




Dissociated amplitude and phase effects of alpha oscillation in a nested structure of rhythm- and sequence-based temporal expectation

Zhongbin Su ^{1,2,3}, Xiaolin Zhou ^{4,5,*}, Lihui Wang ^{1,2,6,*}

¹Shanghai Key Laboratory of Psychotic Disorders, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai 200030, China,

²Institute of Psychology and Behavioral Science, Shanghai Jiao Tong University, Shanghai 200030, China,

³Beijing Key Laboratory of Behavior and Mental Health, School of Psychological and Cognitive Sciences, Peking University, Beijing 100871, China,

⁴Shanghai Key Laboratory of Mental Health and Psychological Crisis Intervention, School of Psychology and Cognitive Science, East China Normal University, Shanghai 200062, China,

⁵PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing 100871, China,

⁶Shanghai Center for Brain Science and Brain-Inspired Intelligence Technology, Shanghai 201602, China

*Corresponding authors: Xiaolin Zhou, School of Psychology and Cognitive Science, East China Normal University, Shanghai 200062, China.

Email: xz104@psy.ecnu.edu.cn. Lihui Wang, Institute of Psychology and Behavioral Science, Shanghai Jiao Tong University, Shanghai 200030, China.

Email: lihui.wang@sjtu.edu.cn

The human brain can utilize various information to form temporal expectations and optimize perceptual performance. Here we show dissociated amplitude and phase effects of prestimulus alpha oscillation in a nested structure of rhythm- and sequence-based expectation. A visual stream of rhythmic stimuli was presented in a fixed sequence such that their temporal positions could be predicted by either the low-frequency rhythm, the sequence, or the combination. The behavioral modeling indicated that rhythmic and sequence information additively led to increased accumulation speed of sensory evidence and alleviated threshold for the perceptual discrimination of the expected stimulus. The electroencephalographical results showed that the alpha amplitude was modulated mainly by rhythmic information, with the amplitude fluctuating with the phase of the low-frequency rhythm (i.e. phase-amplitude coupling). The alpha phase, however, was affected by both rhythmic and sequence information. Importantly, rhythm-based expectation improved the perceptual performance by decreasing the alpha amplitude, whereas sequence-based expectation did not further decrease the amplitude on top of rhythm-based expectation. Moreover, rhythm-based and sequence-based expectations collaboratively improved the perceptual performance by biasing the alpha oscillation toward the optimal phase. Our findings suggested flexible coordination of multiscale brain oscillations in dealing with a complex environment.

Key words: temporal expectation; neural entrainment; prestimulus oscillation; alpha amplitude; alpha phase.

Introduction

We live in a world teemed with temporal regularities: day and night alternate with each other approximately every 12 h; the thunder is always heard after the seeing of the lightning. Such temporal regularity affords the temporal expectation of an event, which avails to guide perception and action that serve the current goal (Jones 1976). A few minutes after leaving home for work in a morning routine, for instance, one can quickly spot the cafe and grab a coffee without paying attention to the surroundings. In laboratory settings, a stimulus stream with temporal regularity can lead to improved task performance such as increased accuracy and/or facilitated reaction times (RTs) in discriminating the perceptual attributes of the expected stimulus (Sanabria et al. 2011; Cravo et al. 2013; Morillon et al. 2016). The improved perceptual performance was supported by sharpened neural representation (de Lange et al. 2018) and earlier neural excitation (Anderson and Sheinberg 2008) for the expected event.

Temporal expectation can be built upon multiple bases such as a fixed rhythm or a fixed sequence of a stimulus stream (Nobre and van Ede 2018). For instance, a red circle can appear every

800 ms (i.e. a regular rhythm of 1.25 Hz), or always come after two successive blue circles (i.e. a regular sequence). In the case of a regular sequence, the successive stimuli were not necessarily presented in a certain rhythm, and hence sequence-based temporal expectation is dissociable from the rhythm-based expectation. Although either the rhythmic or the sequence regularity alone can be predictive of a stimulus, and there is already evidence suggesting both common and dissociable mechanisms of different forms of temporal expectation (Correa et al. 2014; Breska and Deouell 2017; Bouwer et al. 2020), it remains unclear how the nested rhythmic and sequence information are combined to affect the sensorimotor processing and the corresponding behavioral performance. This is of significance as different sources of temporal regularity are often nested in a real-world situation.

The rhythmic regularity has been associated with neural entrainment, a process where the intrinsic neural oscillations are tuned to the rhythm of the external stimuli (Jones 2010; Calderone et al. 2014; Haegens and Golumbic 2018). By entrainment, a stimulus gains optimized processing when aligned with the high-excitability phase of the oscillatory neural signals relative to the low-excitability phase of the oscillatory signals, and the

perceptual performance was dependent on the phase of the entrained activity (Stefanics et al. 2010; Cravo et al. 2013). It has been shown that the phase synchronization of the low-frequency (i.e. delta-band) oscillation not only underlined the temporal expectation afforded by the rhythmic stimuli, but also the temporal expectation based on memory (Breska and Deouell 2017), suggesting a general role of phase synchronization in multiple forms of temporal expectation.

Another important component of brain activity that has been linked to temporal expectation is the alpha-band (~8–12 Hz) oscillation, which is suggested as causal for the timing and perceptual processing of the upcoming stimulus (Romei et al. 2010; Klimesch 2012). On the one hand, the prestimulus alpha oscillation showed decreased amplitude during temporal expectation, indicating increased neural excitation in preparation for the expected stimulus (Rohenkohl and Nobre 2011; van Diepen et al. 2015). On the other hand, it has been shown that the temporal expectation of a stimulus improved the perceptual performance by tuning the prestimulus alpha oscillation into an optimal phase (Busch et al. 2009; Samaha et al. 2015). Of note, there is also debate concerning whether alpha amplitude or alpha phase was responsible for the temporal expectation and the enhanced sensorimotor processing (van Diepen et al. 2015). One possible account to explain the 2-fold results is that the amplitude decrease and the phase concentration of prestimulus alpha activity may play dissociable roles in different forms of temporal expectation.

In the present study, we asked how the nested temporal expectation by rhythmic and sequence regularity affected the perceptual performance of the expected stimulus. For this purpose, a visual stream of rhythmic stimuli was presented in a fixed sequence such that their temporal positions could be predicted by either the low-frequency rhythm, the sequence, or the combination. For each of the three conditions, a visual stream of arrhythmic was presented as the control. We hypothesized that the combination of sequence-based expectation and rhythm-based expectation would benefit the perceptual performance more than when only one kind of expectation was available. We predicted that the rhythmic and sequence regularity would have a combined effect in improving the perceptual performance of the predicted stimulus, leading to increased accuracy or/and decreased RT. A perceptual performance engages multiple cognitive modules. In a common perceptual decision-making task, the perceptual attribute has to be recognized and transformed into motor responses based on the stimulus–response mapping defined by the current task. These cognitive modules can be estimated with computational modeling such as the drift-diffusion model (DDM) (Ratcliff et al. 2016). With the combination of the DDM, we assessed whether and to which extent the accumulation speed of perceptual evidence, the threshold of the perceptual decision, and the speed of motor implementation were affected by different structures of temporal expectation.

At the neural level, electroencephalographical (EEG) activity was simultaneously recorded to uncover the neural mechanism. We hypothesized that the rhythmic stimuli would lead to neural entrainment, with the onbeat target aligned with the high-excitability phase and the offbeat target aligned with the low-excitability phase. To test this hypothesis, we estimated the phase synchronization of the stimuli frequency both during and after the stimuli. If the neural oscillation would be entrained by the rhythmic stimuli, the phase synchronization of the rhythmic stimuli would be stronger than the arrhythmic stimuli, and this oscillation would outlast after the offset of the stimuli sequence. In a second step, we investigated if the low-frequency rhythm

of the visual stimuli communicated with the prestimulus alpha activity. As it has been suggested that the high-frequency neural activity can be modulated by the phase of the low-frequency activity (Lakatos et al. 2008), we expected that there would be cross-frequency couplings between the low-frequency activity and the prestimulus alpha activity. Crucially, we aimed to tease apart the potentially different roles of alpha amplitude and alpha phase in different temporal expectations. For this purpose, in a third step, we dissociated the analysis of alpha amplitude and alpha phase and assessed if the amplitude and the phase would be distinctively modulated by different structures of temporal expectation. The DDM was also included to assess to which extent alpha amplitude and alpha phase could be related to the cognitive components (Samaha et al. 2020).

Materials and methods

Participants

A total of 26 right-handed university students participated in this study. One participant was excluded from analysis due to incomplete data caused by technical errors during the experiment, and another participant was excluded from data analysis due to excessive artifacts (50% of total trials) of the EEG signals, resulting in 24 participants (10 females, mean age 20.5 years old). All participants had normal or corrected-to-normal visual acuity and normal color vision and reported no history of psychiatric or neurological disorders. The experiment was carried out in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the School of Psychological and Cognitive Sciences, Peking University (#2019-12-01).

Stimuli and apparatus

Stimuli were created and presented using Psychtoolbox-3 extension for MATLAB (Brainard 1997). Stimuli were presented on a Display++ monitor (1920*1080 spatial resolution, 120 Hz refresh rate) against a gray background (Red-Green-Blue color spaces, RGB:125, 125, 125). The eye-to-monitor distance was fixed at 70 cm. A chin rest was used to maintain the head position and a constant viewing distance. Responses were collected using a standard keyboard.

Stimuli were presented in a stream flow at the center of the screen, one at a time (Fig. 1A). Two types of stimuli were used in the experiments: standard stimuli and target stimuli. Standard stimuli were disks of Gaussian noise (radius: 2° of visual angle), each of which was surrounded by a blue ring (RGB: 0, 0, 200). The disks were created by smoothing Gaussian patches with a two-dimensional kernel. The smoothing dimension (0.083° of visual degree) and the root-mean-square contrast of the noise patch were fixed across experiments and participants. Each target stimulus was a Gabor patch surrounded by a red ring (RGB:200, 0, 0) and was embedded in the standard stimuli. The orientation of the Gabor patch was either 135° or 45° relative to the horizontal axis, and the spatial frequency of the Gabor patches was 2 cycles/degree of visual angle. The above-mentioned parameters of the stimuli were chosen following a previous study (Rohenkohl et al. 2012). Prior to the experiment, a psychophysical test was conducted to estimate the contrast threshold with 75% accuracy in discriminating the orientation of the Gabor patches, with an adaptive staircase procedure (Kaernbach 1991). The contrast that 10^{0.1} times of contrast threshold corresponding to 75% accuracy was used in the formal experiment across all conditions.

