the ROT or PSE from all participants against zero using
773 one-sample *t*-tests (α = .05 / 3 = .017). The steepness of the psychometric function was
774 computed as the reciprocal of the difference between fitted timing or luminance difference
775 values at 75% and 25% proportion "too late" or "brighter" answers, respectively.

776 *Linear model*

777 To test whether reaction times were dependent on the timing difference of the reappearing
778 stimulus in each task, we averaged across all luminance differences within each timing
779 difference bin in each condition. We then utilized a second-order (quadratic) polynomial
780 regression model with timing difference as predictor for reaction times and computed the first-
781 and second-order coefficients for each participant in each condition. The coefficients from all
782 participants were then tested against zero using one-sample *t*-tests in all conditions (α = .05 /
783 3 = .017).

784 *Spectral power*

785 We decomposed the MEG recordings into time-frequency representations by convolving
786 the data with complex Morlet's wavelets (Cohen, 2014). The recording of each trial and
787 channel was convoluted with 40 complex wavelets, logarithmically spaced between 0.5 to 100
788 Hz. With increasing frequency, the number of cycles for each wavelet logarithmically increased
789 from two to ten cycles. For all analyses of the MEG data, we considered subjectively correct
790 trials only, i.e., trials in which participants answered correctly based on their individual ROT.
791 To correct for trial count differences between the tasks, we stratified the number of trials for
792 each participant for the three different conditions by randomly selecting as many trials for each
793 condition as the number available from the condition with lowest trial count.

794 Since the temporal dependencies between the movement onset, disappearance behind
795 the occluder and reappearance of the stimulus varied strongly between trials, averaging across

trials would heavily smear the power estimates of the different stages within each trial. To obtain an estimate of spectral power modulations related to the different events in our experimental paradigm, we cut each trial further into four separate, partly overlapping windows (see Figure 2A): a “Baseline” window from -550 to -50 ms earlier to movement onset; a “Movement” window from -50 to 950 ms relative to the movement onset; a “Disappearance” window from -350 to 950 ms relative to complete disappearance of the stimulus behind the occluder; and a “Reappearance” window from -350 to 450 ms relative to the (first frame) reappearance of the stimulus. Spectral power estimates were then averaged across all trials belonging to the same condition in each window and binned into time windows 100 ms (centered on each full deci-second). All power estimates were normalized using the pre-stimulus baseline window from -500 to -200 ms earlier to movement onset.

For all statistical analyses on sensor level, we first flipped all sensors of participants, who saw the stimulus moving from right to left, at the sagittal midline, i.e., the anterior-posterior axis. This made sure that lateralized activity due to the lateralized stimulation was comparable across groups. From this on, we considered all participants as if for everyone the stimulus was moving from the left to the right side. Channels that did not have a counterpart on the opposite site were excluded from further analyses. In order to obtain an overview of the spectral power modulations related to the different events within the trials, we then averaged the power estimates across all channels and conditions (grand average) and tested each time-frequency pair of the Movement, Disappearance and Reappearance windows against the pre-stimulus baseline using paired-sample *t*-tests. We controlled for multiple comparisons by employing cluster-based permutation statistics as implemented in FieldTrip (Maris and Oostenveld, 2007). In this procedure, neighboring time-frequency bins with an uncorrected *p*-value below 0.05 are combined into clusters, for which the sum of t-values is computed. A null-distribution is created through permutations of data across participants (n = 1,000 permutations), which defines the maximum cluster-level test statistics and corrected p-values for each cluster. For each window, a separate cluster-permutation test was performed (α = .05; liberally chosen to observe all ongoing power modulations; see Results section).

824 Since we were most interested in differences between the conditions during the
825 disappearance time, we subsequently compared the spectral power estimates averaged within
826 the beta range (13–30 Hz; see Results section) at each time point within the disappearance
827 window and all channels from the visual or tactile temporal prediction task with the working
828 memory task. We again employed cluster-permutation statistics, this time by clustering
829 neighboring channels and time points. We used a one-sided α = .025 / 2 = .0125, since
830 negative and positive clusters were tested separately, and to adjust for the two separate
831 comparisons between the conditions (used throughout the study unless stated differently).

