Rhythmic Sampling and Competition of Target and Distractor in a Motion Detection Task

Changhao Xiong[1*], Nathan M. Petro[2*], Ke Bo[3], Lihan Cui[1], Andreas Keil[4**], Mingzhou Ding[1**]

1. J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL 32611

2 Boys Town National Research Hospital, Boys Town, NE 68010

3. Department of Psychological and Brain Science, Dartmouth College, Hanover, NH 03755

4. Department of Psychology, University of Florida, Gainesville, FL 32611

*Authors contributed equally.

**Correspondence: mding@bme.ufl.edu and akeil@ufl.edu

Abstract

It has been suggested that the visual system samples attended information rhythmically. Does rhythmic sampling also apply to distracting information? How do attended information and distracting information compete temporally for neural representations? We recorded electroencephalography from participants who detected instances of coherent motion in a random dot kinematogram (RDK; the target stimulus), overlayed on different categories (pleasant, neutral, and unpleasant) of affective images from the International Affective System (IAPS) (the distractor). The moving dots were flickered at 4.29 Hz whereas the IAPS pictures were flickered at 6 Hz. The time course of spectral power at 4.29 Hz (dot response) was taken to index the temporal dynamics of target processing. The spatial pattern of the power at 6 Hz was similarly extracted and subjected to a MVPA decoding analysis to index the temporal dynamics of processing pleasant, neutral, or unpleasant distractor pictures. We found that (1) both target processing and distractor processing exhibited rhythmicity at ~1 Hz and (2) the phase difference between the two rhythmic time courses were related to task performance, i.e., relative phase closer to π predicted a higher rate of coherent motion detection whereas relative phase closer to 0 predicted a lower rate of coherent motion detection. These results suggest that (1) in a target-distractor scenario, both attended and distracting information were sampled rhythmically and (2) the more target sampling and distractor sampling were separated in time within a sampling cycle, the less distraction effects were observed, both at the neural and the behavioral level.

Keywords: Attention, rhythmic sampling; distractor; EEG; multivariate pattern analysis; SSVEP

Introduction

Sustained visual attention is required in many real-life scenarios such as driving a vehicle or operating machinery. It is characterized by limited capacity, in which not all information available to the visual system can be processed in-depth. Recent work has suggested that to manage limited capacity, the visual system samples the attended information in a rhythmic fashion, mediated by low-frequency intrinsic brain oscillations (Chota et al., 2022; Dugué et al., 2015; Fiebelkorn et al., 2013, 2018; Fiebelkorn & Kastner, 2019; Helfrich et al., 2018; Michel et al., 2022; Re et al., 2019; VanRullen, 2013; Zalta et al., 2020). In this view, the cycle of an intrinsic brain oscillation can be divided into two phases: a high excitability phase and a low excitability phase. When a stimulus occurs during the high excitability phase, behavioral performance tends to be better than average; conversely, if the stimulus occurs during the low excitability phase, performance is generally worse than average (Lakatos et al., 2008; VanRullen, 2013). Behavioral performance may thus exhibit rhythmic fluctuations at the frequency of the aforementioned intrinsic brain oscillation. One paradigm that has been used to test the idea of rhythmic visual sampling is the cue-target paradigm (Posner, 1980; Posner et al., 1987, 1988). The cue at the beginning of each trial, in addition to providing instructions on how the impending target stimulus should be responded to, helps to reset the phase of the low-frequency oscillation such that all the trials start at approximately the same phase. By varying the stimulus onset asynchrony (SOA) between the cue and the target, one obtains the behavioral response (e.g., accuracy and/or reaction time) as a function of the SOA. The rhythmic nature and the frequency of this function can then be assessed by applying time-domain and/or spectral-domain analysis.

When attending to one object in isolation, the frequency of rhythmic sampling tends to be in the high theta or low alpha frequency range, i.e., around 8 Hz (Fiebelkorn et al., 2013; Senoussi

et al., 2019; Van Der Werf et al., 2023). It has also been argued that when attention is directed to multiple objects in the environment, rather than sampling all the objects simultaneously, the brain appears to sample the objects in a serial fashion (Cohen et al., 1990; Wyart et al., 2012). This would then lead to a slower rhythmic sampling of any given object, in the low range of the theta frequency band, i.e., around 4 Hz (Thigpen et al., 2019). For example, when participants were cued to attend one visual hemifield but were asked to detect the appearance of a weak stimulus in either the cued or the un-cued visual hemifield, the rhythmic detection rate for the target appearing in a given visual hemifield decreased from 8 Hz to 4 Hz (Chota et al., 2022; Fiebelkorn et al., 2013; VanRullen, 2013). Interestingly, when the detection rate functions of the cued and uncued targets were compared, a 180 degree relative phase was apparent, suggesting that the visual system indeed sampled the two visual hemifields in a serial, alternating fashion (Fiebelkorn et al., 2013; Jiang et al., 2024). In another example, two spatially overlapping clouds of moving dots, one in red color and the other in blue color, moved in orthogonal directions (Re et al., 2019), and the participant was cued to attend both the red dots and the blue dots and instructed to report the change in either the red dots or the blue dots as soon as it occurred. When there was only one cloud of moving dots, the detection accuracy exhibited rhythmic fluctuations as a function of the SOA at a frequency around 8 Hz. When both clouds of moving dots were present, rhythmic fluctuations in the accuracy of detecting changes in a given cloud of moving dots were again identified, and the sampling frequency was reduced to 4 Hz. In this case, however, no apparent 180 degree relative phase between the rhythmic behavioral response functions to the red dots and to the blue dots was found, suggesting that there was no serial, alternating sampling between the two attended objects if they appeared at the same spatial location.