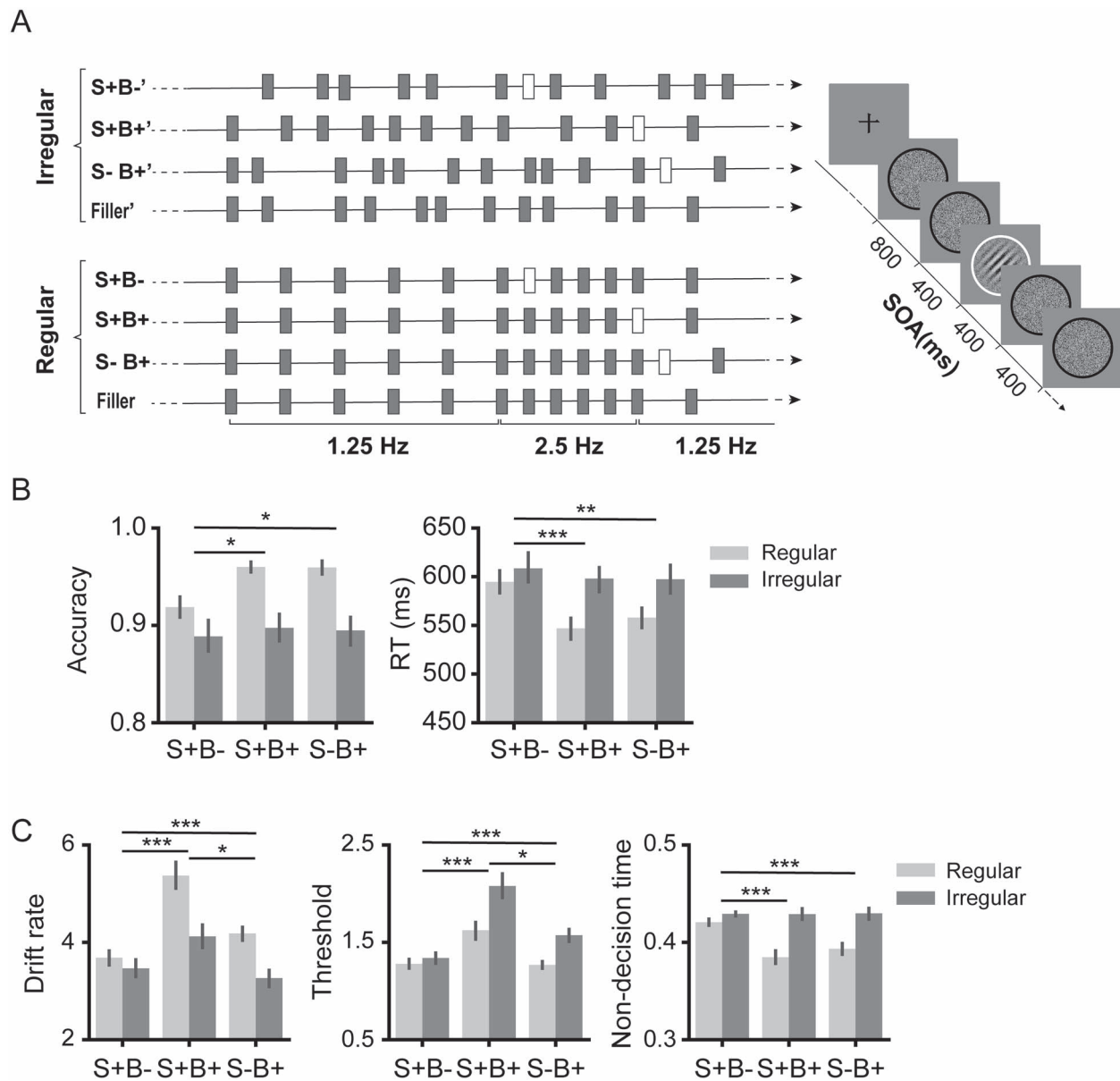


Fig. 1. Experimental design and behavioral results. (A) Left: The example of stimulus stream in different conditions. The dark gray bars indicate standard stimuli and the white bars indicate target stimuli. Right: The sequence of stimuli in an example trial. The target was a near-threshold Gabor patch surrounded by a red ring (illustrated as a white ring here), and the standard stimuli were visual noise surrounded by a blue ring (illustrated as a black ring here). Participants were asked to make a discriminative response to the orientation of the Gabor patch. (B) Accuracies (left) and RTs (right) are shown as a function of the experimental conditions. (C) Parameters estimated in the hierarchical drift-diffusion model (HDDM) are shown as a function of the experimental conditions. Error bars indicate standard errors of the mean (SEM) across participants. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Experimental design and procedure

There were two types of stimulus streams: the regular stream and the irregular stream. In both streams, each stimulus remained on the screen for 50 ms. Within the regular stream, stimuli were presented at two alternating frequencies (i.e. 1.25 Hz, and 2.5 Hz). The stimulus onset asynchrony (SOA) between successive stimuli changed with a fixed rule: five SOAs of 800 ms followed by five SOAs of 400 ms that was then followed by five 800 ms SOAs and so on. Every 10 stimuli were grouped into a unit where the long-SOA stimuli were always followed by the short-SOA stimuli. Within the irregular stream, the SOA between each successive two stimuli was randomly chosen from 300 ms, 400 ms, 500 ms, 600 ms, 700 ms, 800 ms, and 900 ms. Within the irregular stream, every 10 successive stimuli were also grouped into a unit. There

were 17 units in each stream, and the two consecutive units were connected in an end-to-end manner to avoid hazard rate effects (Luce 1986).

In each unit, there was a 75% probability that one target stimulus was embedded in the stream while a 25% probability that no target was presented at all (no-target, filler unit). No target was presented in the very first unit of each trial. In the regular stream, the target was presented at one of the three temporal positions with equal probability (as shown in Fig. 1A): (i) the first position of the short-SOA stimuli following the long-SOA stimuli; (ii) the last position of the short-SOA stimuli that followed by the long-SOA stimuli in the next unit; (iii) an extra position inserted after the last position of the short-SOA stimuli that followed by the long-SOA stimuli. In these three conditions,

there was always rhythmic information during the presentation of the stimuli. However, whether the target could be predicted by the sequence information and whether the target was onbeat or offbeat differed among the three conditions. In the first of the above three conditions, the target was highly predictable because of the fixed sequence structure while nevertheless being located at an offbeat position of the 1.25 Hz rhythm (S+B-). In the second condition, the target was also highly predictable while being located at an onbeat position of the 2.5 Hz rhythm (S+B+). In the third condition, the target was not expected in the sequence while being located at an onbeat position of the 2.5 Hz rhythm (S-B+). In the irregular stream, stimuli were presented with varying inter-stimulus intervals. The three positions within the irregular stream (S'+B'-, S'+B'+, S'-B'+) were matched with the positions within the regular stream in the way that the numbers of stimuli between two target stimuli were the same in the two streams, and the time intervals before and after the positions within the irregular stream were consistent with the regular stream. Given that both the stimulus number and the pre-target intervals were controlled between regular and irregular conditions, the observed difference between the two conditions cannot be due to that the accumulated amount of information was different.

Each trial started with a black central fixation cross (RGB: 0, 0, 0, 1° of visual angle), which remained on the screen until a space key was pressed by the participant to begin the current trial. The stimulus stream was immediately presented after the key press. Participants were asked to respond to the orientation of the target stimulus (left vs. right) with the left and right index fingers, respectively. They were asked to make responses as quickly and accurately as possible. A white central fixation cross (RGB: 255, 255, 255, 1° of visual angle) was presented at the end of each trial and remained on the screen for 10 s until the next trial started.

In each trial, there were 158 standard stimuli and 12 target stimuli, and the target was presented at one of the three positions with equal probability. Units with different target positions were mixed and presented in a pseudorandom order such that no more than two consecutive units had the same target position. There were 20 trials of the regular stream and 20 trials of irregular stream, resulting in 240 targets for each of the two streams. The two types of trials were presented in two separate blocks, with the order of the two blocks counterbalanced across participants. After every five trials, there was a 1-min break.

Therefore, the experiment had a 2 (Regularity: regular vs. irregular) × 3 (Target position: S+B-, S+B+, vs. S-B+) within-subject design.

Statistical analysis of behavioral data

For each participant, incorrect responses, omissions, and responses with RTs beyond mean RT ± 3 standard deviation (1.80%) in each condition were excluded. The mean RT of the remaining responses for each condition was calculated. The accuracy for each condition was calculated as the proportion of the number of correct responses against the total number of targets in each condition. Accuracy and RTs were subjected to a repeated-measure ANOVA with target position and stream regularity as within-subject factors. Given a significant interaction, further post-hoc tests were conducted to show if the improved behavioral performances by temporal regularity (i.e. RT and accuracy difference between regular and irregular conditions) were larger at one particular position than the other. Note that we did not test the simple effects of the RTs/accuracies at different positions (e.g. accuracy at S+B- regular vs. accuracy at S+B+ regular) because of the different numbers of stimuli before the target and the different lengths of the interval after the target. As the

number of stimuli and the interval were matched between regular and irregular conditions, the tests on the RT/accuracy difference between regular and irregular conditions were not subject to the different stimuli numbers or different intervals.

The DDM that characterizes the evolving perceptual decisions in a two-alternative forced-choice task (Sun and Landy 2016; Tavares et al. 2017; Stafford et al. 2020) was used to model the behavioral data. Here we estimated three parameters, drift rate, threshold, and nonddecision time to simulate the perceptual decision-making processes using the hierarchical drift-diffusion modeling (HDDM) 0.6.0 toolbox (Wiecki et al. 2013). The HDDM model is a hierarchical Bayesian estimation of drift-diffusion parameters and generates parameter estimates at both the individual level and the group level. Omission trials were excluded prior to modeling and the probability of outlier was set to 5%. In each experimental condition, for the convergence of the parameters, each of the three parameters was estimated with an independent model. Four Markov Monte-Carlo chains were used to estimate the parameters, with 10,000 samples in each chain while the first 2,000 samples were discarded as burn-in to achieve convergence. We computed the R-hat (Gelman-Rubin) convergence statistics to ensure the convergence of the models (Gelman and Rubin 1992). Individual parameters were averaged across the remaining 32,000 samples for further analysis. The same repeated-measure ANOVA and post-hoc comparisons as above were performed on the HDDM parameters, with target position and stream regularity as the within-subject factors.

EEG recording and preprocessing

EEG signals were recorded by 64 Ag/AgCl electrodes mounted in an elastic cap (Easy-cap Brain Products, Germany) according to an extending 10–20 system. Vertical electrooculograms (EOG) were recorded by an electrode placed below the center of the right eye. The impedance of all electrodes was kept below 5 kΩ. The EEG and EOG signals were amplified by two Brain-Amp amplifiers (Brain Product, Germany), digitalized to a sample of 500 Hz, and were online filtered by a band-pass filter of 0.016–100 Hz. EEG signals were online referenced to the FCz electrode. Preprocessing (denoising and segmenting) was conducted with EEGLAB toolbox (Delorme and Makeig 2004). The offline data were band-pass filtered between 0.5 and 60 Hz and re-referenced to the averaged signal of right and left mastoid electrodes. Independent component analysis was performed to remove eye-movement and other artifact components (Drisdelle et al. 2017). Stimulus-locked epochs were extracted from the interval of -4,000 to 1,500 ms relative to the target onset. The long epoch was selected to avoid potential edge effects for the analysis in the frequency domain. Epochs with amplitude exceeding 100 μV were removed for further analysis.

Phase-locking values of low-frequency neural oscillation

To show the neural entrainment by the rhythmic visual stimuli, the PLVs of the pre-target activity were calculated over the visual cortex. The PLVs were expected to be higher in the regular condition than in the irregular condition for the entrained frequencies. For each of the occipital channels (Oz, O1, O2, POz, PO3, PO4, PO7, PO8), the stimulus-locked data (time range: -4,000 to 1,500 ms relative to target onset) was filtered using a two-way least-squares finite impulse response (FIR) filter (S+B-: 0.75–1.75 Hz, center frequency 1.25 Hz; S+B+ and S-B+: 2.0–3.0 Hz, center frequency 2.5 Hz; *eeffit*, EEGLAB). The phase of the filtered data was obtained with Hilbert transformation. For each participant and

each condition, the PLVs at each time point were calculated as the length of the mean vector of the phases across the targets of a specific condition. For each of the three positions, the difference in PLVs between the regular and the irregular condition was tested with a *paired-t*-test. This test was performed for each time point of the -400 to 50 ms interval relative to the target onset, with *cluster-based permutation* (number of iterations = 1,000, minimum time duration = 20 ms, cluster level $\alpha = 0.05$) for correcting multiple comparisons across the time points.

An alternative account could be that the observed PLVs were solely due to the transient visual stimulations of the individual stimuli rather than a result of neural oscillation entrained by the rhythmic stream. To exclude this alternative, we tested if the low-frequency oscillation sustained in the absence of the visual stimulations (van Bree et al. 2021). We compared the PLVs calculated for the data after the last stimulus of each trial in the regular condition (i.e. posttrial PLVs, Fig. 2C) with the PLVs for the data before the start of each trial in the regular condition (i.e. pretrial PLVs). To avoid any smearing effect of filtering on PLV estimations, the posttrial phase was estimated using Fast-Fourier Transform in a sliding time window of 1,200 ms ($-4,000$ to $3,498$ ms relative to the last stimulus, step size 2 ms). Zero-padding method was used to improve the frequency resolution to $5/12$ Hz ($1/3$ of 1.25 Hz and $1/6$ of 2.5 Hz). To avoid the impact of the last stimulus on the PLVs, posttrial PLVs were calculated within the time interval of 600 – $2,096$ ms relative to the “onset” of the first omitted stimulus for the trials ended with 1.25 Hz and the time interval of 600 – $2,496$ ms relative to the “onset” of the first omitted stimulus for the trials end with 2.5 Hz. Pretrial PLVs were calculated within the time interval of $-2,996$ to $-1,500$ ms relative to the first stimulus for the trial ended with 1.25 Hz and the time interval of $-3,396$ to $-1,500$ ms relative to the first stimulus for the trial ended with 2.5 Hz. The time intervals here referred to the sliding windows, and each time point within the interval referred to the center of the sliding time window. We avoid choosing the time interval immediately before the first stimulus to avoid any confounding factor related to motor preparation (the first stimulus was presented immediately after the button press for initiating the current trial). The alternative account can be ruled out if the posttrial PLVs were stronger than the pretrial PLVs. Considering the number of trials (20 trials in the regular condition for each participant), we compared the PLVs between posttrial and pretrial conditions using a Bootstrapping method (with 1,000 iterations) to ensure statistical reliability. For each iteration, the trials from a subset of participants (21 out of 24, with replacement) were randomly sampled to calculate the PLVs, with each estimation based on at least 168 trials. Statistical significance was assessed with the 95% confidence interval (CI) of the estimated PLVs.