832 To estimate spectral power in source space, we computed separate leadfields for each
833 recording session and participant based on each participant's mean head position in each
834 session and individual magnetic resonance images. We used the single-shell volume
835 conductor model (Nolte, 2003) with a 5,003 voxel grid that was aligned to the MNI152 template
836 brain (Montreal Neurological Institute, MNI; http://www.mni.mcgill.ca) as implemented in the
837 METH toolbox. Cross-spectral density (CSD) matrices were computed from the complex
838 wavelet convoluted data in steps of 100 ms in the same time windows as outlined above. To
839 avoid biases in source projection, common adaptive linear spatial filters (DICS beamformer
840 (Gross et al., 2001)) pointing into the direction of maximal variance were computed from CSD
841 matrices averaged across all time bins and conditions for each frequency.

842 All time-frequency resolved CSD matrices were then multiplied with the spatial filters to
843 estimate spectral power in each of the 5,003 voxels and normalized with the pre-stimulus
844 baseline window. Analogous to sensor space, we first flipped all voxels at the y-axis (anterior-
845 posterior axis) for the half of participants that saw the stimulus moving from right to left earlier
846 to further statistical analysis. We then averaged across all time bins within the disappearance
847 window and utilized cluster-based permutation statistics to identify clusters of voxels that show
848 statistical difference in beta power between each of the temporal prediction tasks and the
849 working memory task.

Inter-trial phase consistency

We computed ITPC estimates from the complex time-frequency representations obtained from the wavelet convolution as described in the *Spectral power* section above. In each time sample and trial, the phase of the complex data was extracted (using the function angle.m in MATLAB). ITPC was then computed across all subjectively correct and stratified trials within each of the four time windows in all frequencies as

$$ITPC_{tf} = \left| n^{-1} \sum_{r=1}^{n} e^{ik_{tfr}} \right|$$

where *n* is the number of trials and *k* the phase angle in trial *r* at time-frequency point *tf* (Cohen, 2014). In other words, ITPC is the length of the mean vector from all phase vectors with length 1 across all trials at a given time-frequency point. Values for ITPC can vary between 0 and 1, where 0 means that at a given time-frequency point there is no phase consistency across trials at all and 1 means all trials show the exact same phase. Similar to spectral power, we averaged ITPC estimates again in bins of 100 ms and plotted all time windows averaged across all channels and conditions to obtain a general overview of ITPC estimates at all events during the trial.

Since we were most interested in ITPC related to stimulus disappearance behind the occluder, we subsequently computed ITPC in a longer time window from -1,900 ms to 1,900 ms centered around time of complete stimulus disappearance behind the occluder. Thus, we took advantage of the fact that the onset of other events within each trial, such as the movement onset and the reappearance of the stimulus, strongly jittered across all trials and strong contributions of these events to ITPC could thereby be reduced (see Figure S4). For statistical analysis, we first averaged ITPC estimates within a frequency band of 0.5 to 3 Hz (see Results) and then computed cluster-based permutation statistics across all 100 ms time bins within the 3,800 ms long window and all sensors between each of the temporal prediction tasks and the working memory task.

875 ITPC on source level was computed using the same leadfields and common beamformer
876 filters as for spectral power (see above). The complex time-frequency representations obtained
877 from the wavelet convolution within the 3,800 ms long window on sensor level were multiplied
878 with the filters to obtain the time-frequency representations in each of the 5,003 voxels. ITPC
879 was computed for each time sample and frequency and then averaged within the time window
880 showing statistically significant difference between the temporal prediction tasks and the
881 working memory task on sensor level and within the pre-defined frequency band of 0.5 to 3
882 Hz. Cluster-based permutation statistics were employed to identify clusters of voxels showing
883 statistically significant differences in ITPC between the conditions on source level.

884 Correlations between condition-wise source level ITPC estimates and the steepness of
885 each individual's psychometric function were computed using Pearson correlations in each of
886 the 5,003 voxels within the grid. For this analysis, we averaged ITPC estimates from time bins
887 of 0 to 1,500 ms with respect to the disappearance of the stimulus within the pre-defined delta
888 band of 0.5 to 3 Hz. Multiple comparisons were accounted for by using cluster-based
889 permutation statistics as implemented in FieldTrip (α = .025 / 3 = .008)

890 *Delta phase clustering at ROT*

891 To determine whether each participant's subjective ROT was associated with a specific
892 phase in the delta band, we extracted the phase at each individual's ROT from sensors
893 showing the strongest ITPC effect and computed the distance from this distribution to a uniform
894 distribution over all possible phases. The procedure was as follows.