The real world visual environment contains both task-relevant information (target) and task-irrelevant (distractor) information. It is well established that in the presence of a distractor, the processing of the target is negatively impacted, leading to reduced task performance (Lavie, 2005; Murphy et al., 2016). This implies that the distractor, despite the need for it to be suppressed by the brain's executive control system (Kastner et al., 1998, 1999; Kastner & Pinsk, 2004; Seidl et al., 2012; Ungerleider, 2000), is nevertheless processed in the brain, and the competition between the target and the distractor at the neural representational level causes the detriment in behavioral performance. Does the rhythmic sampling theory extend to the target-distractor scenario? If so, what is the temporal relationship between the rhythmic sampling of the two types of stimuli? These questions have hitherto not been addressed. Part of the reason is that the majority of the studies on rhythmic environmental sampling focuses on behavioral evidence, e.g., rhythmicity in the aforementioned performance-vs-SOA function (Fiebelkorn & Kastner, 2019; Landau & Fries, 2012). Since the distractor is not responded to, its sampling by the visual system cannot be inferred purely on the basis of response behavior, and consequently, it is also not possible to study how the target and the distractor might compete for neural representations pure behaviorally.

In this study we addressed these limitations by investigating rhythmic sampling during a target-distractor scenario using steady-state visual evoked potential (SSVEP) frequency tagging. The stimuli were a cloud of randomly moving dots (the target) superimposed on emotional images from the International Affective Picture System (IAPS; Lang et al., 1997) (the distractor). The target and the distractor were flickered at two different frequencies for an extended duration of ~12 seconds. The participants were asked to focus on the randomly moving dots and report the number of times the dots moved coherently. In this paradigm, the onset of the stimulus array is the

event that resets the phase of the putative low-frequency brain oscillation underlying rhythmic sampling, and the time from the stimulus array onset, referred to as time-from-onset (TFO), is analogous to the SOA in the traditional cue-target paradigm. It is worth noting that, although this paradigm has been extensively used in studies of target-distractor competition with electroencephalography (Hindi Attar & Müller, 2012; Muller et al., 2008), it has not yet been examined in the context of rhythmic sampling. Aided by frequency tagging, from the EEG data, we extracted neural representations of target and distractor processing separately as a function of TFO. By examining the rhythmicity of these representations as functions of TFO and the phase relationship between these functions, we assessed (1) whether the target and the distractor were sampled rhythmically and (2) how their temporal competition for neural representations impacted behavioral performance.

Materials and Methods

Participants

The experimental protocol was approved by the Institutional Review Board of the University of Florida. Thirty undergraduate students from the University of Florida gave written informed consent and participated in the experiment to earn credit in an introductory psychology course. Because the EEG data were recorded inside the MRI scanner (simultaneous EEG-fMRI), participants underwent screening for ferromagnetic implants, claustrophobia, and personal or family history of epilepsy or photic seizures. Female participants were also administered a pregnancy test before participation. Three participants were excluded due to excessive movements during recording. The EEG data from n=27 participants (18 women, 9 men, mean age = 19.2 ± 1.1 years) were analyzed and reported here.

Stimuli

The stimulus comprised a random-dot kinematogram (RDK) overlaid on affective images selected from the International Affective Picture System (IAPS) database. The RDK consisted of 175 yellow dots randomly distributed within a circular aperture in the center of the screen, with each dot spanning <0.5 degrees of visual angle. The IAPS images portrayed three broad categories of emotions: pleasant, neutral, and unpleasant. They were similar in overall composition and rated complexity and matched in picture file size to minimize confounds across categories. The stimulus was presented on a 30-inch MR-compatible LCD monitor placed approximately 230 cm from the participant's head outside the bore of the MRI scanner. A white fixation dot was displayed at the center of the screen throughout the experiment.