To test if S+B– was at an antiphase of an optimal phase of 1.25 Hz entrained neural activity, the difference between the phase of 1.25 Hz at the target onset and the phase of 1.25 Hz at the onset of the preceding stimulus before the target was obtained for each participant. The preceding stimulus was expected to be at an optimal phase of the entrained neural activity at 1.25 Hz. Rayleigh test was used to test whether the phase difference was uniformly distributed or centered $\sim 180^\circ$. For S+B+ and S-B+, the phase difference of 2.5 Hz at the target onset between the two positions was calculated for each participant. The two positions were expected to be at an optimal phase of the entrained neural activity at 2.5 Hz. Rayleigh test was used to test whether the phase difference of 2.5 Hz between S+B+ and S-B+ was uniformly distributed or centered $\sim 0^\circ$.

Phase-amplitude coupling analysis

To investigate if the high-frequency activities were affected by the rhythmic regularity, phase-amplitude coupling (PAC) was performed to test the coupling relationship between the phase of low frequencies (0.5–5 Hz, in step of 0.25 Hz) and the amplitude of high frequencies (5–30 Hz, in step of 1 Hz) before the target. The phase and amplitude time courses for each frequency were extracted from the Hilbert transformations of the epoch filtered with a two-way least-squares FIR filter (*eegfilt*, EEGLAB, frequency range). The amplitudes of high frequencies were transformed to z-scores across the time points for each epoch. The PAC was calculated within the interval of $-2,000$ to 0 ms relative to target onset at S+B– and within the interval of $-1,000$ to 0 ms at S+B+ and S-B+. A vector of phase-amplitude time course was constructed, with the phase of the low frequencies as angle and the amplitude of the high frequencies as length. For each condition and epoch, the mean vector was calculated over the selected time intervals. The length of the mean vector was defined as the PAC value for the low-frequency phase and high-frequency amplitude. To avoid the bias that may be caused by the nonuniform distribution of the phase, we normalized the PAC value with nonparametric permutations (Cohen 2014). Specifically, for each epoch, each amplitude time course was cut at a random time point, and the data points before and after the cutting point were swapped. The PAC value based on the swapped data was calculated. This calculation was repeated 1,000 times, rendering a distribution of permuted PAC values. The normalized Z value of the PAC (PAC_z) was then obtained by subtracting the mean of the permuted PAC values from the un-permuted PAC values and divided by the standard deviation of the permuted PAC values. The PAC_z values were averaged across epochs and compared between the regular and irregular conditions with a *paired-t*-test. Cluster-based permutation (number of iterations = 1,000, minimum number of clusters = 30, cluster level $\alpha = 0.05$) was used for correcting multiple comparisons across phases and frequencies.

Although PAC analysis showed a coupling relationship between the phase of entrained frequencies and alpha amplitude (see results and Fig. 3A), an alternative account for the PAC between the low-frequency phase and alpha amplitude could be that the alpha amplitude was simply reset by the individual stimuli rather than fluctuating with the phase of the low-frequency oscillation. To test this alternative account, we compared the phase preferred by alpha (8–12 Hz) amplitude (alpha-preferred phase) at S+B– (1.25 Hz, regular condition) with the alpha-preferred phase at S+B+ (2.5 Hz, regular condition), and with the alpha-preferred at S-B+ (2.5 Hz, regular condition). Here the alpha-preferred phase in each condition was identified as the phase of the low-frequency oscillation where the alpha amplitude reached the peak. If the coupling between the alpha amplitude and the phase of the low-frequency oscillation was just a by-product of the physical stimulation, the alpha-preferred phase should be different between S+B– (1.25 Hz) and S+B+ (2.5 Hz), and between S+B– (1.25 Hz) and S-B+ (2.5 Hz). To calculate the alpha-preferred phase, for each participant, three cycles of the low-frequency oscillation ($-2,400$ to 0 ms for 1.25 Hz at S+B–; $-1,200$ to 0 ms for 2.5 Hz at S+B+ and S-B+) before the target onset were included. The alpha amplitudes were sorted according to the phases of the low-frequency oscillation and binned with a sliding phase window of $\pi/3$ in radians in a step of $\pi/18$. Then the alpha amplitude was averaged across epochs for each condition, and the phase at the peak (maximum value) of the alpha amplitude was defined as alpha-preferred phase (Fig. 3B). The difference between alpha-preferred phases was calculated and tested against zero

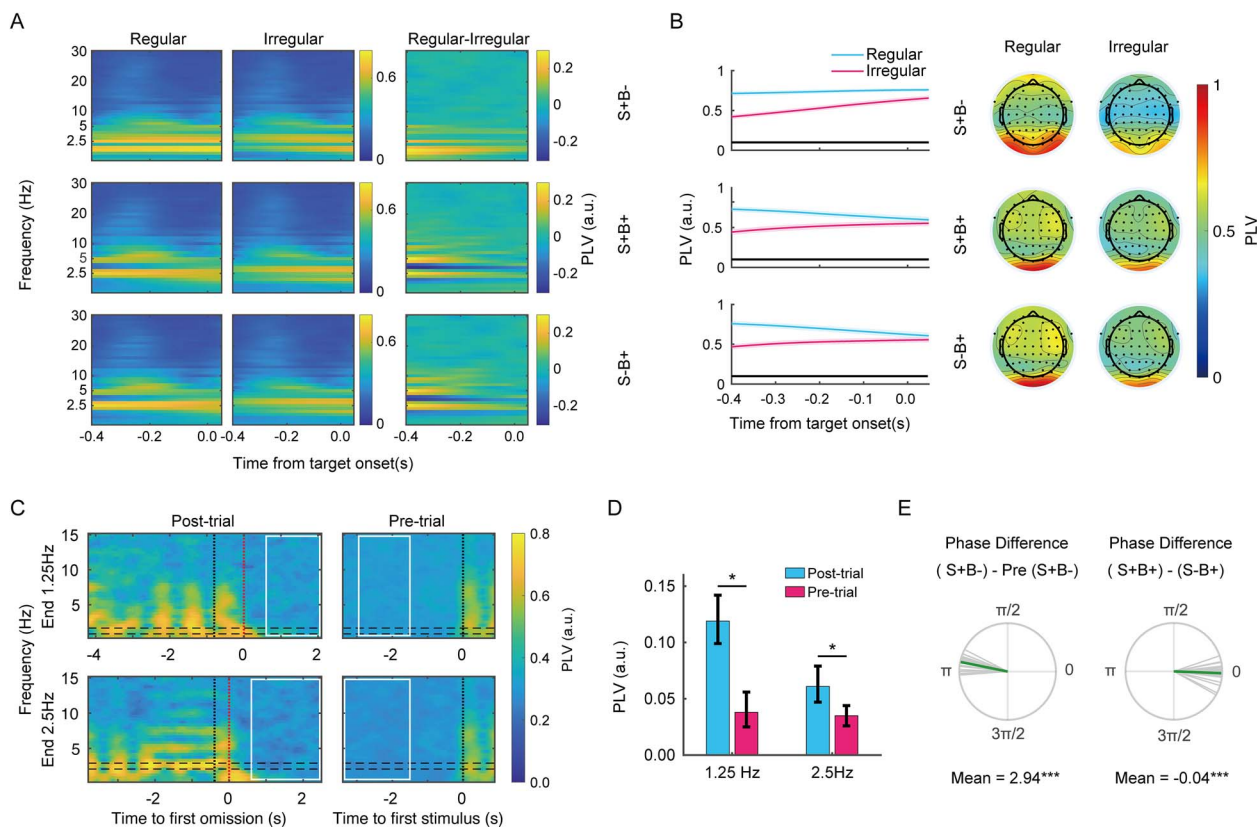


Fig. 2. Neural entrainment over the occipital cortex. (A) PLVs before target onset are shown as a function of frequency and time (left column: Regular conditions, middle column: Irregular conditions, right column: Regular vs. irregular conditions; upper row: S+B-, middle row: S+B+, lower row: S+B+). (B) Left panel: PLVs are shown as a function of the time relative to the target onset in each condition. The black line at the bottom of each graph indicates the time range where the PLV showed a significant difference between regular and irregular conditions, cluster-based permutation corrected at $P < 0.05$. Right panel: The topographical distribution of the PLVs at the target onset. (C) Group-mean PLVs during the posttrial and pretrial periods in the regular condition. Note that during the posttrial period, the zero point (marked by the red dashed line) refers to the “onset” of the first omitted stimulus at the end of the regular stream. The vertical black dashed line in the posttrial period indicates the onset of the last stimulus. The vertical black dashed line in the pretrial period indicates the onset of the first stimulus. The horizontal black dashed lines indicate the two stimuli frequencies (1.25 Hz and 2.5 Hz). White frames indicate the time windows for the comparison of posttrial versus pretrial PLVs. (D) The PLVs are shown as a function of the stimuli frequency and the time period. The error bars indicate the 95% CIs. $*P < 0.05$. (E) Left panel: 1.25 Hz phase difference between S+B- and the previous position of S+B- (pre S+B-). The phase differences were clustered ~ 2.94 (radian), Rayleigh test, $***P < 0.001$. Right panel: 2.5 Hz phase difference between S+B+ and S-B+. The phase differences were clustered around -0.04 (radian), Rayleigh test, $***P < 0.001$. The green bar indicates the mean phase difference across participants. Each gray bar indicates an individual phase difference.

among participants with an analog of *one-sample-t-test* for circular data (*circ_mtest*, *CircStat* toolbox, *MATLAB*).

The extraction of alpha amplitude and alpha phase

For each of the occipital electrodes, the stimulus-locked data (time range: $-4,000$ to $1,500$ ms relative to target onset) was filtered (8–12 Hz) using a two-way least-squares FIR filter (*eegfilt*, *EGLAB*). The filtered data were then transformed with a Hilbert transformation (*hilbert*, *MATLAB*). The alpha phases were extracted from the transformation and averaged over the occipital electrodes.

To obtain alpha amplitude, in each condition, the average of the epoched data was firstly subtracted from the data epochs. This was to remove the evoked components in the data set so that the confounding effect of alpha phase and alpha amplitude can be avoided. Then the induced epoch was filtered with a two-way least-squares FIR filter (8–12 Hz, *eegfilt*, *EGLAB*) and transformed with Hilbert transformation (*hilbert*, *MATLAB*). The transformed amplitudes were averaged across the occipital electrodes.

Analysis of pre-target alpha amplitude

For each of the three positions, the averaged alpha amplitudes across epochs were compared between regular and irregular conditions using a *paired-t-test*. This test was performed for each time point of the -400 to 600 ms interval relative to the target onset with *cluster-based permutation* (number of iterations = 1,000, minimum time duration = 20 ms, cluster-level alpha threshold = 0.05) for correcting multiple comparisons across time points. We chose this time range because it covered both the temporal expectation of the target and the sensorimotor processing of the target.