895 For this analysis, we only considered trials in which the stimulus reappeared later than
896 each individual's ROT and the participant answered subjectively correct. By this, we prevented
897 possible phase distortions by the external stimulation earlier to or at ROT. Trials were again
898 stratified across conditions. Moreover, to make sure that we reduced also activity that was
899 related to external stimulations after each individual's ROT, we first aligned all trials from the
900 same condition to the time point of stimulus reappearance, computed the average across trials

901 (event-related field, ERF) and subtracted the ERF caused by the reappearance from all trials
902 in that condition. Subsequently, in each trial we centered a 2,500 ms long window on each
903 participant's ROT, computed a complex wavelet convolution for all frequencies between 0.5
904 and 3 Hz (14 frequencies; same procedure and frequencies as above) in all channels, and
905 computed the mean phase angle at ROT, i.e., the center time bin, across all considered trials
906 in each condition. This procedure is similar to computing ITPC as described above, except for
907 extracting the angle of the mean phase vector instead of the length. Since for the working
908 memory task we did not have an estimate of each individual's ROT, we applied the estimate
909 of ROT from the visual prediction task also to the working memory trials, since based on their
910 equal physical appearance temporal predictions should also be equal.

911 As a next step, from the result of the cluster-based permutation statistics on ITPC
912 estimates described above, we determined the sensors that showed the strongest ITPC effect
913 for the two contrasts between the temporal prediction tasks and the working memory task for
914 a time window between 0 and 1,500 ms after disappearance behind the occluder. For the
915 contrast between the visual prediction and the working memory task, we considered the
916 sensors showing the top 20% of *t*-values (37 channels). To keep the number of sensors
917 comparable, we also considered the top 37 sensors from the contrast of the tactile prediction
918 task against working memory.

919 Within these channels, for each individual participant we determined the frequency within
920 the 0.5 to 3 Hz delta band, which showed the strongest ITPC for the visual or the tactile
921 prediction as compared to the working memory task, respectively, in the same time window of
922 0 to 1.500 ms. To determine the frequencies also for the working memory condition, we here
923 extracted the frequencies showing the strongest estimates of ITPC in the working memory as
924 compared to the visual temporal prediction task. For these individual frequencies, we plotted
925 the phase angle at ROT (as described above) from all the considered channels and all
926 participants in a histogram (in bins of 10°; see Figure 5). If there was no relationship between
927 each individual's phase at each individual's ROT, this histogram would show a uniform
928 distribution across all possible phases (-180° – 180°). Therefore, we computed the distance

929 from the observed phase distribution to a uniform distribution using a discrete and normalized
930 version of the Kullback-Leibler distance, i.e., the modulation index (MI) (Tort et al., 2010).

931 For statistical analysis, we repeated the same procedure as described above for 10,000
932 times and randomly picked any frequency from the 14 frequencies within the 0.5 to 3 Hz band
933 in each repetition. By that we obtained a distribution of surrogate MI estimates (but still based
934 on real data from all individual participants), from which we computed the percentile
935 determined by the MI that was observed using the individually strongest ITPC frequency. MI
936 estimates above the 95th percentile were considered significantly stronger as compared to the
937 randomly obtained surrogate MIs (*p*-value = 1 – percentile).

Acknowledgements

We thank Florian Göschl, Tessa Rusch, Marina Fiene and Guido Nolte for valuable discussions. This work was funded by grants from the DFG (SFB TRR 169/B1 and SFB 936/A3 to A.K.E.).

Author Contributions

Conceptualization, J.D., A.K.E, P.W., A.M.; Methodology, J.D., A.K.E.; Software, J.D.; Formal Analysis, J.D.; Investigation, J.D.; Writing – Original Draft, J.D.; Writing – Review & Editing, A.K.E., P.W., A.M., D.Z.; Visualization, J.D.; Funding Acquisition, A.K.E.; Supervision, A.K.E.; Project Administration, A.K.E., D.Z.; Resources, A.K.E.

Declaration of Interests

The authors declare no competing interests.

References

Arnal, L.H., and Giraud, A.L. (2012). Cortical oscillations and sensory predictions. Trends Cogn. Sci. *16*, 390–398.

Arnal, L.H., Doelling, K.B., and Poeppel, D. (2015). Delta-beta coupled oscillations underlie temporal prediction accuracy. Cereb. Cortex *25*, 3077–3085.

