Delta phase resets mediate 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 upcoming 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 luminance matching 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.

Keywords

Temporal prediction; crossmodal prediction; neural oscillations; delta band; beta band; inter-trial phase coherence; spectral power; phase reset; cerebellum; magnetoencephalography

Introduction

Neural oscillations reflect alternating states of higher or lower neural excitability, modulating the efficiency by which coupled neurons engage in mutual interactions [1]. As a result, neural communication and information processing has been shown to occur in a phase-dependent manner [2,3], reflected for example by fluctuations in perception thresholds correlating with the phase of ongoing oscillations [4]. Based on these assumptions, oscillations were also linked to temporal predictions of upcoming relevant information [2,5,6]. Studies have shown that animals can utilize predictive aspects of environmental stimuli in a way that reaction times are reduced [7–10] or stimulus processing is enhanced [11,12]. 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 stimulation in order to optimize behavior [13].

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 [14,15]. Also in the visual domain, studies showed that neural oscillations align to rhythmic visual input [8,11,16,17]. However, whether temporal predictions indeed involve phase resets of endogenous neural oscillations remains a matter of debate [18–20]. Despite their ecological relevance, using rhythms for the investigation of an involvement of oscillations in temporal predictions entails methodological and conceptual challenges. Rhythmic input 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 [21]. Rather than phase resets of endogenous neural oscillations, temporal predictions could therefore also be reflected by stimulus-induced potentials that appear to be rhythmic during rhythmic stimulation [18]. Conclusive evidence that temporal predictions involve phase resets of endogenous neural oscillations rather than stimulus evoked potentials is still lacking.

Moreover, using only rhythmic stimulation excludes 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.

For this reason, we set out to investigate whether phase adjustments of neural oscillations can be observed for non-rhythmic, but predictable visual motion stimuli. We measured magnetoencephalography (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, participants were required to temporally predict the correct time point of reappearance to be able to accomplish the task. Participants further performed a control task, in which the task was to judge the luminance of the reappearing stimulus instead of its timing. Importantly, physical appearance of both conditions was exactly the same in all aspects of the stimulation. Any purely stimulus-related, bottom-up activity should therefore 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 [22,23], we further introduced a third condition, in which a tactile instead of a visual stimulus was presented at reappearance. By contrasting it to the luminance matching 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.

In the two temporal prediction tasks, as compared to luminance matching, we observed stronger delta band inter-trial phase consistency (ITPC) within time windows between disappearance and expected reappearance in frontoparietal brain areas. Enhanced delta ITPC specifically in these time windows reflected phase resets of ongoing oscillations at disappearance of the stimulus, where temporal prediction might be initialized. By introducing a novel design, in which physical stimulation was exactly the same between the visual temporal prediction and the luminance matching task, we provide profound evidence that purely bottom-up evoked processes could not explain observed differences in ITPC between the condition. In the crossmodal setting, we show that temporal information provided to the visual modality leads to phase adjustments also in the tactile modality. Moreover, participants who showed a consistent judgment of reappearance timing, as represented by a steep psychometric function, also showed stronger delta ITPC during temporal predictions. This confirms that a consistent timing judgment across trials also involves a consistent phase across trials. We further observed a phase clustering at ±90° within the delta oscillation showing the strongest ITPC in each participant at the individual subjective time points of predicted reappearance. This strongly suggests that the phase of ongoing oscillations serves as a subjective marker for the individual estimation of elapsed time.

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 conditions and "point of subjective equivalence" (PSE) in the luminance matching 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 luminance matching task (ΔRGB (PSE) = -1.29 ± 4.54 RGB; $t(22) = -1.36$; $p = .19$), respectively (Fig. 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 luminance matching 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 (Fig. 1C), we averaged across all luminance differences and fitted a linear model to reaction time data in each condition. 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 luminance matching 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 luminance matching 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.