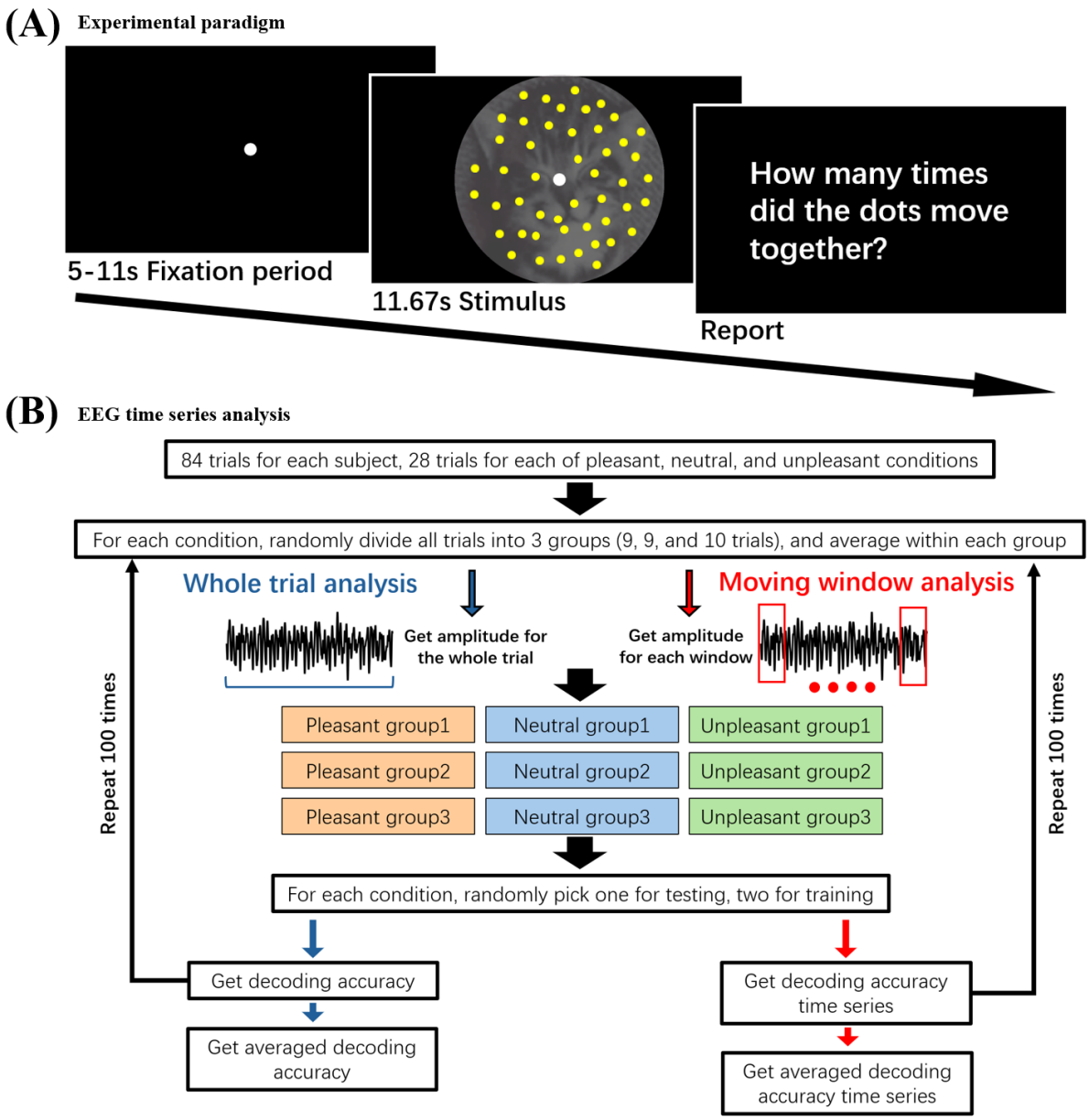

Figure 1. Experimental paradigm and general approach for EEG data analysis. (A) Motion detection task. (B) EEG time series were subject to (1) whole trial analysis and (2) moving window analysis.

Procedure

See Figure 1(A) for the schematic illustration of the experimental task. After 5 to 11 seconds of fixation, the participant was presented the compound stimulus array consisting of the

randomly moving dots (RDK) superimposed on the IAPS pictures for a duration of 11.667 seconds. The moving dots and the background pictures were flickered on and off at 4.29 Hz and 6 Hz, respectively. For each 4.29 Hz flicker cycle, the moving dots were displayed for 100ms, which was followed by a 133ms off period. Similarly, for each 6 Hz flicker cycle, the IAPS background picture was shown for 100ms and followed by a 66.7ms off period. During each on-off cycle, the moving dots in the RDK were randomly displaced by 0.3 degrees of visual angle in either random directions or one coherent direction. Coherent motion instances lasted for four on-off cycles (933 ms) and appeared once in 39 trials (13 trials per emotion condition) or twice in 4 trials. The remaining 41 trials contained no instances of coherent motion. Each trial lasted 11.667 seconds (50 moving dots cycles and 70 IAPS background picture cycles). The coherent motion instances occurred in the interval between 2.3 seconds and 10.4 seconds post stimulus array onset. The participant was asked to fixate on the central white dot during the trial, to monitor the motion coherence of the random dots, and report the number of coherent motion instances at the end of the trial. Both the number of coherent motion instances and the underlying emotion category of IAPS image were randomized in each trial. A total 42 IAPS pictures were equally divided into three content categories based on valence: pleasant (erotic couples), neutral (workplace people), or unpleasant (bodily mutilation). Depending on the emotion category of the IAPS picture used in a given trial, the trials are referred to as pleasant, neutral, and unpleasant trials. There was a total of 84 trials: 28 pleasant trials, 28 unpleasant trials, and 28 neutral trials, and each picture was used twice during the experiment.

EEG data collection and preprocessing

EEG data was recorded using a 32-channel MR-compatible EEG recording system (Brain Products, Germany). The system was synchronized to the internal clock of the scanner to facilitate the subsequent scanner noise removal. Thirty-one Ag/AgCl electrodes were located on the scalp according to the 10–20 system via an elastic cap. One additional electrode was located on the participant's upper back to record the electrocardiogram (ECG). Electrode FCz was used as the reference during recording. Impedances were kept below 20kΩ for all scalp electrodes and below 50kΩ for the ECG electrode, as suggested by the manufacturer. EEG data was digitized at 16-bit resolution and sampled at 5kHz with a 0.1- 250 Hz (3dB-point) bandpass filter applied online

(Butterworth, 18 dB/octave roll off). The digitized data was transferred to a laptop computer via a fiber-optic cable.

Artifact removal from electroencephalogram (EEG) data, specifically magnetic gradient and cardioballistic artifacts, was conducted using the Brain Vision Analyzer 2.0 software (Brain Products GmbH). The elimination of magnetic gradient artifacts was based on an algorithm initially proposed by (Allen et al., 2000). The process involves the creation of an artifact template through averaging EEG data over 41 consecutive fMRI volumes, which was subsequently subtracted from the EEG recordings. Additionally, cardioballistic artifacts were removed by employing a technique developed by (Allen et al., 1998), in which R peaks were detected via the EKG electrode, and a corrective template were computed from 21 successive heart beats and subtracted from the EEG data.