Besle, J., Schevon, C.A., Mehta, A.D., Lakatos, P., Goodman, R.R., McKhann, G.M., Emerson, R.G., and Schroeder, C.E. (2011). Tuning of the human neocortex to the temporal dynamics of attended events. J. Neurosci. *31*, 3176–3185.

Brainard, D.H. (1997). The psychophysics toolbox. Spat. Vis. *10*, 433–436.

Breska, A., and Ivry, R.B. (2016). Taxonomies of timing: where does the cerebellum fit in? Curr. Opin. Behav. Sci. *8*, 282–288.

Buzsáki, G. (2006). Rhythms of the brain (New York: Oxford University Press.).

Cohen, M. (2014). Analyzing neural time series data: theory and practice (Cambridge: MIT Press).

Coull, J.T., and Nobre, A.C. (2008). Dissociating explicit timing from temporal expectation with fmri. Curr. Opin. Neurobiol. *18*, 137–144.

Cravo, A.M., Rohenkohl, G., Wyart, V., and Nobre, A.C. (2013). Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. J. Neurosci. *33*, 4002–4010.

Daume, J., Gruber, T., Engel, A.K., and Friese, U. (2017a). Phase-amplitude coupling and long-range phase synchronization reveal frontotemporal interactions during visual working memory. J. Neurosci. *37*, 313–322.

Daume, J., Graetz, S., Gruber, T., Engel, A.K., and Friese, U. (2017b). Cognitive control during audiovisual working memory engages frontotemporal theta-band interactions. Sci. Rep. *7*, 12585.

Doelling, K.B., and Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. Proc. Natl. Acad. Sci. *112*, E6233–E6242.

Doelling, K.B., Assaneo, M.F., Bevilacqua, D., Pesaran, B., and Poeppel, D. (2019). An oscillator model better predicts cortical entrainment to music. Proc. Natl. Acad. Sci. 201816414.

van Ede, F., de Lange, F., Jensen, O., and Maris, E. (2011). Orienting attention to an upcoming tactile event involves a spatially and temporally specific modulation of sensorimotor alpha- and beta-band oscillations. J. Neurosci. *31*, 2016–2024.

Engel, A.K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. Nat. Rev. Neurosci. *2*, 704–716.

Engel, A.K., Gerloff, C., Hilgetag, C.C., and Nolte, G. (2013). Intrinsic coupling modes: multiscale interactions in ongoing brain activity. Neuron *80*, 867–886.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn. Sci. *9*, 474–480.

Fujioka, T., Trainor, L.J., Large, E.W., and Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. J. Neurosci. *32*, 1791–1802.

Giraud, A.-L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. Nat. Neurosci. *15*, 511–517.

Gomez-Ramirez, M., Kelly, S.P., Molholm, S., Sehatpour, P., Schwartz, T.H., and Foxe, J.J. (2011). Oscillatory sensory selection mechanisms during intersensory attention to rhythmic auditory and visual inputs: a human electrocorticographic investigation. J. Neurosci. *31*, 18556–18567.

Gould, I.C., Rushworth, M.F., and Nobre, A.C. (2011). Indexing the graded allocation of visuospatial attention using anticipatory alpha oscillations. J. Neurophysiol. *105*, 1318–1326.

Gross, J., Kujala, J., Hamalainen, M., Timmermann, L., Schnitzler, A., and Salmelin, R. (2001). Dynamic imaging of coherent sources: studying neural interactions in the human brain. Proc. Natl. Acad. Sci. *98*, 694–699.

Gulberti, A., Moll, C.K.E., Hamel, W., Buhmann, C., Koeppen, J.A., Boelmans, K., Zittel, S., Gerloff, C., Westphal, M., Schneider, T.R., et al. (2015). Predictive timing functions of cortical beta oscillations are impaired in Parkinson's disease and influenced by L-DOPA and deep brain stimulation of the subthalamic nucleus impaired beta-band timing functions in PD. NeuroImage Clin. *9*, 436–449.

Henry, M.J., Herrmann, B., and Obleser, J. (2014). Entrained neural oscillations in multiple frequency bands comodulate behavior. Proc. Natl. Acad. Sci. *111*, 14935–14940.

Herrmann, C.S. (2001). Human EEG responses to 1-100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. Exp. Brain Res. *137*, 346–353.