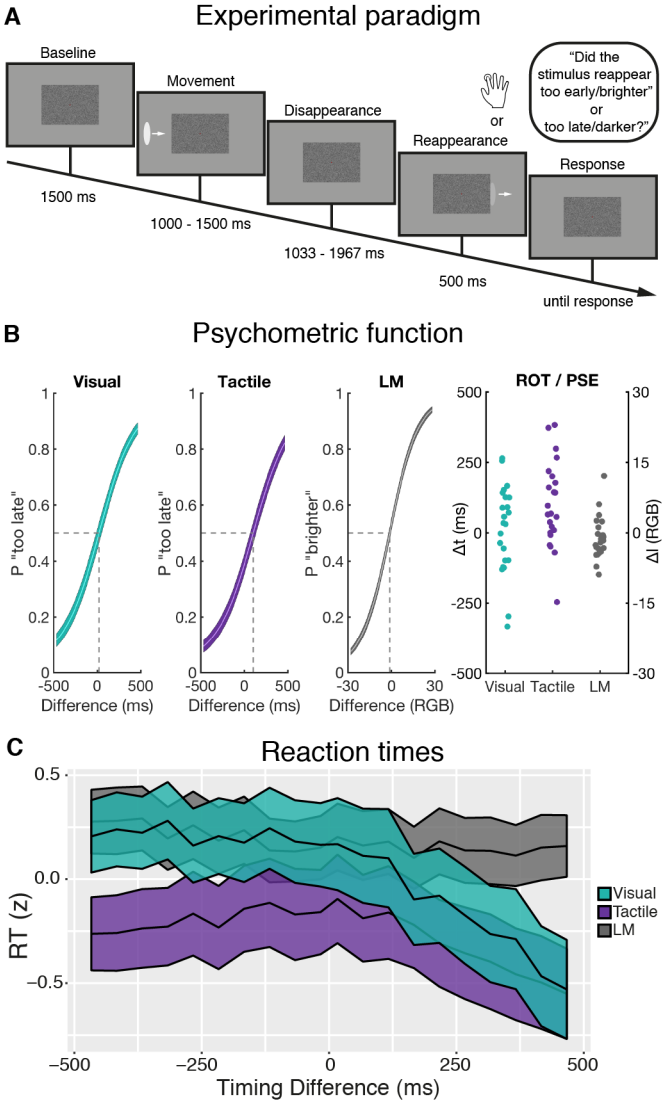

Figure 1. Experimental design and behavioral results. (**A**) A stimulus moved towards the center of the screen until it disappeared behind an occluder. The task was to judge whether the stimulus reappeared *too early* or *too late*. In the luminance matching condition, task was to judge whether the luminance became *brighter* or *darker*. 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 contralateral to the disappearance of the visual stimulus. (**B**) Psychometric functions 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). Colored areas depict standard errors of the mean (SEM). (**C**) Log-transformed and standardized reaction times for all timing differences (mean ± SEM). P = proportion; LM = luminance matching; 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, and tested an average of spectral power across all sensors and conditions against a pre-stimulus baseline window. As a first step, we obtained a general overview of power modulations at each event in the experimental paradigm. Due to the jittered stimulation built into the design (see Materials and Methods), we computed 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"; Fig. 2A).

In time bins around movement onset as well as reappearance (but not disappearance) 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 and gamma range (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 Fig. S1).

Since we were most interested in examining power modulations associated with temporal predictions, i.e., during the disappearance window, we further compared spectral power estimates between the temporal prediction tasks and the luminance matching task in all sensors within the disappearance window while ignoring the other windows. We restricted our analysis to the classical beta band ranging from 13 to 30 Hz, showing the strongest modulation as compared to baseline during the disappearance window. Cluster-based permutation statistics revealed reduced beta power during visual temporal prediction in occipital sensors during all time-bins of the disappearance window (cluster-p = .01). Source level statistics revealed a statistically significant decrease of beta power in a cluster of bilateral occipital voxels (cluster-p = .01). Beta power was further reduced during tactile prediction in a cluster of occipital as well as left lateralized frontocentral sensors (cluster-p = .002). At source level, a significant power reduction in the beta band was most strongly apparent in parts of bilateral visual as well as left-lateralized somatosensory cortex (cluster-p = .01).

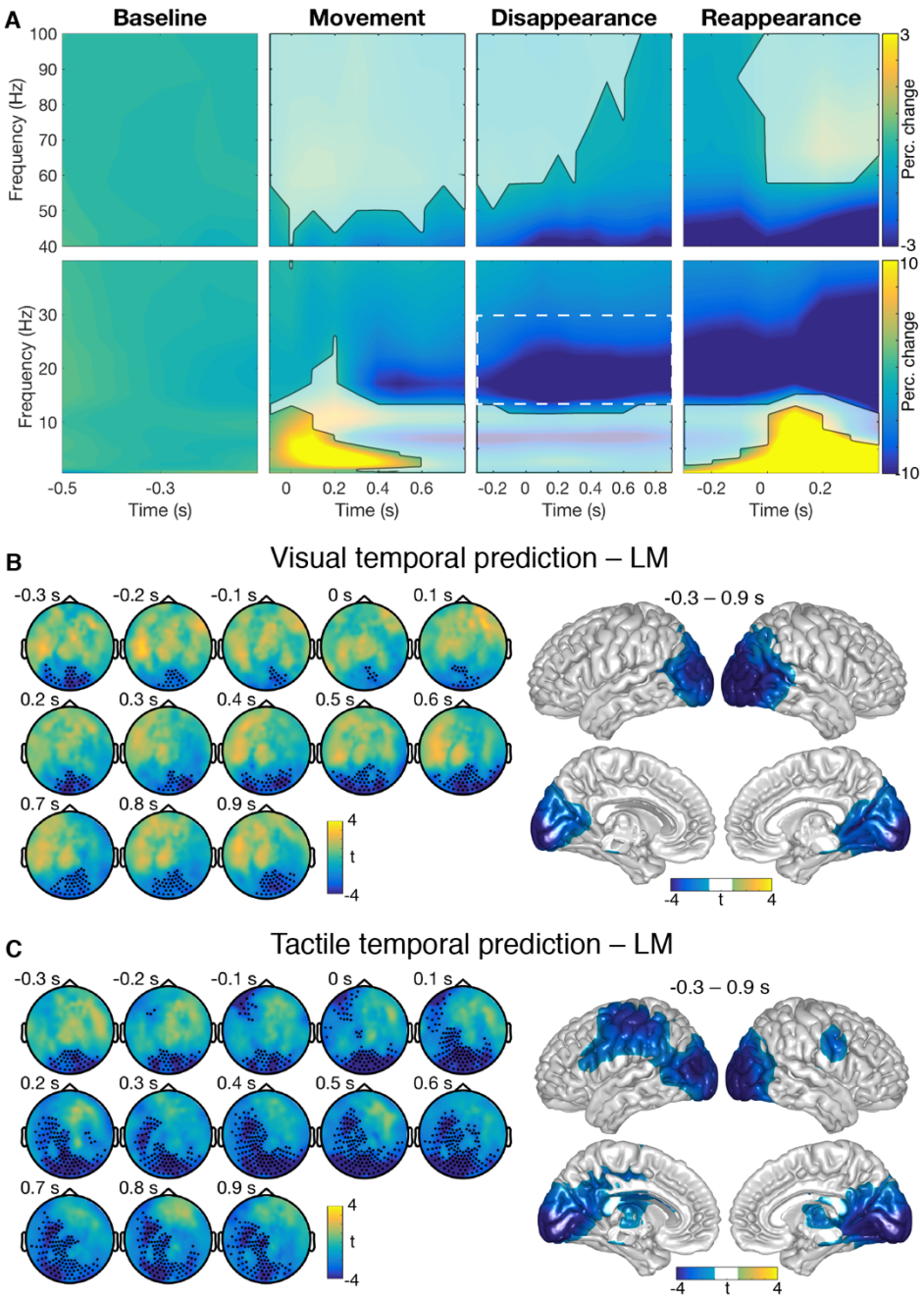

Figure 2. Power modulations during temporal prediction. (**A**) Spectral power averaged across sensors, conditions, and participants. Each window was centered on the different events within the paradigm and normalized with pre-stimulus baseline. Time 0 refers to the onset of each event. Cluster-based permutation statistics revealed significant power modulations as compared to baseline (unmasked colors). See also Fig. S1. (**B,C**) Difference between the two temporal prediction and the luminance matching task, respectively, within the beta band (13 – 30 Hz) in time bins around stimulus disappearance. At source level, cluster-based permutation statistics revealed cluster of voxels showing significant differences between the conditions (colored voxels). See also Fig. S2. LM = luminance matching.