To exclude that the amplitude effect of prestimulus alpha was confounded by the activity related to button presses, the alpha amplitude in the filler units of the regular condition was also analyzed as a control condition. There was no button press in this no-target condition whereas the temporal regularity was still available to form the expectation. To achieve a fair control condition, the stimulus positions in the no-target condition were aligned with the stimulus positions in the regular condition, and the data was extracted from the time interval of -400 to 600 ms where the 0-ms point aligned with the target onset in the regular condition. This alignment and time interval definition were performed, respectively, for S+B- and S+B+. The same

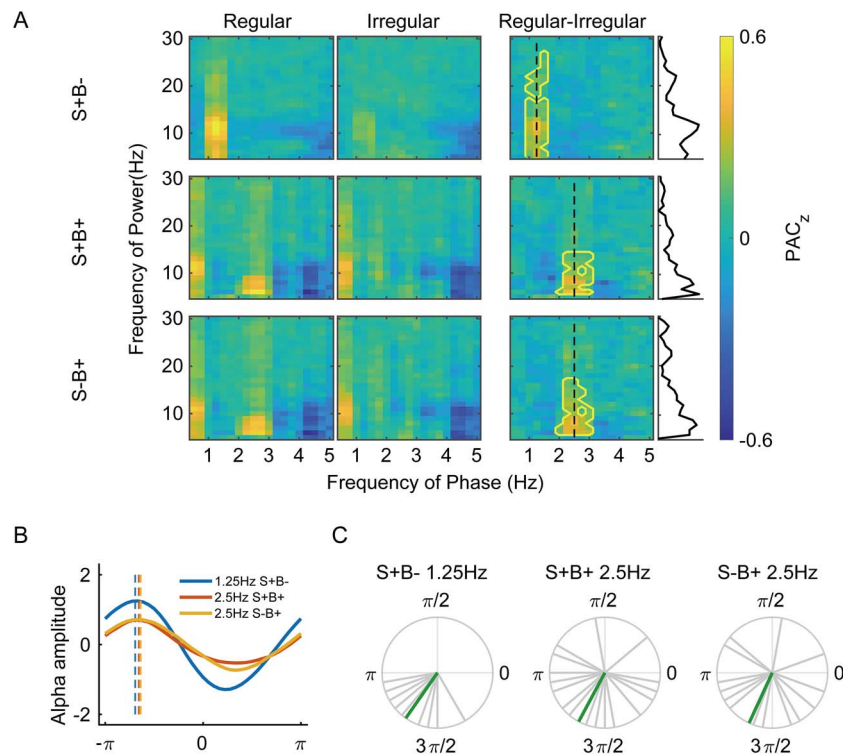


Fig. 3. The coupling between alpha amplitude and the phase of the low-frequency oscillation. (A) The PAC_z are shown as a function of frequencies for phase (0.5–5 Hz) and amplitude (5–30 Hz) (left column: Regular conditions, middle column: Irregular conditions, right column: Regular vs. irregular conditions; upper row: S + B–, middle row: S + B+, lower row: S–B+). The black dashed lines (right column) indicate the stimuli frequencies (1.25 Hz at S + B–, 2.5 Hz at S + B+ and S–B+). The black line on the right of each spectrum (PAC_z difference between regular and irregular conditions) denotes the difference of PAC_z values for the phase of stimuli frequencies. (B) The alpha (8–12 Hz) amplitudes are shown as a function of the phase of the stimuli frequency. The dashed line indicates the peak of alpha amplitude at each position. Note that the alpha amplitudes were transformed to z-scores only for visualization. (D) The low-frequency phase preferred by alpha amplitude at S + B– (1.25 Hz, left panel), S + B+ (2.5 Hz, middle panel) S–B+ (2.5 Hz, right panel). The green bar indicates the mean of the alpha-preferred phase across participants. Each gray bar indicates the alpha-preferred phase for a specific participant.

paired-t-tests as above were also performed to compare the alpha amplitudes between regular and no-target conditions, and between irregular and no-target conditions. Due to the fact that the target at S–B+ was added in extra and this extra position was unavailable in the no-target units, these control analyses were not performed for S–B+.

To examine if the alpha amplitude could be related to the perceptual performance, we tested if the HDDM parameters (Drift rate, Threshold, Nondecision time) could be predicted by the alpha amplitude. For each of the three positions, the amplitude difference between the regular and irregular conditions, with the time ranges where significant alpha amplitude differences were found between regular and irregular conditions, was included in a regression model that was combined with the HDDM model (*HDDMRegressor*, HDDM). By including the regular condition as the intercept (baseline) and the irregular condition as a fixed regressor, the regression model estimated to which extent the three HDDM parameters (drift rate, threshold, nondecision time) difference could be predicted by the amplitude difference between regular and irregular conditions. The regression coefficient was estimated using a hierarchical Bayesian estimation method and the posterior distribution of the regression coefficient was estimated using the Markov chain Monte Carlo techniques (four chains, 10,000 samples for each chain, the first 2,000 samples were discarded). The Gelman-Rubin R-hat statistics were used to assess the convergence of the model parameters. For each of the three positions, statistical testing was conducted by calculating the probability of the regression coefficient in the posterior

distribution. The significance criterion (Type-I error threshold) was set as 0.05. A $P > 0.975$ suggested that the difference of an HDDM parameter could be positively predicted by the amplitude difference, and $P < 0.025$ suggested that an HDDM parameter could be negatively predicted by the amplitude difference. In addition, to assess if the predictability of alpha amplitude on parameters was affected by different positions, the regression coefficients of the subject level at different positions were submitted to a repeated measures of ANOVA.

Analysis of pre-target alpha phase

We also investigated if and how the phase of alpha oscillation was modulated by the different forms of temporal expectation. First, we tested if the phase coherence of the prestimulus alpha oscillation was modulated by the regularity of the steams. The PLVs of the alpha band (8–12 Hz) were calculated using the same method as the low frequencies with the interval of –400 to 600 ms relative to the target onset. For each time point, paired-t-test was used to test the difference of the alpha PLVs between regular and irregular conditions, with *cluster-based permutation* (number of iterations = 1,000, minimum temporal cluster = 20 ms, cluster-level alpha threshold = 0.05) for correcting multiple comparisons across time points. As the amplitude analysis above, the PLVs of the alpha oscillation were also calculated in the no-target condition (i.e. filler units). As a control analysis to exclude the confounding activities related to button presses, the difference of alpha PLVs between the no-target condition and the regular/irregular condition was also tested.

Next, for each of the three positions, we tested if the alpha phase after the stimulus preceding the target (i.e. pre-target alpha phase) was different from the alpha phase after the stimulus two positions before the target (i.e. pre-pre-target alpha phase). The phase difference was conducted in the regular conditions provided that the time intervals were random in the irregular condition. The tests would further disentangle the phase effect contributed by the sequence-based expectation from the phase effect contributed by the rhythm. For the pre-target alpha phase, the 400 ms interval was time-locked to the onset of the pre-target stimulus; for the pre-pre-target alpha phase, the 400 ms interval was time-locked to the pre-pre-target stimulus (i.e. the stimulus two positions before the target). For each time point, the normalized alpha phases (averaged across epochs) were compared with the Watson-Williams test (Watson and Williams 1956; Stephens 1969), a circular analog of the t-test, which tests if the two samples of angles have different phase distributions. Cluster-based permutation correction was used to correct the multiple comparisons across time points.

To investigate the relationship between the alpha phase and the perceptual performance, “optimal” and “non-optimal” alpha phases were differentiated based on the behavioral response, with the expectation that the target onset of trials with “good” responses would be more likely to be localized at an optimal phase (Samaha et al. 2015). For each participant, trials with correct responses and incorrect responses were differentiated, and the correct trials were sorted based on RTs in ascending order. To achieve comparable numbers of trials, the “good trials” were defined as the upper quartiles of the sorted correct trials (number of trials: mean = 109.2, SE = 7.05), and the “bad trials” were defined as the last 1/6 of the sorted correct trials plus the incorrect trials (number of trials: mean = 111.2, SE = 18.19). The alpha phases at the target onset in “good trials” were averaged as the “optimal” phase; the alpha phases at the target onset in “bad trials” were averaged as the “non-optimal” phase. The alpha phase difference between the “good trials” and the “bad trials” was calculated and tested against zero with an analog of one-sample t-test for circular data (circ_mtest, CircStat toolbox, MATLAB). For each epoch, the phase distance was calculated as the difference between the phase at the target onset and the “optimal” phase. To avoid the risk of double-dipping, we did not test the difference in phase distance between different conditions. Instead, we focused on the predictability of the phase distance on the single-trial perceptual performance. The same regression models were fitted to estimate to which extent the HDDM parameters could be predicted by the phase distance.

Results

Accuracy and RTs

A 2 (Regularity: regular vs. irregular) \times 3 (Target position: S+B-, S+B+, vs. S-B+) repeated-measures ANOVA revealed that the accuracy was higher in the regular condition (94.6%) than in the irregular condition (89.4%), $F(1, 23) = 25.74$, $P < 0.001$, $\eta_p^2 = 0.528$ (Fig. 1B, left). The main effect of position was significant, $F(2, 46) = 8.58$, $P < 0.001$, $\eta_p^2 = 0.272$. The accuracy was lower at S+B- (90.4%) than the accuracies at S+B+ (92.9%), $p_{\text{bonferroni}} = 0.007$, and S-B+ (92.7%), $p_{\text{bonferroni}} = 0.023$, whereas the difference between S+B+ and S-B+ did not reach significance, $p_{\text{bonferroni}} > 0.999$. There was an interaction between regularity and target position, $F(2, 46) = 5.55$, $P = 0.007$, $\eta_p^2 = 0.194$. Further tests showed that the increased accuracy by regularity was larger at S+B+ (accuracy difference between regular and irregular conditions: 6.3%) than

that at S+B- (3.0%), $t(23) = 3.17$, $p_{\text{bonferroni}} = 0.013$. The increased accuracy by regularity was larger at S-B+ (6.5%) than that at S+B-, $t(23) = 2.78$, $p_{\text{bonferroni}} = 0.032$. However, the difference between S+B+ and S-B+ did not reach significance, $t < 1$.

The ANOVA on RTs showed that responses were faster in the regular condition (566 ms) than the responses in the irregular condition (601 ms), $F(1, 23) = 15.9$, $P = 0.001$, $\eta_p^2 = 0.408$ (Fig. 1B, right). The main effect of position was also significant, $F(2, 46) = 34.5$, $P < 0.001$, $\eta_p^2 = 0.600$. Pair-wise comparisons showed that responses were faster at S+B+ (572 ms), $p_{\text{bonferroni}} < 0.001$, and S-B+ (578 ms), $p_{\text{bonferroni}} < 0.001$, than responses at S+B- (602 ms), whereas the difference between S+B+ and S-B+ did not reach significance, $p_{\text{bonferroni}} = 0.141$. There was a significant interaction between regularity and position, $F(2, 26) = 19.3$, $P < 0.001$, $\eta_p^2 = 0.456$. Further tests showed that the facilitated response by regularity was larger at S+B+ (RT difference between irregular and regular conditions: 51 ms) than that at S+B- (14 ms), $t(23) = 6.31$, $p_{\text{bonferroni}} < 0.001$, and larger at S-B+ (39 ms) than S+B-, $t(23) = 3.60$, $p_{\text{bonferroni}} = 0.005$. However, the difference between S+B+ and S-B+ did not reach significance, $t(23) = 2.24$, $p_{\text{bonferroni}} = 0.105$. Considering the long tail of the RTs distribution which might have confounded the results, the RT results were verified by log-transforming the RTs. The same pattern of results was observed with the transformed data (Supplementary Fig. S1).

Hierarchical drift-diffusion modeling

We further modeled the behavioral data with the HDDM (Wiecki et al. 2013) to elucidate how the multiple cognitive components were affected by the rhythmic and sequence regularity. The HDDM estimated three parameters: drift rate, decision threshold, and nondesideration time. Specifically, the drift rate quantified the accumulation speed of the sensory evidence; the decision threshold quantified the criteria for the perceptual decision; the nondesideration time quantified the combined time of early stimulus encoding and the late motor implementation of the decision (Wiecki et al. 2013). Convergence was achieved for all of the estimated DDM parameters, $R\text{-hats} < 1.01$.

The 2 \times 3 ANOVA revealed that the drift rate was higher in the regular condition (4.41) than in the irregular condition (3.62), $F(1, 23) = 21.5$, $P < 0.001$, $\eta_p^2 = 0.483$ (Fig. 1C, left). The main effect of position was significant, $F(2, 46) = 18.9$, $P < 0.001$, $\eta_p^2 = 0.451$. Pair-wise comparisons showed that the drift rate was higher at S+B+ (4.75) than the drift rates at S+B- (3.57), $p_{\text{bonferroni}} < 0.001$, and S-B+ (3.72), $p_{\text{bonferroni}} < 0.001$, whereas no significant difference between S+B- and S-B+ was observed, $p_{\text{bonferroni}} = 0.630$. The interaction between regularity and position was significant, $F(2, 46) = 30.6$, $P < 0.001$, $\eta_p^2 = 0.571$. Further tests showed that the enhanced drift rate by regularity was higher at S+B+ (1.25) than S+B- (0.22), $t(23) = 6.82$, $p_{\text{bonferroni}} < 0.001$, and S-B+ (0.92), $t(23) = 2.85$, $p_{\text{bonferroni}} = 0.027$, and the enhancement was also stronger at S-B+ than S+B-, $t(23) = 5.24$, $p_{\text{bonferroni}} < 0.001$.