Hipp, J.F., and Siegel, M. (2013). Dissociating neuronal gamma-band activity from cranial and ocular muscle activity in EEG. Front. Hum. Neurosci. *7*, 338.

Ivry, R.B., and Keele, S.W. (1989). Timing functions of the cerebellum. J. Cogn. Neurosci. *1*, 136–152.

Kösem, A., Bosker, H.R., Takashima, A., Meyer, A., Jensen, O., and Hagoort, P. (2018). Neural entrainment determines the words we hear. Curr. Biol. *28*, 2867-2875.e3.

Lakatos, P., Chen, C.-M., O'Connell, M.N., Mills, A., and Schroeder, C.E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. Neuron *53*, 279–292.

Lakatos, P., Karmos, G., Mehta, A.D., Ulbert, I., and Schroeder, C.E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. Science *320*, 110–113.

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

Mercier, M.R., Foxe, J.J., Fiebelkorn, I.C., Butler, J.S., Schwartz, T.H., and Molholm, S. (2013). Auditory-driven phase reset in visual cortex: human electrocorticography reveals mechanisms of early multisensory integration. Neuroimage *79*, 19–29.

Nolte, G. (2003). The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. Phys. Med. Biol. *48*, 3637–3652.

ten Oever, S., and Sack, A.T. (2015). Oscillatory phase shapes syllable perception. Proc. Natl. Acad. Sci. *112*, 15833–15837.

ten Oever, S., Schroeder, C.E., Poeppel, D., van Atteveldt, N., Mehta, A.D., Mégevand, P., Groppe, D.M., and Zion-Golumbic, E. (2017). Low-frequency cortical oscillations entrain to subthreshold rhythmic auditory stimuli. J. Neurosci. *37*, 4903–4912.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. *2011*, 1–9.

R Core Team (2014). R: a language and environment for statistical computing. R Found. Stat. Comput. Vienna Austria.

Rimmele, J.M., Morillon, B., Poeppel, D., and Arnal, L.H. (2018). Proactive sensing of periodic and aperiodic auditory patterns. Trends Cogn. Sci. *22*, 870–882.

Rohenkohl, G., and Nobre, A.C. (2011). Alpha oscillations related to anticipatory attention follow temporal expectations. J. Neurosci. *31*, 14076–14084.

Roth, M.J., Synofzik, M., and Lindner, A. (2013). The cerebellum optimizes perceptual predictions about external sensory events. Curr. Biol. *23*, 930–935.

Saleh, M., Reimer, J., Penn, R., Ojakangas, C.L., and Hatsopoulos, N.G. (2010). Fast and slow oscillations in human primary motor cortex predict oncoming behaviorally relevant cues. Neuron *65*, 461–471.

Samaha, J., Bauer, P., Cimaroli, S., and Postle, B.R. (2015). Top-down control of the phase of alpha-band oscillations as a mechanism for temporal prediction. Proc. Natl. Acad. Sci. *112*, 8439–8444.

Schroeder, C.E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci. *32*, 9–18.

Stefanics, G., Hangya, B., Hernádi, I., Winkler, I., Lakatos, P., Ulbert, I., Hernadi, I., Winkler, I., Lakatos, P., and Ulbert, I. (2010). Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. J. Neurosci. *30*, 13578–13585.

Stolk, A., Todorovic, A., Schoffelen, J.-M., and Oostenveld, R. (2013). Online and offline tools for head movement compensation in MEG. Neuroimage *68*, 39–48.

Tort, A.B.L., Komorowski, R., Eichenbaum, H., and Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. J. Neurophysiol. *104*, 1195–1210.

VanRullen, R. (2016). Perceptual cycles. Trends Cogn. Sci. *20*, 723–735.

Wiener, M., Turkeltaub, P., and Coslett, H.B. (2010). The image of time: a voxel-wise meta-analysis. Neuroimage *49*, 1728–1740.

Wilsch, A., Henry, M.J., Herrmann, B., Maess, B., and Obleser, J. (2015). Slow-delta phase concentration marks improved temporal expectations based on the passage of time. Psychophysiology *52*, 910–918.

Zoefel, B., ten Oever, S., and Sack, A.T. (2018). The involvement of endogenous neural oscillations in the processing of rhythmic input: more than a regular repetition of evoked neural responses. Front. Neurosci. *12*, 1–13.