Inter-trial phase consistency between conditions

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

Hence, the delta band showed no increase in power but the strongest increase in ITPC as compared to baseline during the disappearance window for an average across all conditions (see Fig. 2A, 3A, and S1). For further statistical comparisons between conditions, we therefore restricted our analyses to an average of frequencies between 0.5 to 3 Hz. 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 (see Fig. S3; for condition-specific ITPC differences during disappearance to baseline, see Fig. S4).

We found two clusters that showed significantly stronger ITPC during visual temporal predictions as compared to luminance matching (Fig. 3B). 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).

ITPC was also significantly enhanced in bilateral temporal sensors during tactile temporal predictions, evolving around -400 ms in right temporal sensors and shifting towards left hemisphere with ongoing disappearance time (cluster p < .001; Fig. 3C). 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 luminance matching (cluster p < .001).

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. Eye movements as well as ITPC computed from the ET data did not show any differences between the conditions (see Fig. S5A,B,C). Moreover, we did not observe significant correlations between ITPC values computed from the ET and the MEG signal in any of the conditions across participants (see Fig. S5D).

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 luminance matching between 0 and 1,500 ms (see Fig. 3B; similar results were obtained for sensors showing the top 10% or 5% of t values, see Fig. S3D). ITPC also increased in the luminance matching 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 stayed elevated throughout the entire disappearance window.

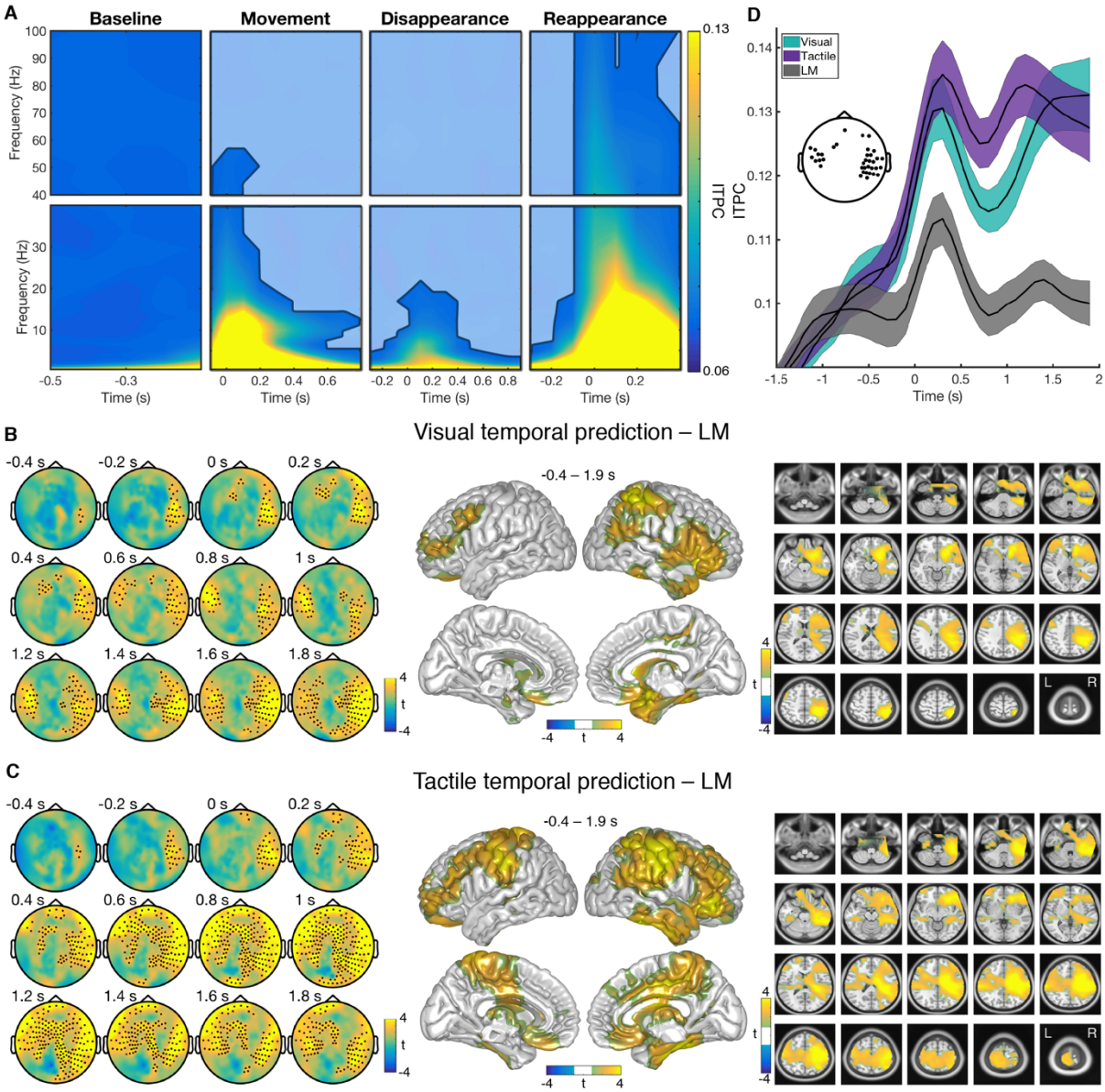

Figure 3. ITPC during temporal prediction as compared to luminance matching. (**A**) ITPC estimates averaged across sensors, conditions, and participants. Masked colors refer to non-significant ITPC modulations as compared to baseline. (**B,C**) Difference in ITPC between the visual or tactile prediction and the luminance matching task, respectively, within the delta band. For clarity, only every second time bin was plotted. On source level, clusters of voxels showing significant differences between the conditions are colored. See also Fig. S3, 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 luminance matching task (see topography). LM = luminance matching.