The ANOVA showed that the decision threshold was higher in the irregular condition (1.66) than in the regular condition (1.39), $F(1, 23) = 17.1$, $P < 0.001$, $\eta_p^2 = 0.426$ (Fig. 1C, middle). The main effect of position was significant, $F(2, 46) = 15.1$, $P < 0.001$, $\eta_p^2 = 0.396$. Pair-wise comparisons showed that the decision threshold was higher at S+B+ (1.85) than the thresholds at S-B+ (1.42), $p_{\text{bonferroni}} = 0.002$, and S+B- (1.31), $p_{\text{bonferroni}} = 0.001$, whereas no significant difference between S+B- and S-B+ was observed, $p_{\text{bonferroni}} > 0.167$. The interaction between regularity and position was significant, $F(2, 46) = 22.8$, $P < 0.001$, $\eta_p^2 = 0.498$. Further tests showed that the lowered threshold by regularity was larger at S+B+ (difference of threshold between irregular

and regular conditions: 0.45) than that at S+B- (0.06), $t(23) = 5.59$, $p_{\text{bonferroni}} < 0.001$, and S-B+ (0.30), $t(23) = 2.67$, $p_{\text{bonferroni}} = 0.041$, and the threshold decrease was larger at S-B+ than S+B-, $t(23) = 5.07$, $p_{\text{bonferroni}} < 0.001$.

The ANOVA on nondecision time showed that the nondecision time was shorter in the regular condition (0.400) than in the irregular condition (0.429), $F(1, 23) = 28.2$, $P < 0.001$, $\eta_p^2 = 0.551$, (Fig. 1C, right). The main effect of position was significant, $F(2, 46) = 7.0$, $P = 0.002$, $\eta_p^2 = 0.233$. Pair-wise comparisons showed that the nondecision time was longer at S+B- (0.425) than that at S+B+ (0.407), $p_{\text{bonferroni}} < 0.014$, whereas the difference between S+B- and S-B+ (0.412), $p_{\text{bonferroni}} = 0.074$, and the difference between S+B+ and S-B+, $p_{\text{bonferroni}} = 0.533$, did not reach significance. The interaction between regularity and position was significant, $F(2, 46) = 25.6$, $P < 0.001$, $\eta_p^2 = 0.527$. Further tests showed that the facilitated nondecision time by regularity was larger at S+B+ (difference of nondecision time between irregular and regular conditions: 0.044), $t(23) = 6.51$, $p_{\text{bonferroni}} < 0.001$, and S-B+ (0.036), $t(23) = 5.98$, $p_{\text{bonferroni}} < 0.001$, than the facilitated nondecision time at S+B- (0.009), whereas the difference between S+B+ and S-B+ did not reach significance, $t(23) = 1.41$, $p_{\text{bonferroni}} = 0.516$.

The low-frequency neural oscillation entrained by the rhythmic stimuli

The phase-locking values (PLVs) covering the frequency of 0.5–30 Hz are shown in Fig. 2A. We have assumed that the rhythmic stimuli would induce low-frequency (i.e. 1.25 Hz for S+B-, 2.5 Hz for S+B+ and S-B+) neural entrainment over the visual cortex. As shown in Fig. 2B, the 1.25 Hz PLVs at S+B-, the 2.5 Hz PLVs at S+B+ and S-B+ were higher in the regular condition than in the irregular condition (*paired-t-test*, $P < 0.05$, with *cluster-based permutation correction*), demonstrating the typical neural entrainment synchronized with the rhythmic stimuli. It should be noted that the stimulus timings (e.g. the intervals both before and after the target) at each of the three positions were kept the same in regular and irregular conditions. Therefore, the enhanced PLVs in the regular condition than the irregular condition cannot be simply due to the phase-resetting effect induced by the preceding stimulus. The same pattern of results was observed when a shorter filter was used to estimate the low-frequency PLVs (two cycles, *Supplementary Fig. S3*), suggesting that the above results cannot be due to confounding effects caused by long filters.

Moreover, as shown in Fig. 2C, the oscillation at the stimulus frequency was still observed after the last stimulus of the regular stream, when no stimulus was presented. Further analysis showed that, for trials ended with the 1.25 Hz stimuli, the 1.25 Hz PLVs during the posttrial period (see Materials and methods and Fig. 2C, upper) were stronger than the PLV during the pretrial period (Fig. 2D, left), posttrial: mean = 0.119, 95% CI = [0.099, 0.142], pretrial: mean = 0.038, 95% CI = [0.025, 0.056]. For trials ended with the 2.5 Hz stimuli, the 2.5 Hz PLVs during the posttrial period (Fig. 2C, bottom) were stronger than the PLVs during the pretrial period (Fig. 2D, right), posttrial: mean = 0.061, CI = [0.047, 0.079], pretrial: mean = 0.035, CI = [0.026, 0.044]. To verify if the higher PLVs in the posttrial interval were specific to the low-frequency rhythmic stimuli, we conducted the same comparison on a control frequency of 25/6 Hz (10 times the frequency resolution). We chose this frequency because it was still in the low-frequency band, and was not an integer multiple of the entrained frequencies (i.e. 1.25 Hz and 2.5 Hz). The results showed no significant difference between posttrial and pretrial intervals. For trials ended with the 1.25 Hz stimuli, posttrial: mean = 0.065, 95% CI = [0.050, 0.078], pretrial: mean = 0.061, 95% CI = [0.045, 0.082].

For trials ended with the 2.5 Hz stimuli, posttrial: mean = 0.059, 95% CI = [0.049, 0.073], pretrial: mean = 0.049, 95% CI = [0.038, 0.063]. Collectively, these results suggested that the neural oscillation sustained even after the termination of the visual stimulation, a pattern consistent with van Bree et al. (2021). Therefore, the stronger PLVs in the regular condition cannot be simply due to the transient responses to the individual stimuli, but rather a result of neural entrainment by the rhythmic stream. However, the rhythmic neural responses during the stimuli presentation cannot be taken as completely unaffected by the visual stimulations.

For the phase comparison, the 1.25 Hz phase of S+B- was at an opposite phase (antiphase) of an optimal phase predicted by the 1.25 Hz neural entrainment, as shown by the phase difference (centered mean = 2.94) between S+B- and its preceding position, $P < 0.001$ (Rayleigh Test, radian, Fig. 2E left). By contrast, the 2.5 Hz phases of S+B+ and S-B+ were both at an optimal phase predicted by the 2.5 Hz neural entrainment. The phases were the same, as shown by the phase difference (centered mean = -0.04) between these two positions, $P < 0.001$ (Rayleigh Test, radian, Fig. 2E right).

The coupling between low-frequency oscillation and alpha activity

To test if there were couplings between the low-frequency and the high-frequency oscillations, we calculated the phase-amplitude coherence (PAC) between the phase of the low-frequency (0.5–5 Hz) oscillations and the amplitude of the high-frequency (5–30 Hz) activities before the target. The regular condition showed stronger PAC_z (Z scored PAC values) than the irregular condition between the phase of 1–1.5 Hz and the amplitude of 6–27 Hz at S+B- (cluster-based permutation corrected at $P < 0.05$, Fig. 3A upper right), stronger PAC_z between the phase of 2.25–2.75 Hz and the amplitude of 7–14 Hz at S+B+ (cluster-based permutation corrected at $P < 0.05$, Fig. 3A middle right), and stronger PAC_z between the phase of 2.25–2.75 Hz and the amplitude of 7–17 Hz at S-B+ (cluster-based permutation corrected at $P < 0.05$, Fig. 3A bottom right). These results suggested that there was a significant cross-frequency coupling between the phase of the low-frequency oscillations (1.25 Hz and 2.5 Hz) and the amplitude of the high-frequency oscillations, which manifested at the alpha band. The periodical characteristics of the neural oscillations at the three positions can be seen in *Supplementary Fig. S2*. The same pattern of results was observed when a shorter filter (two cycles for low frequency, three cycles for high frequency) was used to estimate the PAC_z (*Supplementary Fig. S3*).

As shown in Fig. 3B and C, the alpha-preferred phase at S+B- (1.25 Hz) was not significantly different from the alpha-preferred phase at S+B+ (2.5 Hz, phase difference = -0.113, CI = [-0.868, 0.643]). Similarly, the alpha-preferred phase at S+B- (1.25 Hz) was not significantly different from the alpha-preferred phase at S-B+ (2.5 Hz, phase difference = -0.072, CI = [-0.551, 0.407]). The results suggested that the coupling between the alpha amplitude and the phase of the low-frequency oscillations cannot be simply due to the transient stimulations of the individual stimuli.

Prestimulus alpha amplitude under different structures of temporal expectation

As shown in Fig. 4A, the amplitude of the prestimulus alpha activity over the visual cortex was higher in the regular condition than in the irregular condition at S+B-, with a significant temporal cluster of -400 to -290 ms relative target onset (*cluster-based permutation corrected* at $P < 0.001$). By contrast, the alpha amplitude

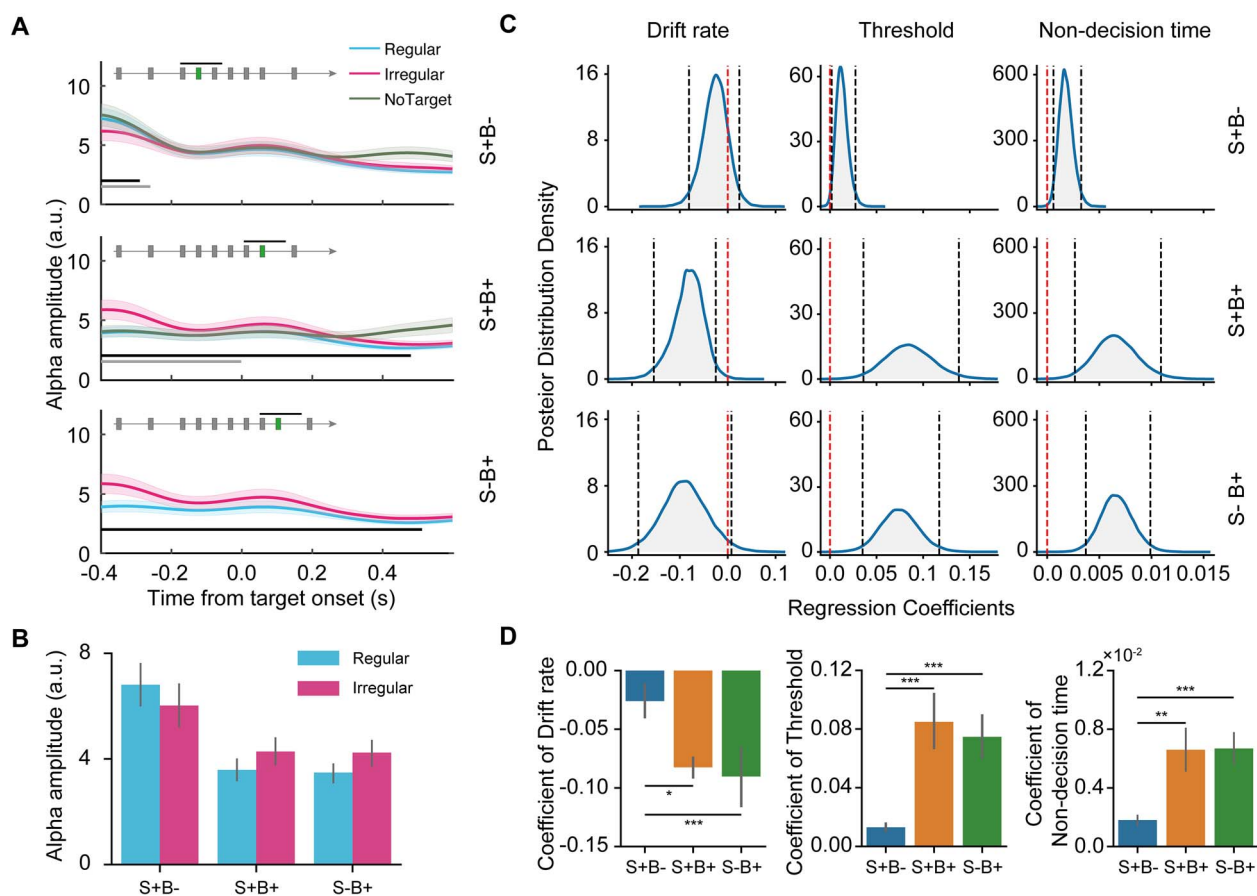


Fig. 4. Results of alpha amplitudes. (A) Alpha amplitudes are shown as a function of time points relative to the target onset for each experimental condition. The black bar at the bottom of each graph indicates a significant difference between regular and irregular conditions; the gray bar indicates a significant difference between no-target and irregular conditions (cluster-based permutation corrected at $P < 0.05$). The shadows denote SEM across participants. The example of the stimulus stream on the top of each graph illustrates the interested temporal locations. (B) Alpha amplitudes averaged over the significant cluster in (A) are shown as a function of the experimental conditions error bars indicate SEM across participants. To avoid “double-dipping,” we did not test the difference between regular and irregular conditions, but rather compared the absolute difference (regular vs. irregular) of amplitudes between positions (see Results). (C) The group-level posterior distribution of the regression coefficient was estimated with the regression model for each target position and each one of the HDDM parameters. The regression model quantified to which extent the HDDM parameters could be predicted by alpha amplitude. The black dashed line denotes a 95% CI, and the red dashed line denotes 0 point. (D) The individual-level regression coefficients (mean value with SEM) estimated with the regression model are shown as a function of the target position. Error bars indicate SEM across participants. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

was lower in the regular condition than in the irregular condition at S+B+ and S-B+ (both $P < 0.001$ with *cluster-based permutation correction*): a significant temporal cluster of -400 to 480 ms at S+B+ and a significant temporal cluster of -400 to 512 ms at S-B+. The same pattern of results was observed when a shorter filter (three cycles) was used to estimate the alpha amplitude (Supplementary Fig. S3).