Subsequent to scanner artifact removal, data was downsampled to 500 Hz and exported into EEGLab software (Delorme & Makeig, 2004). The data underwent further filtering using a 0.1 to 40 Hz band-pass Butterworth filter. Independent Components Analysis (ICA) was applied to remove components associated with eye blinks, horizontal eye movements, and residual cardioballistic artifacts. The data were then converted to the average reference.

EEG Data analysis

Overview: According to the task design, the IAPS pictures were behaviorally irrelevant and thus the distractor to be ignored, while the moving dots were behaviorally relevant and thus the target to be attended. To minimize the transient effect resulting from the stimulus array onset and the possible effect resulting from anticipating the end of a trial, the EEG data from the beginning and the end of a trial were discarded, namely, the analyzed EEG data came from the period from 2 to 11 seconds post array onset, which contained the period from 2.3 to 10.4 seconds post stimulus array onset during which instances of coherent motion in the moving dots took place.

Quantifying target processing: The moving dots were flickered at 4.29 Hz. For a given type of emotion trials (i.e., pleasant, neutral, or unpleasant), the SSVEP was computed by averaging all the trials within the type. Filtering the SSVEP between 4.29 - 0.5 Hz and 4.29 + 0.5 Hz yielded the data specific for target processing. Obtaining the magnitude of the band-pass filtered data at the whole trial level allowed the assessment of the overall strength of target

processing; see Figure 1(B). To assess target processing as a function of time-from-onset (TFO), the magnitude of the band-passed filtered data was obtained using a moving window approach, where the window duration was 0.5 second and the step size was 0.25 second. See Figure 1(B) for illustration.

Quantifying distractor processing: The IAPS pictures were flickered at 6 Hz. Band-pass filtering the EEG data between 6 - 0.5 Hz and 6 + 0.5 Hz resulted in signals that were specific to distractor processing. Following a recent study where we showed that the emotion category of IAPS pictures can be decoded from scalp EEG data using the MVPA method (Bo et al., 2022), we assessed distractor processing using a MVPA decoding approach at the whole trial level as well as at the level of moving windows. The MVPA analysis was conducted with the linear support vector machine (SVM) method as implemented in the LibSVM package (http://www.csie.ntu.edu.tw/~cjlin/libsvm/) (Chang & Lin, 2011). The decoding was between two types of emotion trials (e.g., pleasant vs neutral) or between all three types of emotion trials based on an one-vs-all strategy. Above-chance decoding accuracy (50% for pairwise decoding and 33.3% for three-way decoding) is taken to indicate distractor processing in the brain with higher decoding accuracy indicating stronger distractor processing. For both the whole trial and moving window analysis, the trials from each of the three different emotional categories were divided into three subsets of trials randomly. We averaged the trials within each subset to yield the subset SSVEP. For the whole trial analysis, we calculated the 6 Hz SSVEP over the whole trial, whereas for the moving window analysis, the 6 Hz SSVEP amplitude was obtained for each 0.5-second analysis window. For the decoding strategy, the SSVEP amplitude from the two subsets within each emotion category served as training data for constructing a support vector machine classifier, while the SSVEP amplitude from the third subset was used as testing data for calculating decoding accuracy. This process was iterated 100 times to ensure the stability of the decoding result and the average of the decoding accuracy values was analyzed and reported (Bae & Luck, 2019; Haxby et al., 2014; Zhang et al., 2024). See Figure 1(B) for an illustration of the method.

Quantifying the relation between target and distractor processing: To investigate the temporal relationship between target processing and distractor processing, we calculated the phase relation between the target amplitude time series from the moving window approach which quantified the temporal fluctuation of the strength of target processing and the distractor decoding

accuracy time series which quantified the temporal fluctuation of the strength of distractor processing. To investigate the effect of temporal competition between target processing and distractor processing, we correlated the relative phase relation between the target processing time series and the distractor processing time series with behavioral performance.

Results

Behavioral analysis

The compound stimulus array consisted of a cloud of moving dots (RDK) flickered at 4.29 Hz superimposed on IAPS pictures flickered at 6 Hz. The participant attended the moving dots, ignored the IAPS pictures, and reported the number of instances of coherent motion in the moving dots at the end of the trial (0, 1 or 2). The overall detection accuracy was 55.73% ± 2.94%, with that for pleasant, neutral, and unpleasant trials being 55.67% ± 2.76%, 55.03% ± 3.10%, and 56.48% ± 3.61%, respectively. A one-way ANOVA found no significant difference in behavioral performance between the three types of trials ($F_{2, 78} = 0.053$, $p = 0.949$), suggesting that the three types of distractors exerted similar distracting influence on the detection of coherent motion, irrespective of their emotional significance.