Correlation of ITPC to 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 also exhibit stronger ITPC during temporal predictions, 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 – the steeper the psychometric function, the more consistent the answers of the participant. We computed Pearson correlations of source level delta ITPC with the steepness of the psychometric function across participants and found statistically significant positive correlations in the visual (cluster $p = .003$) as well as in the tactile temporal prediction task (cluster $p = .002$; Fig. 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 luminance matching task (all cluster $p > .1$).

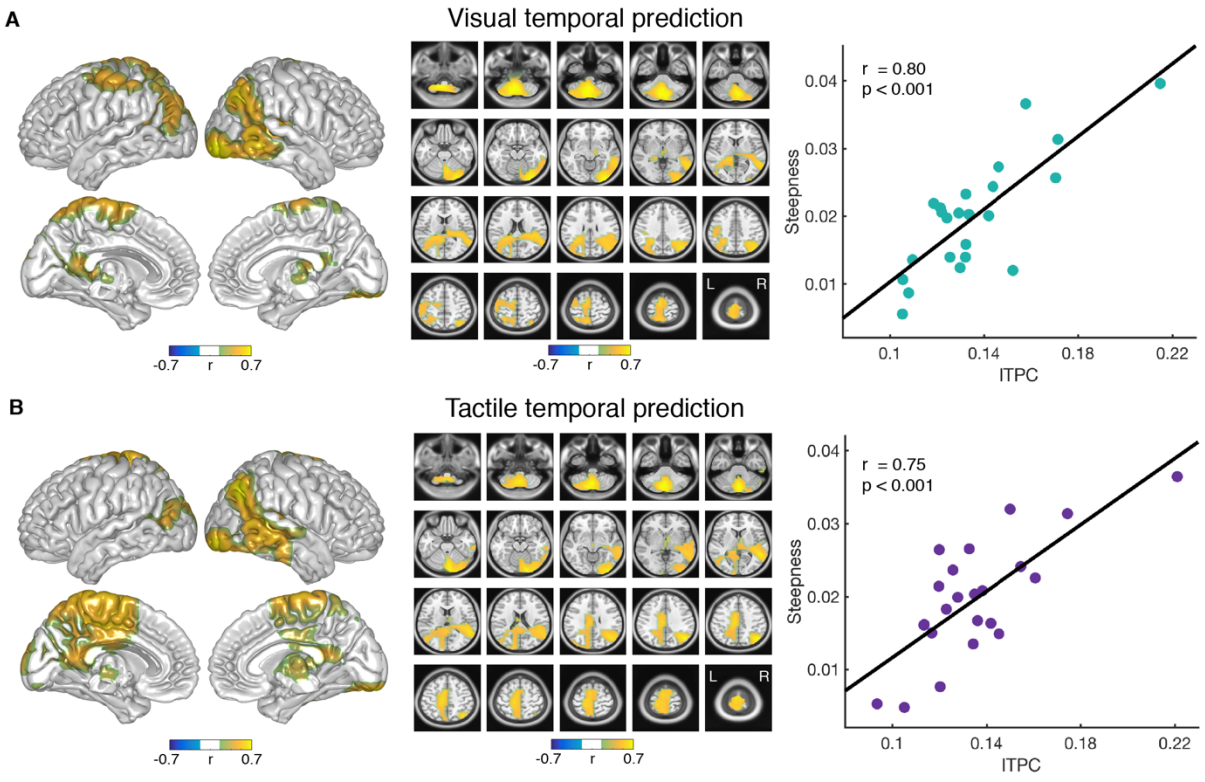

Figure 4. Correlation of ITPC to behavior. (A,B) Correlation of individual ITPC estimates with the individual steepness of the psychometric function within all voxels, shown in (A) for the visual prediction, and in (B) for the tactile prediction condition. ITPC estimates were averaged within the delta 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. In the scatter plots, each dot represents one participant and ITPC estimates were averaged across all voxels within the clusters of significant correlations. There was no significant correlation observed for the luminance matching condition.

Delta phase clustering at individually predicted reappearance time points

Furthermore, if the phase of oscillations indeed codes for the predicted time point of reappearance, a clustering of a specific phase should be observed, when extracting the phase at each individual ROT, i.e., the time point of each individual's estimation for the correct reappearance of the stimulus. Such a clustering at subjective ROT estimates would provide strong evidence that the phase of ongoing oscillations codes for the subjective estimation of elapsed time. That is, 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, since individual ROTs strongly differed across participants as well (see Fig. 1B).

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 luminance matching task at ROT in each participant. We again used the sensors that showed the strongest statistical differences in ITPC for the contrasts of each prediction task to the luminance matching (see Fig. 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 combined and plotted into a histogram for each condition (Fig. 5, upper row; each plot shows participants x channel data). We quantified the distance of the observed distribution to a uniform distribution by means of the modulation index [MI; 24].

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 luminance matching task, no significant clustering at a specific phase was found (p = .96).

Our reaction time analysis revealed that also in the luminance matching 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 luminance matching task, since physical stimulation was identical in both tasks. We repeated the above described analysis for the luminance matching 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 (Fig. 5, Panel 4: LM (VP)). With frequencies obtained from the visual prediction task, the MI observed for the luminance matching task was significantly stronger than MIs obtained from the random repetitions (p = .02).