The reversed pattern of alpha amplitude between S+B- and S+B+/S-B+ conditions suggested that the alpha amplitude was mainly modulated by the phase of the low-frequency oscillation but was not additionally modulated by the sequence-based expectation. To further test this hypothesis, we compared the amplitude difference between regular and irregular conditions at S+B- with the amplitude difference between irregular and regular conditions (i.e. the reverse of the difference between regular and irregular conditions) at S+B+/S-B+. For each position, the alpha amplitudes were extracted from the significant cluster to calculate the mean difference between regular and irregular conditions (Fig. 4B). To avoid making statistical inferences based on nonsignificant P-values, we performed the Bayes Factor analysis to quantify the likelihood of the null hypothesis against the

alternative hypothesis (Keysers et al. 2020). The results showed that the hypothesis “the decreased amplitude at S+B+ was equivalent to the decreased amplitude at S-B+” was 4.46 times more likely to be true than the alternative hypothesis “the decreased amplitude at S+B+ was different from the decreased amplitude at S-B+”. Moreover, the hypothesis “the decreased amplitude at S+B+ was equivalent to the increased amplitude at S+B-” was 3.60 times more likely to be true than the alternative hypothesis “the decreased amplitude at S+B+ was different from the increased amplitude at S+B-”.

In addition, at S+B-, the alpha amplitude was higher in the no-target condition than in the irregular condition, with a significant temporal cluster of -400 to -260 ms relative to target onset (cluster-based permutation corrected at $P < 0.05$), whereas there was no difference between the regular and the no-target condition (Fig. 4A). At S+B+, the alpha amplitude was lower in the no-target condition than in the irregular condition, with a significant temporal cluster of -400 to 194 ms relative to target onset (cluster-based permutation corrected at $P < 0.05$), whereas there was no difference between the regular and the no-target condition (Fig. 4A). These results suggested that the increased/decreased

alpha amplitude in the regular (vs. irregular) condition cannot be related to the button press, as the same pattern was observed in the no-target condition.

To assess if and how the amplitude of pre-target alpha activity could be related to the perceptual performance of the target, we tested if the HDDM parameters (drift rate, threshold, and nondesideration time) could be predicted by the alpha amplitude. For each position and each HDDM parameter, a regression model that combined with the HDDM was constructed to estimate to which extent the amplitude difference between regular and irregular conditions could predict the difference of the HDDM parameter between regular and irregular conditions. Convergence was achieved for all estimated regression coefficients, R -hats < 1.01 . The results of the regression models are shown in Fig. 4C. The drift rate could be negatively predicted by the alpha amplitude at S+B+ (mean regression coefficient = -0.082), P (coefficient > 0) = 0.004 , with the drift rate increasing linearly as the alpha amplitude was decreased by regularity (Fig. 4C, left). The predictability did not reach significance at S+B- (mean regression coefficient = -0.026), P (coefficient > 0) = 0.152 , whereas was marginally significant at S-B+ (mean regression coefficient = -0.09), P (coefficient > 0) = 0.034 (Type-I error threshold of 0.025 given the two directions of correlation, see Materials and methods). In addition, the ANOVA on the coefficients showed a main effect of position, $F(2, 46) = 7.18$, $P = 0.002$, which was due to higher predictability at S+B+ ($p_{\text{bonferroni}} < 0.001$) and S-B+ ($p_{\text{bonferroni}} = 0.012$) than at S+B-. No significant difference was observed between S+B+ and S-B+, $p_{\text{bonferroni}} > 0.999$ (see Fig. 4D, left).

The decision threshold could be positively predicted by the alpha amplitude at all of the three positions: S+B- (mean regression coefficient = 0.013), P (coefficient > 0) = 0.990 , S+B+ (mean regression coefficient = 0.085), P (coefficient > 0) > 0.999 , and S-B+ (mean regression coefficient = 0.075), P (coefficient > 0) > 0.999 (see Fig. 4C, middle), with the decision threshold decreasing linearly as the alpha amplitude was decreased by regularity. In addition, the ANOVA on coefficients showed a main effect of position, $F(2, 46) = 15.0$, $P < 0.001$. The predictability was higher at S+B+, $p_{\text{bonferroni}} = 0.001$ and S-B+, $p_{\text{bonferroni}} < 0.001$ than S+B-. No significant difference was observed between S+B+ and S-B+, $p_{\text{bonferroni}} > 0.999$ (Fig. 4D, middle).

The nondesideration time could be positively predicted by the alpha amplitude at all of the three positions: S+B- (mean regression coefficient = 0.002), P (coefficient > 0) = 0.999 , S+B+ (mean regression coefficient = 0.007), P (coefficient > 0) = 0.999 , and S-B+ (mean regression coefficient = 0.007), P (coefficient > 0) > 0.999 (Fig. 4C, right), with the nondesideration time decreasing linearly as the alpha amplitude was decreased by regularity. In addition, the ANOVA on the coefficients showed a main effect of position, $F(2, 46) = 13.6$, $P < 0.001$. The predictability was higher at S+B+, $p_{\text{bonferroni}} = 0.004$ and S-B+, $p_{\text{bonferroni}} < 0.001$ than S+B-. No significant difference was observed between S+B+ and S-B+, $p_{\text{bonferroni}} > 0.999$ (Fig. 4D, right).

Prestimulus alpha phase under different structures of temporal expectation

At S+B-, higher alpha PLVs were observed in the regular condition than in the irregular condition at the interval of 52–292 ms relative to target onset (cluster-based permutation corrected at $P < 0.05$, Fig. 5A). At S+B+, higher alpha PLVs were observed in the regular condition than in the irregular condition at the interval of -312 to -58 ms relative to target onset (cluster-based permutation corrected at $P < 0.05$), and at the interval of 84–316 ms relative to

the target onset (cluster-based permutation corrected at $P < 0.05$). At S-B+, higher alpha PLVs were observed in the regular condition than in the irregular condition at the interval of -356 to -82 ms relative to the target onset (cluster-based permutation corrected at $P < 0.05$). While the increased alpha PLVs before target onset suggested phase effects of prestimulus alpha for the target, the increased alpha PLVs after target onset suggested phase effects of prestimulus alpha for the stimulus immediately after the target. In addition, at S+B+, higher PLVs were observed in the no-target condition than in the irregular condition at the interval of -226 to -106 ms relative to target onset, suggesting that the observed phase effect cannot be simply due to button presses. The same pattern of results was observed when a shorter filter (three cycles) was used to estimate the alpha phase (Supplementary Fig. S3).

Here the enhanced PLVs of pre-target alpha at both S+B+ and S-B+ but not at S+B- suggested that the alpha phase was critically modulated by rhythm-based expectation. The pre-target phase effect was not observed at S+B-, which might lead to the suggestion that the alpha phase was not modulated by sequence-based expectation. However, this null effect at S+B- could be a result of counteracting phase effects contributed by the sequence-based expectation and the antiphase of the stimulus rhythm. To test if the alpha phase was modulated by sequence-based expectation, we compared the alpha phase during the time interval before the target (pre-target alpha phase, regular condition) with the alpha phase prior to the stimulus immediately before the target (pre-pre-target alpha phase, regular condition) for each of the three positions. During these two intervals, the neural entrainment induced by the rhythmic stimuli was locked to the same low-frequency phase so that any observed difference should be attributed to the sequence-based expectation. To consider the different lengths of the two intervals at S+B-, the comparison was performed on the 400 ms range that was time-locked to the onset of the stimulus before the two intervals (i.e. the onset of the pre-target stimulus for the pre-target alpha phase, the onset of the pre-pre-target stimulus for pre-pre-target phase). At S+B-, the alpha phases showed significant differences during the time interval of 324–400 ms relative to the stimulus onset, $P < 0.001$ (Watson-Williams test, with cluster-based correction) (Fig. 5B). At S+B+, the alpha phases showed significant differences during the time interval of 320–382 ms relative to the stimulus onset, $P < 0.001$ (with cluster-based correction). At S-B+, however, no significant difference in the alpha phase was observed. These results suggested that sequence-based expectation (S+B-, S+B+) of the target changed the pre-target alpha phase.

The alpha phase showed a significant difference between “good trials” and “bad trials,” with a mean phase difference = 0.36 (95% CI = $[0.02, 0.70]$, circ_mttest, CircStat toolbox, MTLAB), confirming a valid differentiation between the optimal and nonoptimal alpha phase. Then the trial-by-trial phase distance relative to the optimal phase modeled to assess to which extent the HDDM parameters could be predicted by the phase distance. Convergence was achieved for all regression coefficients, R -hats < 1.01 . The drift rate could be negatively predicted by the phase distance at S+B+ (mean regression coefficient = -0.307), P (coefficient > 0) < 0.001 , and S-B+ (mean regression coefficient = -0.202), P (coefficient > 0) < 0.001 , (Fig. 4C, left), with the drift rate increasing linearly as the alpha phase was closer to the optimal alpha phase, whereas the predictability at S+B- (mean regression coefficient = -0.139) did not reach significance, p (coefficient > 0) = 0.037 (> 0.025). The ANOVA on the coefficients showed a main effect of position, $F(2, 46) = 28.2$, $P < 0.001$, which