SSVEP analysis at the whole trial level

The grand average SSVEP at Oz and its Fourier spectrum are shown in Figure 2(A) and 2(B). From Figure 2(B), spectral peaks corresponding to the flicker frequencies of 4.29 Hz (target) and 6 Hz (distractor) are clearly seen. Averaging target power and distractor power across all electrodes, the 4.29 Hz amplitude was significantly greater than the 6 Hz amplitude ($p=2.6 \times 10^{-4}$); see Figure 2(C). SSVEP amplitude topographies for target and

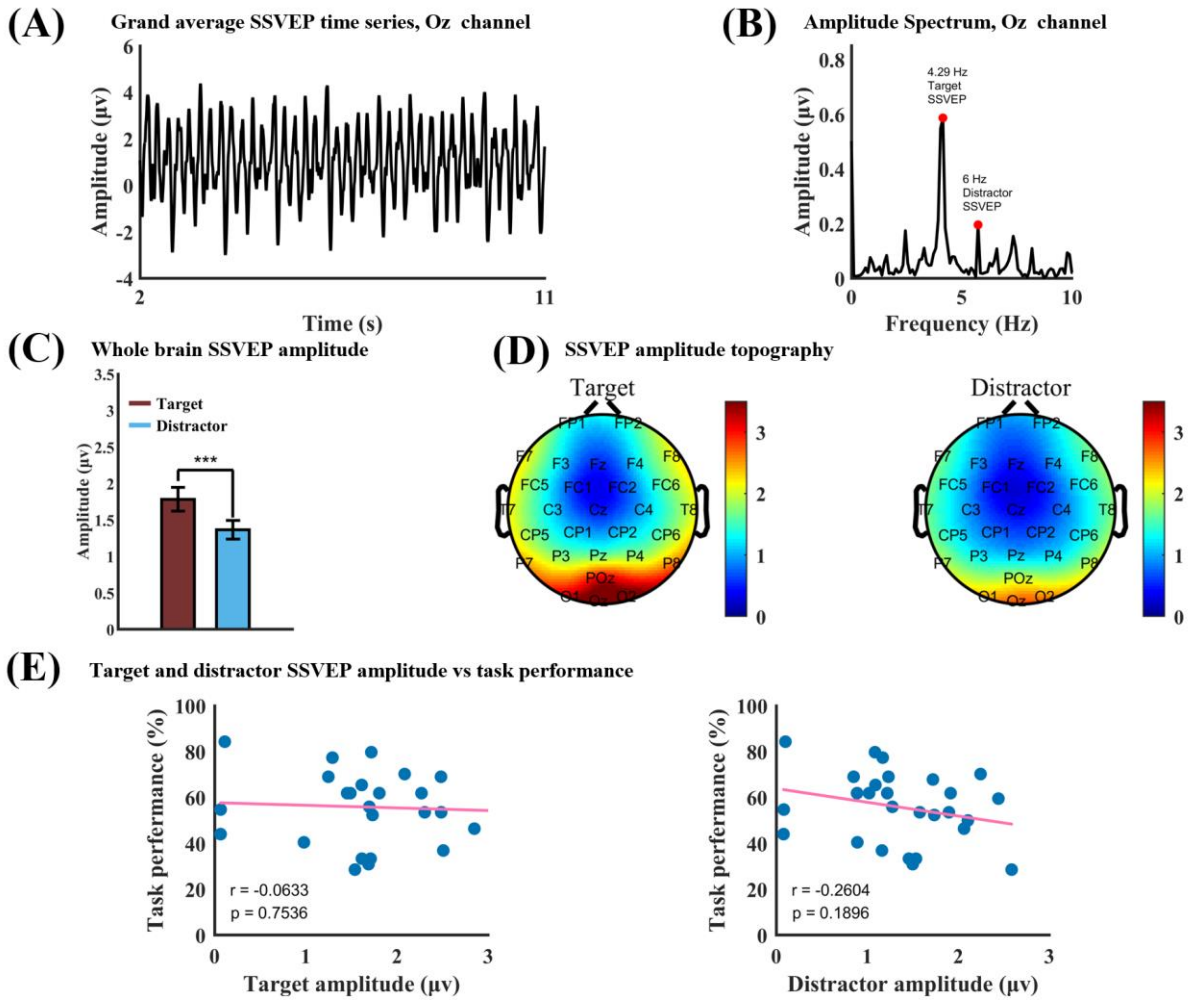

Figure 2. SSVEP analysis at the whole trial level. (A) Grand average SSVEP at Oz. (B) Fourier spectrum of the data in Figure 2(A). (C) Target amplitude across all electrodes is significantly larger than distractor amplitude at $p = 2.6 \times 10^{-4}$. (D) Topographical distributions of target and distractor amplitude. (E) Correlation between target SSVEP amplitude and task performance (left) and between distractor SSVEP amplitude and task performance (right). Both correlation values are not significant.

distractor in Figure 2(D) showed that the strongest response for both frequencies was concentrated in the occipital channels. In Figure 2(E) we assessed the relation between SSVEP amplitude and task performance. Across participants, there was no correlation between target SSVEP amplitude and task performance (p=0.7536); see Figure 2(E) left. The SSVEP amplitude of the distractor has a slight negative correlation with task performance, indicating that the stronger the distractor processing, the worse the performance, but it is not statistically significant (p=0.1896).