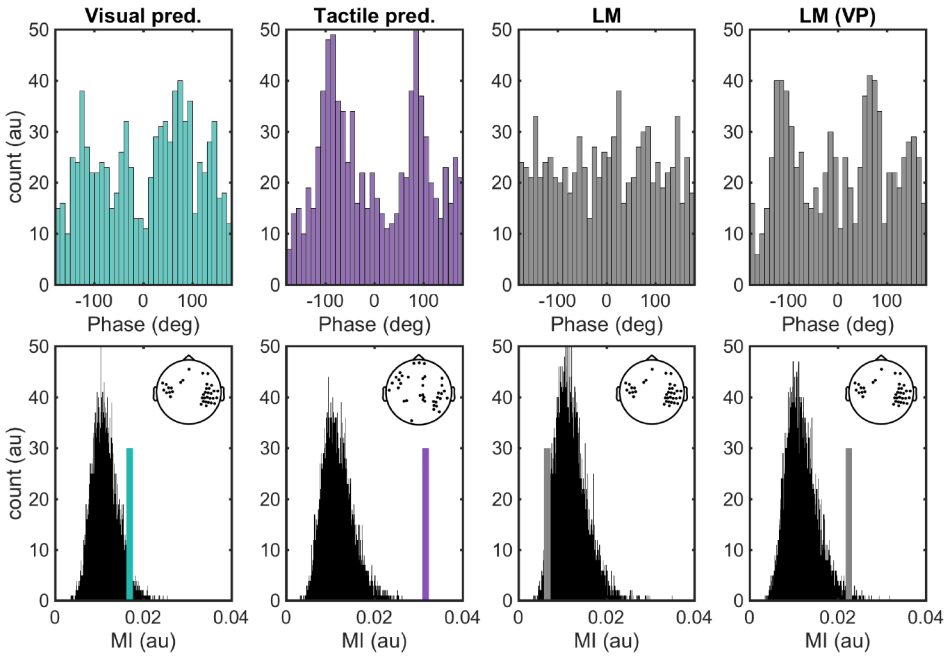

Figure 5. Delta band phase clustering at individual ROT. In each condition, the mean phase observed at individual ROT for each participant was extracted from the top 20% of channels (see Fig. 2 and topographies) and from the delta frequency showing the strongest differences in ITPC to luminance matching. 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). LM = luminance matching; LM (VP) = data from luminance matching condition with frequencies determined in the visual prediction condition (see main text); MI = modulation index.

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 [5]. 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. On the one hand, studies found that beta power was even increased shortly before the onset of the expected stimulus in auditory [25] and visual rhythmic stimulation [16]. On the other hand, van Ede et al. [26] 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. [26]. 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 luminance matching 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.

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 Fig. S2). 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.

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

337 Studies found that neural oscillations entrain towards rhythmic sensory input to track the
338 low-frequency temporal regularities of the stimulation, especially in the auditory domain [14].
339 Such phase entrainment does not only occur in the delta band but can flexibly adapt to the
340 frequency of the external input also at higher frequencies such as the theta or the alpha band
341 during auditory stimulation [15]. However, in the visual system, evidence for the tracking of
342 temporally predictive input by neural oscillations is not as clear. On the one hand, studies
343 showed that the phase of neural oscillations is involved in temporal predictions of low-
344 frequency visual input [11,12,16]. On the other hand, studies suggested that temporal predictions
345 in the visual system were specific to the alpha band, although sensory input was provided in
346 lower frequencies [10,27]. Rohenkohl and Nobre [10], for instance, used rhythmically presented
347 visual stimuli at 2.5 and 1.25 Hz moving across the screen until it disappeared behind an
348 occluder. Nevertheless, neural oscillations exclusively from the alpha band showed modulated
349 activity associated with temporal predictions during the disappearance time. They found no
350 phase locking of oscillations in lower frequencies.

351 In the current study, we provide further evidence that neural oscillations from the delta
352 band show enhanced phase alignment during visual temporal predictions across trials. In
353 order to adapt to the temporal regularity of the presented visual stimulus, delta frequencies in
354 a wide network of parietal and frontal brain areas exerted more consistent phase resets at
355 around the time point of disappearance of a visual stimulus as compared to a luminance
356 matching control condition. The strength of this phase adjustment in each participant
357 correlated with the consistency in judging a reappearance of the visual stimulus as too early or
358 too late. This was the case only in the temporal prediction tasks, which underlines the
359 behavioral relevance of the observed phase adjustments for temporal predictions.

360 Moreover, within each participant's neural oscillation that showed the strongest ITPC
361 during temporal predictions, we found a clustering of phases roughly around ±90° at each
362 participant's ROT. This was not the case when using the frequencies showing the strongest
363 ITPC in the luminance matching condition, where timing was not as important. The bimodal
364 distribution with peaks at 90° as well as -90° was most likely caused by analyzing the data
365 from all participants as well as sensors from both hemispheres together. Possibly differently
366 oriented generators in each participant as well as flips of the phase across hemispheres make
367 it difficult to differentiate between excitable and inhibitory phases of the oscillation using
368 whole-head scalp recordings. Nevertheless, the peaks at ±90° provide strong support for the
369 notion that in the context of temporal predictions the phase of delta oscillations adjusts to the
370 temporal structure of the stimulation to code for the timing of the predicted reappearance. We
371 propose that within each individual's subjective temporal framework, neural oscillations
372 adjusted their phase to the external stimulation such that a phase of high excitability
373 eventually coincided with each individual's predicted time point of reappearance. Our results
374 are in line with results reported by Cravo et al. [11], who showed that contrast sensitivity was a
375 function of the phase of entrained delta oscillations. In their study, the strongest contrast
376 sensitivity for visual stimuli was also observed at a delta phase around 90°. This phase range
377 might therefore indicate an optimal phase for processes related to temporal prediction.

378 Importantly, our study suggests that the mechanism of phase adjustments for temporal
379 predictions can be extended to external stimulation that does not as such involve rhythms. We
380 found that low-frequency oscillations can adjust their phase also to the temporal structure of
381 external stimulation that had to be inferred from motion. Many natural stimuli comprise
382 highly predictable regularities, but not all of them are intrinsically rhythmic. Our results

383 therefore indicate that the framework of phase adjustments during temporal predictions might
384 be generalized to all forms of predictive stimulation.

385 **Enhanced ITPC cannot be explained by stimulus-driven processes**

386 In earlier investigations of phase adjustments to external stimulation participants were
387 mostly presented with streams of rhythmic input. However, rhythmic input also causes
388 evoked brain activity within the same frequency range, which makes it difficult to disentangle
389 streams of evoked activity from entrained endogenous neural oscillations [18,21].