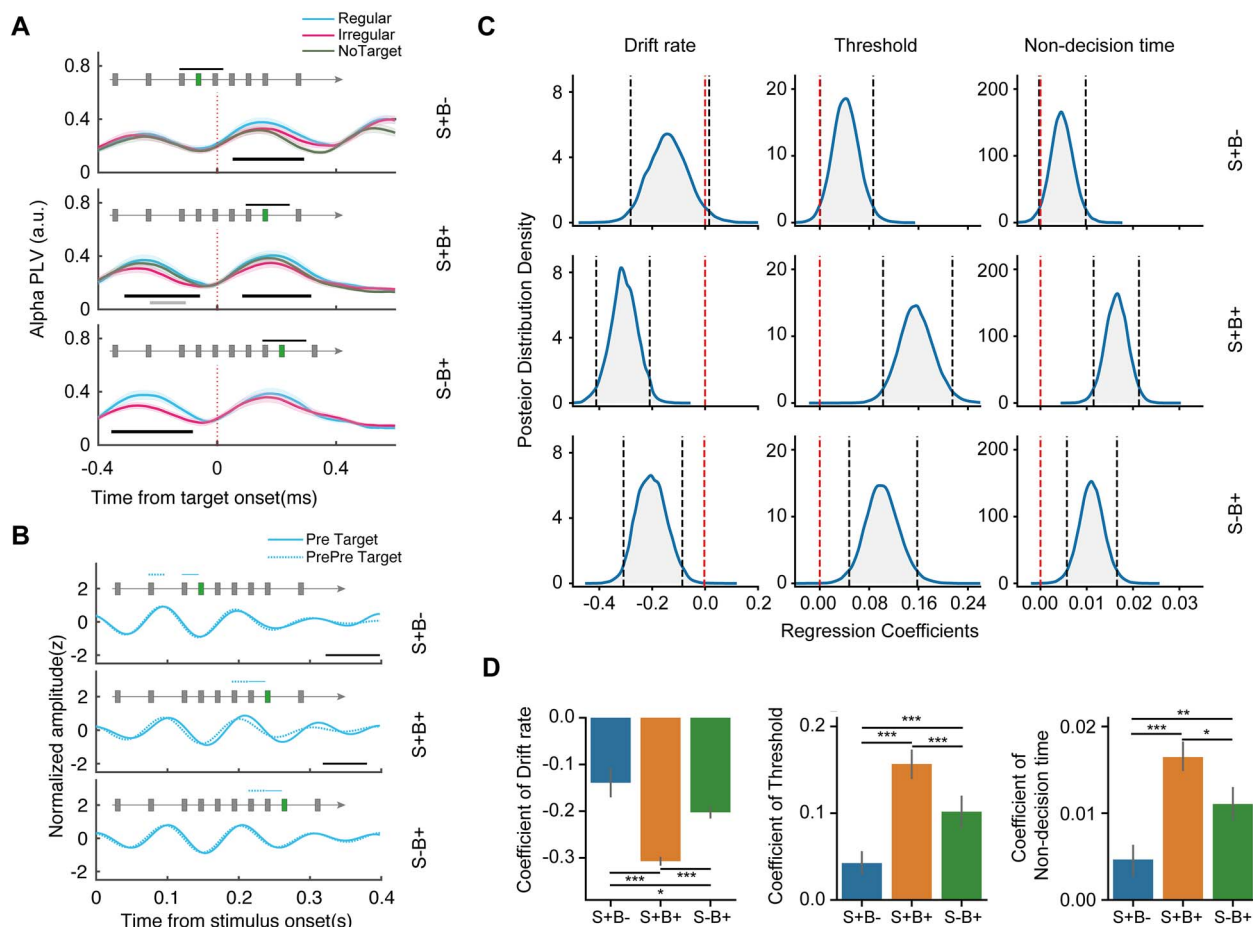


Fig. 5. Results of the alpha phase. (A) PLVs of alpha are shown as a function of time points relative to the target onset for each experimental condition. The example of the stimulus stream on the top of each graph illustrates the range of the time courses. The black bar at the bottom of the graph indicates the temporal cluster with a significant difference between regular and irregular conditions; the gray bar at the bottom of the graph indicates the temporal cluster with a significant difference between the no-target and irregular conditions, cluster-based permutation corrected at $P < 0.05$. (B) Filtered and normalized alpha amplitudes in the regular condition are shown as a function of time. For each target position (S+B-, S+B+, and S-B+), two time courses are shown and were compared: The time-course locked to the onset of the stimulus immediately before the target (pre-target), and the time course locked to the onset of the stimulus two positions before the target (pre-pre-target). The example of the stimulus stream on the top of each graph illustrates the temporal locations of the two time courses. The black bar at the bottom of the graph indicates the temporal cluster with a significant difference between the two time courses, cluster-based permutation corrected at $P < 0.05$. (C) The group-level posterior distribution of the regression coefficient estimated with the regression model for each target position and each one of the HDDM parameters. The regression model quantified to which extent the HDDM parameters could be predicted by the distance between the alpha phase at the target onset and the optimal alpha phase. The black dashed line denotes a 95% CI, and the red dashed line denotes 0 point. (D) The individual-level regression coefficients (mean value with SEM) estimated with the regression model are shown as a function of the target position. Error bars indicate SEM across participants. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

was due to higher predictability at S+B+ ($p_{\text{bonferroni}} < 0.001$) and S-B+ ($P_{\text{bonferroni}} = 0.046$) than S+B-, and also higher predictability at S+B+ than S-B+, $P_{\text{bonferroni}} < 0.001$ (Fig. 5D, left).

The decision threshold could be positively predicted by the phase distance at S+B+ (mean regression coefficient = 0.157), $P(\text{coefficient} > 0) > 0.999$, S-B+, (mean regression coefficient = 0.101), $P(\text{coefficient} > 0) > 0.999$, and S+B- (mean regression coefficient = 0.042), $P(\text{coefficient} > 0) = 0.976$ (see Fig. 5C, middle), with the decision threshold decreasing linearly as the alpha phase was closer to the optimal phase. The ANOVA on the coefficients showed a main effect of position, $F(2, 46) = 46.3$, $P < 0.001$. The predictability was higher at S+B+ than S+B- ($P_{\text{bonferroni}} < 0.001$) and S-B+ ($P_{\text{bonferroni}} < 0.001$) and was also higher at S-B+ than S+B-, $P_{\text{bonferroni}} < 0.001$ (Fig. 5D, middle).

The nondesideration time could be positively predicted by the phase distance at S+B+ (mean regression coefficient = 0.016), $P(\text{coefficient} > 0) > 0.999$, and S-B+ (mean regression coefficient = 0.011), $P(\text{coefficient} > 0) > 0.999$ (see Fig. 5C, right), with the

nondesideration time decreasing linearly as the alpha phase was closer to the optimal phase, but not at S+B- (mean regression coefficient = 0.005), $P(\text{coefficient} > 0) = 0.968$ (< 0.975). In addition, the ANOVA on the coefficients showed a main effect of position, $F(2, 46) = 27.9$, $P < 0.001$. The predictability was higher at S+B+ than S+B- ($p_{\text{bonferroni}} < 0.001$) and S-B+ ($p_{\text{bonferroni}} = 0.012$) and was also higher at S-B+ than S+B-, $p_{\text{bonferroni}} = 0.002$ (Fig. 5D, right).

Discussion

In the present study, we investigated how rhythmic and sequence information can be combined to form temporal expectation and optimize the perceptual performance of the expected stimulus. The behavioral results showed that the rhythm-based and sequence-based expectation had an additive effect in improving the perceptual performance of the expected stimulus. The EEG results suggested a dissociation of alpha amplitude and phase in supporting the perceptual performance under different

structures of temporal expectation. While the amplitude of prestimulus alpha oscillation was modulated by the rhythmic information through PAC, the phase of prestimulus alpha oscillation was affected by both the rhythmic and sequence information. At the single-trial level, on the one hand, the rhythm-based expectation improved the perceptual performance by reducing the prestimulus alpha amplitude, whereas the sequence-based did not further affect the alpha amplitude and the corresponding perceptual performance. On the other hand, the rhythm-based and the sequence-based expectation had an additive effect in biasing the alpha phase toward the optimal phase to improve the perceptual performance.

Mounting evidence has demonstrated that perceptual performance can be improved by the temporal expectation of different sources such as rhythm (Rohenkohl and Nobre 2011; Cravo et al. 2013), sequence (Morillon et al. 2016) and temporal cues (Samaha et al. 2015). In previous studies, however, different forms of temporal expectation were often treated separately, with a single form in a specific context. In studies where two forms of temporal expectation were orthogonally manipulated (Breska and Deouell 2014; Bouwer et al. 2020), the findings revealed both common and different mechanisms underlying different temporal expectations. In an extension of previous findings, here we showed that rhythm-based and sequence-based expectations can be combined to have additive effects in optimizing the perceptual performance of the expected stimulus. Moreover, our HDDM results showed that the rhythmic and sequence information had an additive effect on the perceptual sensitivity to the expected stimulus by both increasing the accumulation of sensory evidence and alleviating the threshold for the perceptual decision. The nondesired time of the perceptual decision, however, benefited more from the rhythm-based expectation, which might be due to more efficient motor implementation as highlighted by the link between the motor system and the rhythmic processing (Morillon et al. 2014; Cannon and Patel 2021). These results suggested that the various temporal information in the environment was utilized to affect perceptual decision-making.

It has been well-documented that brain oscillations on multiple time scales are coordinated to modulate sensorimotor processing (Schroeder and Lakatos 2009; Palva and Palva 2018). It has been suggested that the momentary power of high-frequency oscillation is determined by the phase of the low-frequency oscillation through cross-frequency couplings (Lakatos et al. 2008). In accordance with this notion, our results showed a dependence of the high-frequency amplitude on the phase of the low-frequency neural oscillations, and this dependence manifested in the alpha band. Specifically, the prestimulus alpha amplitude fluctuated with the phase change of the low-frequency neural oscillation, and this fluctuating pattern cannot be simply due to the transient neural responses to the individual stimuli. A corresponding benefit is that a stimulus aligned with the high-excitability phase gains improved perceptual processing, whereas a cost is that a stimulus aligned with the low-excitability is less likely to be efficiently recognized (Lakatos et al. 2008). This prediction was supported by our results that the prestimulus alpha amplitude was reduced when the rhythmic stimulus was at an optimal phase of the low-frequency entrainment (onbeat according to preceding rhythm, i.e. S+B+ and S-B+), and the single-trial perceptual decision-making was critically predicted by the alpha amplitude. However, when the rhythmic stimulus was at an antiphase of the optimal phase (offbeat according to preceding rhythm, i.e. S+B-), the alpha amplitude was otherwise increased, indicating a

lowered preparing state for the stimulus. Our results elucidated a potential mechanism that the rhythmic regularity in the environment affected the perceptual performance through the cross-frequency coupling between the phase of the low-frequency oscillations and the amplitude of prestimulus alpha oscillations.

While both the amplitude reduction (Rohenkohl and Nobre 2011; van Diepen et al. 2015; Breska and Deouell 2017) and the shifted phase (Busch et al. 2009; Samaha et al. 2015) of prestimulus alpha oscillation have been suggested as neural mechanisms of temporal expectation, an important finding here is the dissociated functions of alpha amplitude and alpha phase in the different structures of temporal expectation. The current results showed that the alpha amplitude was mainly driven by the rhythmic information, whereas the sequence information did not additively modulate the alpha amplitude. This notion is supported by the following evidence: (i) at an optimal phase of the rhythmic stimuli, the decreased alpha amplitude was equivalent regardless of the presence of the sequence-based expectation (S+B+ vs. S-B+); (ii) the decreased alpha amplitude at an optimal phase (S+B+) was equivalent to the increased alpha amplitude at antiphase (S+B-) of the rhythmic stimuli; (iii) the predictive power of the alpha amplitude on the single-trial perceptual performance was not additionally contributed by the sequence-based expectation (S+B+ vs. S-B+). In contrast, the alpha phase was modulated by both rhythmic and sequence information, leading to combined effects in biasing the alpha oscillation toward an optimal phase where the perceptual performance can be optimized. Specifically, on top of the rhythmic information, the addition of sequence information induced a change in the phase of the pre-target alpha oscillation (i.e. at both S+B- and S+B+), whereas such phase change was not observed when the sequence-based expectation was absent (i.e. at S-B+). And the phase of the pre-target alpha oscillation was also affected by whether the target was at an optimal phase of rhythmic stimuli (i.e. enhanced pre-target alpha PLVs at S+B+ but not at S+B-). At the single-trial level, the perceptual performance was predicted by the extent to which the phase was close to the optimal phase of the alpha oscillation, and the combination of the rhythmic and sequence information rendered the highest predictive power.

Although the alpha amplitude was involved only in the rhythm-based expectation here, it should not be generalized into that alpha amplitude was immune to the sequence-based expectation regardless of task context. Instead, the suggestion based on the current findings is that the alpha amplitude and alpha phase were flexibly coordinated to take effect according to the task. The rhythmic processing has been suggested as a default mode of the brain, the high and low excitability of which can be reset on multiple time scales (Jones et al. 2006; Schroeder and Lakatos 2009). Due to this dominant role of the rhythmic processing and that a high proportion of the stimuli in the regular stream was aligned with the high-excitability phase (i.e. onbeat stimuli), it would be economic to have a fixed relationship between alpha amplitude and the phase of the entrained low-frequency oscillation, with the alpha phase being flexibly regulated to achieve efficient processing. Here the alpha phase was shifted both by the phase-resetting of the rhythmic stimuli, and by the expectation acquired from the repeated sequence. The phase shift regulated by sequence-based expectation can add to improving the processing of the onbeat target (i.e. S+B+), and more importantly, overcome the lowered processing of the offbeat target (i.e. S+B-). Similar adjustment by top-down temporal expectation in compensating for the processing of the offbeat

targets was also shown in [Breska and Deouell \(2016\)](#). When there was a high probability of offbeat targets such that the onset of the offbeat target can be predicted, the contingent negative variation, an event-related potential component of expectation ([Walter et al. 1964](#)), was adjusted toward the expected time of the offbeat target ([Breska and Deouell 2016](#)). The collaborative effects of alpha phase and alpha amplitude in modulating perceptual performance are also consistent with the recent findings that whether alpha phase modulated perceptual performance was critically dependent on the alpha amplitude ([Fakche et al. 2022](#)). Taken together, these results not only helped to settle the debate concerning the alpha oscillation in temporal expectation ([van Diepen et al. 2015](#)), but also suggested that the various temporal information was flexibly utilized to achieve the task goal through the coordination of distinct neural processes.