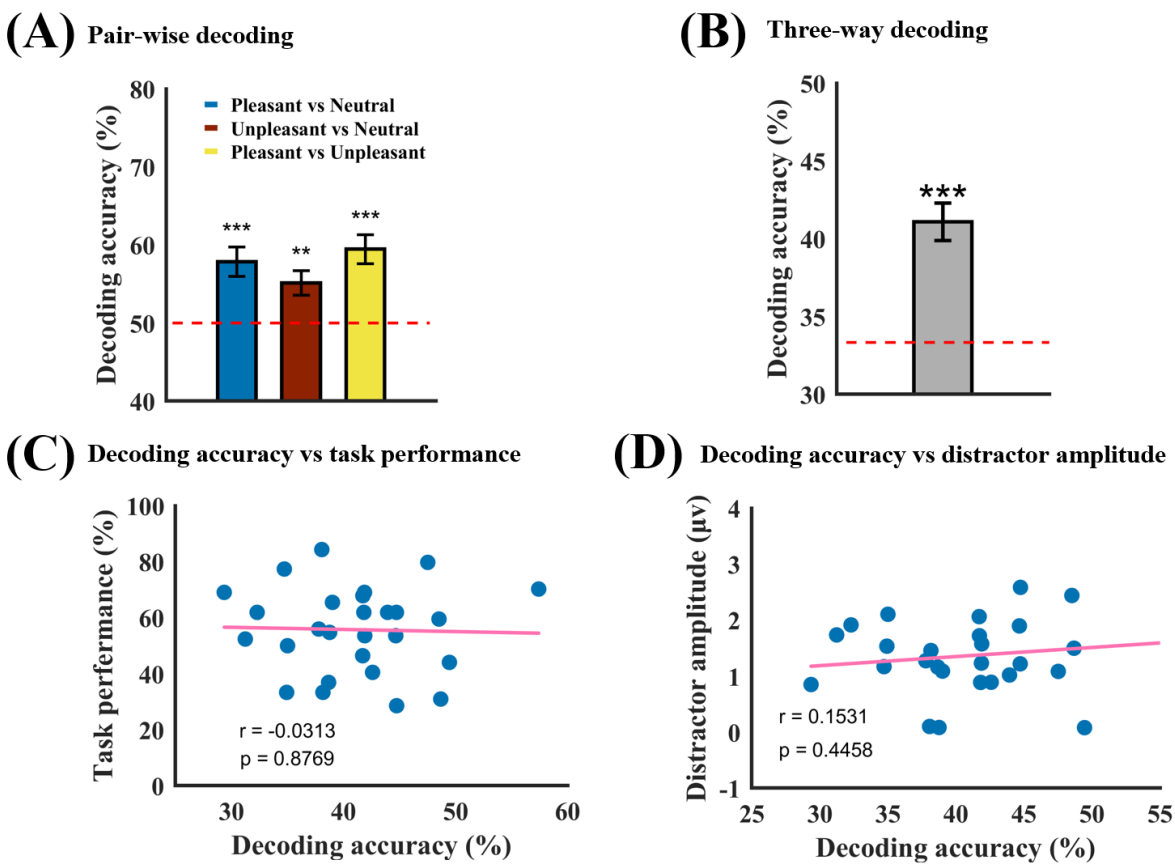

Figure 3. MVPA decoding analysis of distractor processing at the whole trial level. (A) Pair-wise decoding accuracies between pleasant vs neutral, unpleasant vs neutral, and pleasant vs unpleasant are 57.86% ± 9.86%, 55.14% ± 8.17%, and 59.45% ± 9.73%, respectively, which are all significantly above chance level of 50% (red dashed line) at $p=3.2 \times 10^{-4}$, $p=3.0 \times 10^{-3}$, and $p=3.0 \times 10^{-5}$). (B) Three-way decoding accuracy is 41.09% ± 6.25% which is significantly higher than the chance level of 33% (red dashed line) at $p=3.9 \times 10^{-7}$. (C) Decoding accuracy vs task performance. The correlation of r = -0.0313 (p = 0.8769) is not significant. (D) Distractor decoding accuracy vs distractor SSVEP amplitude. The correlation of r = 0.1531 (p = 0.4458) is not significant.

MVPA analysis at the whole trial level

Our previous work has shown that IAPS pictures from different emotion categories evoke distinct spatial patterns in EEG which can be decoded using machine-learning based MVPA methods (Bo et al., 2022). If we were able to decode the emotion categories of the distractor from the spatial patterns of the 6 Hz SSVEP amplitude, the decoding accuracy can then be used to indicate the strength of the distractor representation in the brain, complementing the 6 Hz SSVEP amplitude considered earlier. The decoding was done between different types of emotion trials

(e.g., pleasant vs neutral) using an "ERP" decoding method (Bae & Luck, 2019). See Methods for more details. As shown in Figure 3(A), for pleasant vs neutral, unpleasant vs neutral, and unpleasant vs neutral, the pairwise decoding accuracy was 57.86% ± 9.86%, 55.14% ± 8.17%, and 59.45% ± 9.73%, respectively, which were all significantly above the chance level of 50% at $p=3.2 \times 10^{-4}$, $p=3.0 \times 10^{-3}$, and $p=3.0 \times 10^{-5}$. As shown in Figure 3(B), the three-way decoding accuracy was found to be 41.09% ± 6.25%, which is again significantly above chance level of 33.33% at $p=3.9 \times 10^{-7}$. Similar to distractor SSVEP amplitude, no correlation was found between distractor decoding accuracy and task performance; see Figure 3(C). Also, in order to verify that the distractor decoding accuracy and the distractor amplitude were independent indices of distractor processing, we correlated the two across participants. As Figure 3(D) shows, no correlation was found, suggesting that the two quantities provided complementary characterization of distractor processing (also see Figure S4).

Moving window analysis of target and distractor processing

To examine the temporal dynamics of target processing, the target SSVEP amplitude time series was obtained using the moving window approach, where the window duration was 0.5 seconds and the step size 0.25 seconds. Fourier analysis was then applied to assess the rhythmicity of the time series. The results of these analyses for one representative participant are shown in Figure 4(A)(i). The rhythmic nature of target processing is apparent with a spectral peak at ~1 Hz. Across all participants the averaged Fourier spectrum is shown in Figure 4(A)(ii) where the frequency of the spectral peak was found to be 1.08 ± 0.11 Hz. These results supported the idea that the attended target was sampled rhythmically with a sampling frequency at ~1 Hz (delta frequency band).