390 Our results provide evidence that phase resets of low-frequency oscillations observed
391 during temporal predictions cannot solely be explained by stimulus-evoked, bottom-up brain
392 activity [see also, 21,28]. In the current study, we aimed at reducing such brain responses to a
393 minimum by presenting participants with a continuously moving stimulus instead of several
394 discrete stimuli. We were particularly interested in the time point at which the stimulus
395 transiently disappeared behind an occluder (as opposed to sharp onsets and offsets in
396 rhythms). At disappearance, we did not observe an increase in low-frequency power as
397 compared to pre-stimulus baseline in any of the conditions, which could have been associated
398 with evoked brain activity such as, for instance, the contingent negative variation [CNV; 18].

399 Moreover, by introducing a novel experimental design, in which physical stimulation was
400 exactly the same as during temporal predictions as well as a control condition, we controlled
401 for brain responses that could be driven by bottom-up stimulus processing and are not specific
402 to temporal predictions. Importantly, delta ITPC but not power was stronger during temporal
403 predictions (see also Fig. S1). This provides strong evidence that ongoing, endogenous neural
404 oscillations underwent a phase reset around the time point of disappearance, which was more
405 consistent during temporal predictions than during the luminance matching task. These phase
406 resets can therefore not be solely related to brain responses evoked by the offset of the visual
407 movement, since we did not observe power differences at low frequencies.

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

409 We observed enhanced ITPC values in a network of mostly frontal and parietal brain
410 areas during visual as well as crossmodal temporal predictions. Similarly, Besle et al. [29]
411 observed significant phase entrainment to audiovisual stimulation in a wide network of
412 distributed areas including parietal and inferior frontal areas. These observations support the
413 notion that brain areas involved in temporal predictions may constitute a frontoparietal timing
414 network [6,30].

415 Further, we found enhanced ITPC values also in early somatosensory areas contralateral
416 to the disappearance of the purely visual stimulus during crossmodal temporal predictions,
417 despite the fact that prediction-relevant information was provided only by a moving visual
418 stimulus. This supports evidence reported earlier showing that stimulation within one
419 modality can crossmodally reset the phase of ongoing low-frequency in other modalities,
420 which might be an important mechanism for multisensory integration processes [22,23].

421 Moreover, strong correlations between ITPC and behavior were also observed in the
422 cerebellum, supporting earlier reports on a involvement of the cerebellum in temporal
423 prediction processes [31]. Roth and coworkers [32], for instance, found that cerebellar patients
424 were significantly impaired in recalibrating sensory temporal predictions of a reappearing
425 visual stimulus. This finding is of particular interest as we adapted the authors' experimental
426 paradigm for the use in the current study. Theirs and our results therefore indicate that the

cerebellum might be crucially involved in accurate and consistent judgments of temporal regularities deployed in perceiving object motion.

Conclusions

We provide strong evidence that the phase of neural oscillations can adjust to the temporal regularities of external stimulation and do not arise as a byproduct of bottom-up stimulus processing. Such phase alignments could provide a key mechanism that predicts the onset of upcoming events in order to optimize processing of relevant information and thereby adapt behavior. We show that temporal information provided to one modality leads to phase adjustments in another modality when crossmodal temporal predictions are necessary, providing further evidence that such crossmodal phase resets could be the neuronal basis of multisensory integration processes. Importantly, we observed that these phase adjustments reflected each individual's subjective temporal predictions time points. This supports the notion that the phase of neural oscillation indeed codes for the subjective estimation of elapsed time. Taken together, our results provide important insights into the neural mechanisms that might be utilized by the brain to predict the temporal onsets of upcoming events.

Materials and Methods

An exhaustive description of the methods can be found in the SI.

Participants and experimental procedure

Twenty-three healthy volunteers took part in the study. The ethics committee of the Medical Association Hamburg approved the study protocol and the experiment was carried out in accordance with the approved guidelines and regulations.

The experimental paradigm used in the current study was adopted from an earlier report investigating visual temporal predictions in cerebellar patients [32]. Our experiment consisted of three conditions: a *visual* temporal prediction task, a crossmodal (*tactile*) temporal prediction task, and a *luminance matching* (control) task. The trials of all conditions started with the presentation of a randomly generated, white noise occluder presented in the middle of the screen. We instructed participants to fixate the central fixation dot throughout the entire trial. After 1500 ms, an oval stimulus moved from the periphery towards the occluder with constant speed. The luminance of the stimulus differed in all trials (6 steps). 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 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 luminance matching task, we also manipulated the luminance of the reappearing stimulus relative the luminance the stimulus had before disappearance in each trial (also using 20 different values). 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.

The visual temporal prediction as well as the luminance matching 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*. In the luminance matching task, participants were asked to judge whether the luminance of the reappearing visual stimulus became *brighter* or *darker*.

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. The tactile stimulus was presented by means of a Braille piezostimulator for 200 ms. 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. At the end of each block, they were informed about the overall accuracy of their answers within the last block. Each block consisted of 60 trials, resulting in a total number of 480 trials per condition or 1,440 trials in total.

We used MATLAB R2014b (MathWorks, Natick, USA; RRID: SCR_001622) and Psychtoolbox [33] (RRID: SCR_002881). To mask the sound of the Braille stimulator during tactile stimulation, we presented participants with auditory pink noise at sampling rate of 48

kHz and volume of 85 dB using MEG-compatible in-ear headphones during all experimental blocks.

Data acquisition and pre-processing

MEG was recorded at a sampling rate of 1,200 Hz using a 275-channel whole-head system (CTF MEG International Services LP, Coquitlam, Canada). Online head localizations [34] were used to navigate participants back to their original head position prior to the onset of a new experimental block if their movements exceeded five mm from their initial position.

We analyzed reaction time data using R [35] (RRID: SCR_001905) and RStudio (RStudio Inc., Boston, USA; RRID: SCR_000432). Trials with reaction times longer than three standard deviations were excluded from analysis. Due to the right-skewed nature of reaction times, reaction time data were first log-transformed and then standardized across all trials from each participant.