One might argue that both the effects of alpha amplitude can be simply explained by whether the target was at the onbeat or offbeat position without necessarily introducing the neural entrainment. It should be noted that, however, the onbeat/offbeat relationship was not against or exclusively alternative to the neural oscillations. Specifically, the rhythmic stimuli sequence can lead to neural entrainment at the frequency of the rhythmic stimuli ([Calderone et al. 2014](#); [Haegens and Golumbic 2018](#)). Consistent with this notion, the low-frequency oscillations were observed even after the termination of the visual stimulations. Moreover, the alpha-preferred low-frequency phase did not differ among the three conditions, suggesting that the periodic alpha amplitudes cannot be simply due to the transient responses to the visual stimulations. These results consistently suggested that the alpha amplitudes were, at least partially, modulated by the phase of the low-frequency oscillations. Another point raised here is that the effects observed at S+B- were due to a rhythm change between 1.25 Hz and 2.5 Hz. On one hand, the global rhythm change was not specific to S+B- but was present across all three conditions, as the two kinds of rhythmic stimuli alternated with each other. The inclusion of an offbeat position, on the other hand, necessarily introduced a local rhythm change from the preceding 1.25 Hz to the current 2.5 Hz. Therefore, the current explanations and suggestions were not against or alternative to either the onbeat/offbeat relationship or a local rhythm change.

A nested structure of regularity is fundamental in daily-life activities such as speech and music ([Koelsch et al. 2013](#); [Ding et al. 2015](#)). Our findings shed light on the dynamic interaction of neural oscillations in exploiting a nested structure of temporal regularity. This echoes previous studies of different cognitive contexts ([Palva and Palva 2018](#)). For instance, [Yuan et al. \(2021\)](#) showed that the multiscale rhythmic information in a stimulus stream was used to alleviate the attentional blink ([Raymond et al. 1992](#)), a cognitive “bottleneck” in visual attention, and this behavioral change was critically predicted by the PAC between the multiscale neural oscillations that correspond to the temporal structure of the stimuli ([Yuan et al. 2021](#)). Beyond visual perception, the cross-frequency interaction was also found as crucial for working memory ([Axmacher et al. 2010](#)) and speech segmentation ([Gross et al. 2013](#)). Our results showed that the dynamic interaction of brain oscillations is not only expressed as the cross-frequency coupling (e.g. the PAC), but also as the amplitude-phase coordination within a specific frequency range (e.g. the alpha oscillation). Taken together, these findings suggested that the multiscale brain oscillations are flexibly organized to deal with the complex environment and empower adaptive behavior.

Acknowledgments

We thank an anonymous reviewer concerning the analysis of the alpha-preferred phase at an earlier stage of the manuscript.

Author contributions

Zhongbin Su (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing—original draft, Writing—review & editing), Xiaolin Zhou (Funding acquisition, Supervision, Writing—review & editing), and Lihui Wang (Conceptualization, Funding acquisition, Supervision, Validation, Writing—original draft, Writing—review & editing).

Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

Funding

This study is supported by the National Natural Science Foundation of China (32000779, 31861133012), the Shanghai Sailing Program (20YF1422100), and a Shanghai Jiao Tong University Grant (YG2022QN108).

Conflict of interest statement: None declared.

Data and code availability

Data and codes have been deposited at OSF, accession code: osf.io/h5dww/.

References

- Anderson B, Sheinberg DL. Effects of temporal context and temporal expectancy on neural activity in inferior temporal cortex. *Neuropsychologia*. 2008;46(4):947–957.
- Axmacher N, Henseler MM, Jensen O, Weinreich I, Elger CE, Fell J. Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proc Natl Acad Sci USA*. 2010;107(7):3228–3233.
- Bouwer FL, Honing H, Slagter HA. Beat-based and memory-based temporal expectations in rhythm: similar perceptual effects, different underlying mechanisms. *J Cogn Neurosci*. 2020;32(7):1221–1241.
- Brainard DH. The psychophysics toolbox. *Spat Vis*. 1997;10(4):433–436.
- Breska A, Deouell LY. Automatic bias of temporal expectations following temporally regular input independently of high-level temporal expectation. *J Cogn Neurosci*. 2014;26(7):1555–1571.
- Breska A, Deouell LY. When synchronizing to rhythms is not a good thing: modulations of preparatory and post-target neural activity when shifting attention away from on-beat times of a distracting rhythm. *J Neurosci*. 2016;36(27):7154–7166.
- Breska A, Deouell LY. Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLoS Biol*. 2017;15(2):1–30.
- Busch NA, Dubois J, Vanrullen R. The phase of ongoing EEG oscillations predicts visual perception. *J Neurosci*. 2009;29(24):7869–7876.

- Calderone DJ, Lakatos P, Butler PD, Castellanos FX. Entrainment of neural oscillations as a modifiable substrate of attention. *Trends Cogn Sci*. 2014;18(6):300–309.
- Cannon JJ, Patel AD. How beat perception co-opts motor neurophysiology. *Trends Cogn Sci* 2021;25(2):137–150.
- Cohen MX. *Analyzing neural time series data: theory and practice*. Cambridge (MA): MIT press; 2014
- Correa Á, Cona G, Arbula S, Vallesi A, Bisiacchi P. Neural dissociation of automatic and controlled temporal preparation by transcranial magnetic stimulation. *Neuropsychologia*. 2014;65:131–136.
- Cravo M, Rohenkohl G, Wyart V, Nobre AC. Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *J Neurosci*. 2013;33(9):4002–4010.
- de Lange FP, Heilbron M, Kok P. How do expectations shape perception? *Trends Cogn Sci*. 2018;22(9):764–779.
- Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J Neurosci Methods*. 2004;134(1):9–21.
- Ding N, Melloni L, Zhang H, Tian X, Poeppel D. Cortical tracking of hierarchical linguistic structures in connected speech. *Nat Neurosci*. 2015;19(1):158–164.
- Drisdelle BL, Aubin S, Jolicoeur P. Dealing with ocular artifacts on lateralized ERPs in studies of visual-spatial attention and memory: ICA correction versus epoch rejection. *Psychophysiology*. 2017;54(1):83–99.
- Fakche C, Vanrullen R, Marque P. Dugué I. α phase-amplitude Trade-offs predict visual perception. *eNeuro*. 2022;9(1):ENEURO.0244–ENEU21.2022.
- Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Stat Sci*. 1992;7(4):457–472.
- Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, Garrod S. Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol* 2013;11(12):e1001752.
- Haegens S, Golumbic EZ. Rhythmic facilitation of sensory processing: a critical review. *Neurosci Biobehav Rev*. 2018;86:150–165.
- Jones MR. Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychol Rev*. 1976;83(5):323–355.
- Jones MR. Attending to sound patterns and the role of entrainment. *Attention Time*. 2010;317:330.
- Jones MR, Johnston HM, Puente J. Effects of auditory pattern structure on anticipatory and reactive attending. *Cogn Psychol*. 2006;53(1):59–96.
- Kaernbach C. Simple adaptive testing with the weighted up-down method. *Percept Psychophys*. 1991;49(3):227–229.
- Keysers C, Gazzola V, Wagenmakers EJ. Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nat Neurosci*. 2020;23(7):788–799.
- Klimesch W. Alpha-band oscillations, attention, and controlled access to stored information. *Trends Cogn Sci*. 2012;16(12):606–617.
- Koelsch S, Rohrmeier M, Torrecuso R, Jentschke S. Processing of hierarchical syntactic structure in music. *Proc Natl Acad Sci USA*. 2013;110(38):15443–15448.
- Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE. Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*. 2008;320(5872):110–113.
- Luce RD. *Response times: their role in inferring elementary mental organization*. Oxford: Oxford University Press on Demand; 1986
- Morillon B, Schroeder CE, Wyart V. Motor contributions to the temporal precision of auditory attention. *Nat Commun*. 2014;5(1):1–9.
- Morillon B, Schroeder CE, Wyart V, Arnal LH. Temporal prediction in lieu of periodic stimulation. *J Neurosci*. 2016;36(8):2342–2347.
- Nobre AC, van Ede F. Anticipated moments: temporal structure in attention. *Nat Rev Neurosci*. 2018;19(1):34–48.
- Palva S, Palva JM. Roles of brain criticality and multiscale oscillations in temporal predictions for sensorimotor processing. *Trends Neurosci*. 2018;41(10):729–743.
- Ratcliff R, Smith PL, Brown SD, McKoon G. Diffusion decision model: current issues and history. *Trends Cogn Sci*. 2016;20(4):260–281.
- Raymond JE, Shapiro KL, Arnell KM. Temporary suppression of visual processing in an RSVP task: an attentional blink? *J Exp Psychol Hum Percept Perform*. 1992;18(3):849–860.
- Rohenkohl G, Nobre AC. Alpha oscillations related to anticipatory attention follow temporal expectations. *J Neurosci*. 2011;31(40):14076–14084.
- Rohenkohl G, Cravo AM, Wyart V, Nobre AC. Temporal expectation improves the quality of sensory information. *J Neurosci*. 2012;32(24):8424–8428.
- Romei V, Gross J, Thut G. On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: correlation or causation? *J Neurosci*. 2010;30(25):8692–8697.
- Samaha J, Bauer P, Cimaroli S, Postle BR. Top-down control of the phase of alpha-band oscillations as a mechanism for temporal prediction. *Proc Natl Acad Sci USA*. 2015;112(46):8439–8444.
- Samaha J, Iemi L, Haegens S, Busch NA. Spontaneous brain oscillations and perceptual decision-making. *Trends Cogn Sci*. 2020;24(8):639–653.
- Sanabria D, Capizzi M, Correa A. Rhythms that speed you up. *J Exp Psychol Hum Percept Perform*. 2011;37(1):236–244.
- Schroeder CE, Lakatos P. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci*. 2009;32(1):9–18.
- Stafford T, Pirrone A, Croucher M, Krystalli A. Quantifying the benefits of using decision models with response time and accuracy data. *Behav Res Methods*. 2020;52(5):2142–2155.
- Stefanics G, Hangya B, Hernádi I, Winkler I, Lakatos P, Ulbert I. Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *J Neurosci*. 2010;30(41):13578–13585.
- Stephens MA. Multi-sample tests for the fisher distribution for directions. *Biometrika*. 1969;56(1):169–181.
- Sun P, Landy MS. A two-stage process model of sensory discrimination: an alternative to drift-diffusion. *J Neurosci*. 2016;36(44):11259–11274.
- Tavares G, Perona P, Rangel A. The attentional drift diffusion model of simple perceptual decision-making. *Front Neurosci*. 2017;11(August):1–16.
- van Bree S, Sohoglu E, Davis MH, Zoefel B. Sustained neural rhythms reveal endogenous oscillations supporting speech perception. *PLoS Biol*. 2021;19(2):e3001142.
- van Diepen RM, Cohen MX, Denys D, Mazaheri A. Attention and temporal expectations modulate power, not phase, of ongoing alpha oscillations. *J Cogn Neurosci*. 2015;27(8):1573–1586.
- Walter WG, Cooper R, Aldridge VJ, Mccallum WC, Winter AL. Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. *Nature*. 1964;203(4943):380–384.
- Watson GS, Williams EJ. On the construction of significance tests on the circle and the sphere. *Biometrika*. 1956;43(3/4):344–352.
- Wiecki TV, Sofer I, Frank MJ. HDDM: hierarchical Bayesian estimation of the drift-diffusion model in python. *Front Neuroinform*. 2013;7:14.
- Yuan P, Hu R, Zhang X, Wang Y, Jiang Y. Cortical entrainment to hierarchical contextual rhythms recomposes dynamic attending in visual perception. *elife*. 2021;10:1–21.