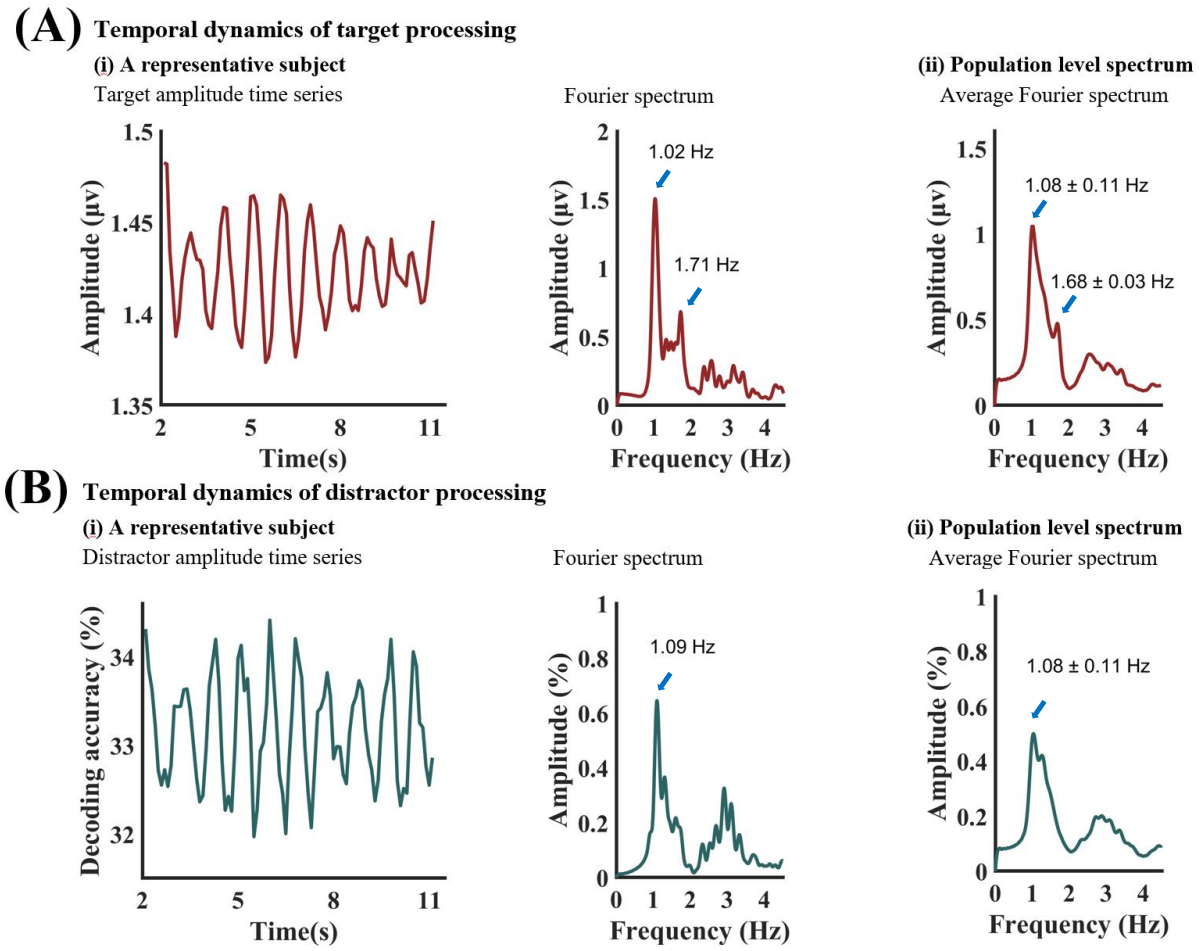

Figure 4. Temporal dynamics of target and distractor processing. (A) (i): Target amplitude time series from the moving window approach for a representative subject (left) and its Fourier spectrum (right). (A) (ii): The average target amplitude spectrum across 27 subjects. (B) (i): Distractor decoding accuracy time series from the moving window approach for a representative subject (left) and its Fourier spectrum (right). (B) (ii): The average distractor decoding accuracy spectrum across 27 subjects.

To examine the temporal dynamics of distractor processing, three-way MVPA decoding was performed for the three types of emotion trials using the moving window approach. The three-way decoding accuracy time series and the Fourier spectrum from one representative participant are shown in Figure 4(B)(i). The rhythmic nature of the decoding accuracy time series is again apparent and the spectral peak is at ~1 Hz. Across all participants the averaged spectrum is shown in Figure 4(B)(ii) where the peak frequency was determined to be 1.08 ± 0.11 Hz. These results supported the idea that the distractor was also sampled rhythmically with a sampling frequency at ~1 Hz (delta frequency band).

Target-distractor competition and task performance

As discussed above, the present evidence suggests that both the target and the distractor were sampled rhythmically. Since the sampling frequency was approximately the same for the two rhythmic time series, the relative phase between them can then be assessed, which characterizes the temporal relationship between the sampling of target and distractor. Figure 5(A) shows the distribution of the relative phase for all participants (mean relative phase = 0.51 ± 0.31π). A Kolmogorov-Smirnov test was applied to the relative phase distribution to see whether it departed from the uniform distribution. A K-S static of 0.10 showed that the relative phase distribution is not different from the uniform distribution at p=0.92, suggesting that there was no systematic relative phase between rhythmic samplings of target vs distractor across participants.