All other data were analyzed using MATLAB R2016b with FieldTrip [36] (RRID: SCR_004849), the MEG and EEG Toolbox Hamburg (METH, Guido Nolte; RRID: SCR_016104), or custom made scripts. Each trial was cut 1,250 ms earlier to stimulus movement onset and 1,250 ms after offset of the reappeared stimulus. Trials containing strong muscle artifacts or jumps were detected by semi-automatic procedures implemented in FieldTrip and excluded from analysis. The remaining trials were filtered with a high-pass filter at 0.5 Hz, a low-pass filter at 170 Hz, and three band-stop filters at 49.5–50.5 Hz, 99.5–100.5 Hz and 149.5–150.5 Hz and subsequently down-sampled to 400 Hz.

We performed an independent component analysis (infomax algorithm) to remove components containing eye-movements, muscle, and cardiac artefacts. As a final step, using procedures described by Stolk *et al.* [34] we identified trials in which the head position of the participant differed by 5 mm from the mean circumcenter of the head position from the whole session and excluded them from further analysis.

Quantification and statistical analysis

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

Psychometric curve

We fitted a psychometric curve to the behavioral data of each participant from all trials in each condition. First, for each timing difference or luminance difference, respectively, we computed the proportion of "too late" or "brighter" answers for each participant. Then, we fitted a binomial logistic regression (psychometric curve) using the glmfit.m and gmlval.m functions provided in MATLAB. The fitted timing or luminance difference value at 50% proportion "too late" or "brighter" answers was determined as ROT or PSE for each participant, respectively. To test for a significant bias towards one of the answers, we tested the ROT or PSE from all participants against zero using one-sample *t*-tests ($\alpha = .05 / 3 =$

.017). The steepness of the psychometric function was computed as the reciprocal of the difference between fitted timing or luminance difference values at 75% and 25% proportion "too late" or "brighter" answers, respectively.

Linear model

We averaged RT across all luminance differences within each timing difference bin in each condition and then utilized a second-order (quadratic) polynomial regression model with timing difference as predictor for reaction times and computed the first- and second-order coefficients for each participant in each condition. The coefficients from all participants were then tested against zero using one-sample *t*-tests in all conditions ($\alpha = .05 / 3 = .017$).

Spectral power

We decomposed the MEG recordings into time-frequency representations by convolving the data with complex 40 Morlet's wavelets [37], logarithmically spaced between 0.5 to 100 Hz and with logarithmically increasing number of cycles from two to ten cycles. For all analyses of the MEG data, we considered subjectively correct trials only, i.e., trials in which participants answered correctly based on their individual ROT. 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 Fig. 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.

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 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 [38]. For each window, a separate cluster-permutation test was performed ($\alpha = .05$; liberally chosen to observe all ongoing power modulations; see Results section).

We subsequently compared the spectral power estimates averaged within the beta range (13–30 Hz; see Results section) at each time point within the disappearance window and all channels from the visual or tactile temporal prediction task with the luminance matching task. We again employed cluster-permutation statistics, this time by clustering neighboring channels and time points. We used a one-sided $\alpha = .025 / 2 = .0125$, since negative and positive clusters were tested separately, and to adjust for the two separate comparisons between the conditions (used throughout the study unless stated differently).

To estimate spectral power in source space, we computed separate leadfields for each recording session and participant based on each participant's mean head position in each session and individual magnetic resonance images. We used the single-shell volume conductor model [39] with a 5,003 voxel grid that was aligned to the MNI152 template brain (Montreal Neurological Institute, MNI; http://www.mni.mcgill.ca) as implemented in the METH toolbox. Cross-spectral density (CSD) matrices were computed from the complex wavelet convolved data in steps of 100 ms in the same time windows as outlined above. To

avoid biases in source projection, common adaptive linear spatial filters (DICS beamformer [40]) pointing into the direction of maximal variance were computed from CSD matrices averaged across all time bins and conditions for each frequency.

All time-frequency resolved CSD matrices were then multiplied with the spatial filters to estimate spectral power in each of the 5,003 voxels and normalized with the pre-stimulus baseline window. We then averaged across all time bins within the disappearance window and utilized cluster-based permutation statistics to identify clusters of voxels that show statistical difference in beta power between each of the temporal prediction tasks and the luminance matching 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 [37]. 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. 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 and all sensors between each of the temporal prediction tasks and the luminance matching task. ITPC on source level was computed using the same leadfields and common beamformer filters as for spectral power (see above).

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

Delta phase clustering at ROT

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

For this analysis, we only considered trials in which the stimulus reappeared later than each individual's ROT and the participant answered subjectively correct. By this, we prevented possible phase distortions by the external stimulation earlier to or at ROT. Moreover, to make sure that we reduced also activity that was related to external stimulations

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

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

Within these channels, for each individual participant we determined the frequency within the 0.5 to 3 Hz delta band, which showed the strongest ITPC for the visual or the tactile prediction as compared to the luminance matching task, respectively, in the same time window of 0 to 1.500 ms. For the luminance matching condition, we extracted the frequencies showing the strongest estimates of ITPC in the luminance matching as compared to the visual temporal prediction task and used individual ROTs from the visual prediction task. For these individual frequencies, we plotted the phase angle at ROT (as described above) from all the considered channels and all participants in a histogram (in bins of 10°; see Fig. 5). We computed the distance from the observed phase distribution to a uniform distribution using a discrete and normalized version of the Kullback-Leibler distance, i.e., the modulation index (MI) [24].

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

Competing interests

The authors declare no competing interests.

Data availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

675 **References**

676 1. Buzsáki, G. *Rhythms of the brain*. (Oxford University Press., 2006).

677 2. Engel, A. K., Fries, P. & Singer, W. Dynamic predictions: oscillations and synchrony
678 in top-down processing. *Nat. Rev. Neurosci.* **2**, 704–16 (2001).

679 3. Fries, P. A mechanism for cognitive dynamics: neuronal communication through
680 neuronal coherence. *Trends Cogn. Sci.* **9**, 474–80 (2005).

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

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

684 6. Rimmele, J. M., Morillon, B., Poeppel, D. & Arnal, L. H. Proactive Sensing of
685 Periodic and Aperiodic Auditory Patterns. *Trends Cogn. Sci.* **22**, 870–882 (2018).

686 7. Stefanics, G. *et al.* Phase Entrainment of Human Delta Oscillations Can Mediate the
687 Effects of Expectation on Reaction Speed. *J. Neurosci.* **30**, 13578–13585 (2010).

688 8. Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I. & Schroeder, C. E. Entrainment of
689 neuronal oscillations as a mechanism of attentional selection. *Science* **320**, 110–3
690 (2008).

691 9. Gould, I. C., Rushworth, M. F. & Nobre, A. C. Indexing the graded allocation of
692 visuospatial attention using anticipatory alpha oscillations. *J. Neurophysiol.* **105**, 1318–
693 1326 (2011).

694 10. Rohenkohl, G. & Nobre, A. C. Alpha Oscillations Related to Anticipatory Attention
695 Follow Temporal Expectations. *J. Neurosci.* **31**, 14076–14084 (2011).

696 11. Cravo, A. M., Rohenkohl, G., Wyart, V. & Nobre, A. C. Temporal Expectation
697 Enhances Contrast Sensitivity by Phase Entrainment of Low-Frequency Oscillations in
698 Visual Cortex. *J. Neurosci.* **33**, 4002–4010 (2013).

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

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

704 14. Giraud, A.-L. & Poeppel, D. Cortical oscillations and speech processing: emerging
705 computational principles and operations. *Nat. Neurosci.* **15**, 511–7 (2012).

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

708 16. Saleh, M., Reimer, J., Penn, R., Ojakangas, C. L. & Hatsopoulos, N. G. Fast and Slow
709 Oscillations in Human Primary Motor Cortex Predict Oncoming Behaviorally Relevant
710 Cues. *Neuron* **65**, 461–471 (2010).

711 17. Gomez-Ramirez, M. *et al.* Oscillatory Sensory Selection Mechanisms during
712 Intersensory Attention to Rhythmic Auditory and Visual Inputs: A Human
713 Electrocorticographic Investigation. *J. Neurosci.* **31**, 18556–18567 (2011).

714 18. Breska, A. & Deouell, L. Y. Neural mechanisms of rhythm-based temporal prediction:
715 Delta phase-locking reflects temporal predictability but not rhythmic entrainment.
716 *PLOS Biol.* **15**, e2001665 (2017).

717 19. Obleser, J., Henry, M. J. & Lakatos, P. What do we talk about when we talk about

718 rhythm? *PLOS Biol.* **15**, e2002794 (2017).

719 20. Novembre, G. & Iannetti, G. D. Tagging the musical beat: Neural entrainment or
720 event-related potentials? *Proc. Natl. Acad. Sci.* **115**, E11002–E11003 (2018).

721 21. Doelling, K. B., Assaneo, M. F., Bevilacqua, D., Pesaran, B. & Poeppel, D. An
722 oscillator model better predicts cortical entrainment to music. *Proc. Natl. Acad. Sci.*
723 201816414 (2019). doi:10.1073/pnas.1816414116

724 22. Lakatos, P., Chen, C.-M., O'Connell, M. N., Mills, A. & Schroeder, C. E. Neuronal
725 oscillations and multisensory interaction in primary auditory cortex. *Neuron* **53**, 279–
726 92 (2007).

727 23. Mercier, M. R. *et al.* Auditory-driven phase reset in visual cortex: Human
728 electrocorticography reveals mechanisms of early multisensory integration.
729 *Neuroimage* **79**, 19–29 (2013).

730 24. Tort, A. B. L., Komorowski, R., Eichenbaum, H. & Kopell, N. Measuring phase-
731 amplitude coupling between neuronal oscillations of different frequencies. *J.*
732 *Neurophysiol.* **104**, 1195–210 (2010).

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

735 26. van Ede, F., de Lange, F., Jensen, O. & Maris, E. Orienting Attention to an Upcoming
736 Tactile Event Involves a Spatially and Temporally Specific Modulation of
737 Sensorimotor Alpha- and Beta-Band Oscillations. *J. Neurosci.* **31**, 2016–2024 (2011).

738 27. Samaha, J., Bauer, P., Cimaroli, S. & Postle, B. R. Top-down control of the phase of
739 alpha-band oscillations as a mechanism for temporal prediction. *Proc. Natl. Acad. Sci.*
740 **112**, 8439–8444 (2015).

741 28. Kösem, A. *et al.* Neural Entrainment Determines the Words We Hear. *Curr. Biol.* **28**,
742 2867-2875.e3 (2018).

743 29. Besle, J. *et al.* Tuning of the Human Neocortex to the Temporal Dynamics of Attended
744 Events. *J. Neurosci.* **31**, 3176–3185 (2011).

745 30. Coull, J. T. & Nobre, A. C. Dissociating explicit timing from temporal expectation
746 with fMRI. *Curr. Opin. Neurobiol.* **18**, 137–144 (2008).

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

749 32. Roth, M. J., Synofzik, M. & Lindner, A. The Cerebellum Optimizes Perceptual
750 Predictions about External Sensory Events. *Curr. Biol.* **23**, 930–935 (2013).

751 33. Brainard, D. H. The psychophysics toolbox. *Spat. Vis.* **10**, 433–6 (1997).

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

754 35. R Core Team. R: A Language and Environment for Statistical Computing. *R*
755 *Foundation for Statistical Computing Vienna Austria* (2014).

756 36. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. FieldTrip: Open Source
757 Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological
758 Data. *Comput. Intell. Neurosci.* **2011**, 1–9 (2011).

759 37. Cohen, M. *Analyzing Neural Time Series Data: Theory and Practice*. (MIT Press,
760 2014).

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

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

766 40. Gross, J. *et al.* Dynamic imaging of coherent sources: Studying neural interactions in
767 the human brain. *Proc. Natl. Acad. Sci.* **98**, 694–699 (2001).

768