Since simultaneously presented target and distractor compete for neural representations and the stronger the competition the worse the task performance, one may expect that if the target sampling and the distractor sampling are well separated in time, namely, if they occur in opposite phases of the 1 Hz brain oscillation, the competition will be minimized, and the task performance will be maximized. Conversely, if the target and the distractor were sampled during the same phase within the 1 Hz cycle, the target-distractor competition will be maximized, and the task performance will be minimized. This notion is illustrated in Figure 5(C), using data from a high-performing participant (accuracy=84.34%) and a low-performing participant (accuracy=33.33%): Time series are shown for the target and distractor decoding accuracy time series from each participant. In the high performer, the two time courses are highly anti-correlated (relative phase is around π), indicating that the target and the distractor were sampled in opposite parts of the cycle, while for the low performer, the two time courses are highly correlated (relative phase is around 0), indicating that the target and the distractor were sampled in the same part of the cycle. Across all participants, as shown in Figure 5(D), a significant positive correlation between relative phase and task performance was observed (r = 0.6041, p = 0.0008), suggesting that the more the target sampling and the distractor sampling are separated in time (i.e., in opposite phases of the cycle), the less the interference between target and distractor processing, the better the task performance.

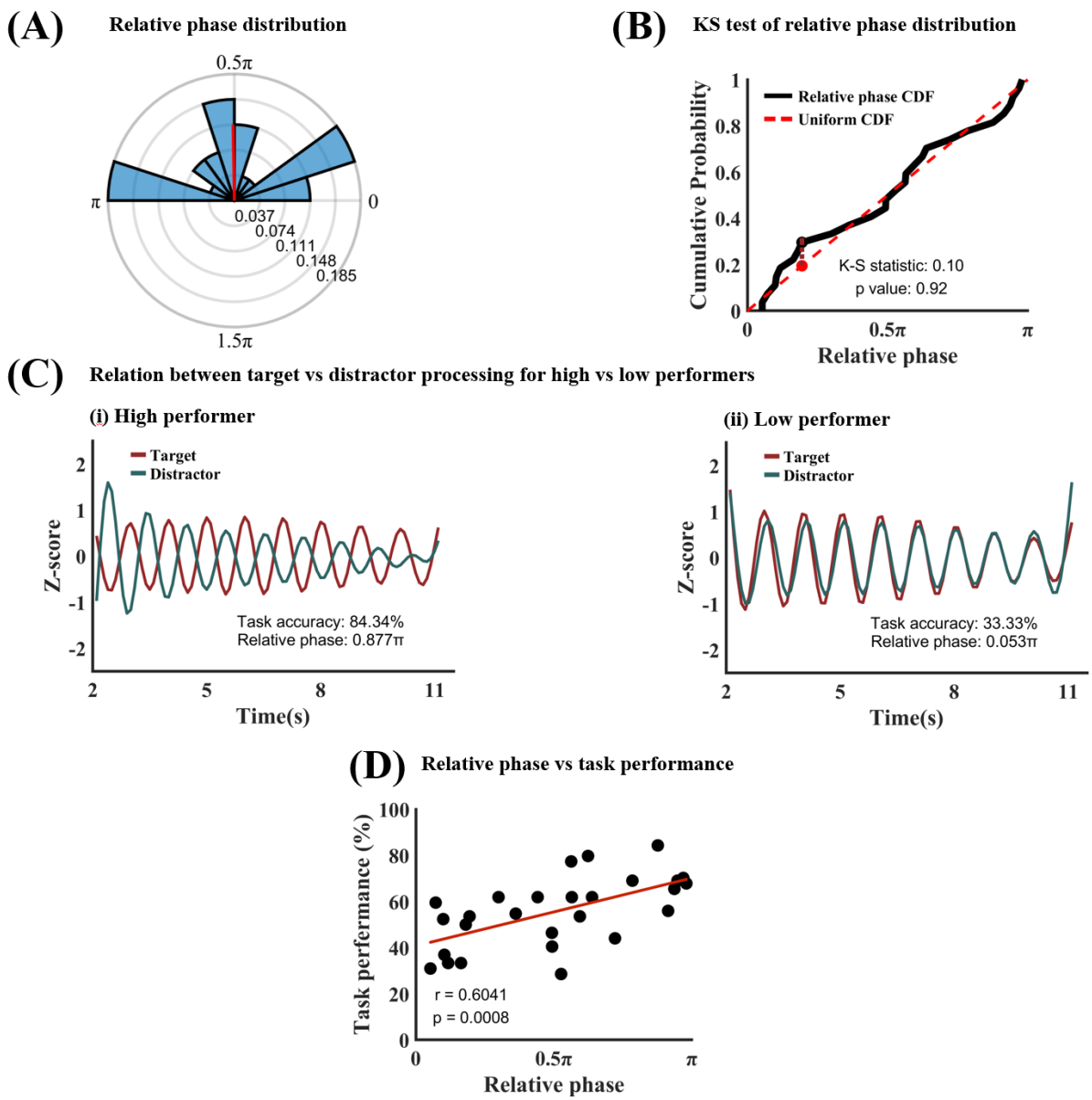

Figure 5. Target-distractor competition analysis. (A) Phase polar histogram for the relative phase between target process time series and distractor processing time series (1 Hz). The average relative phase is 0.51π. (B) Kolmogorov-Smirnov test showed that the relative phase distribution is not different from uniform distribution. (C) Temporal relation between target processing and distractor processing for (i) a high performer (accuracy=83.84%; relative phase=0.877π) and (ii) a low performer (accuracy=33.33%; relative phase=0.053π). (D) Task performance vs 1 Hz relative phase. The significant positive correlation (r=0.6041, p=0.0008) means that the more separated the target and distractor sampling within the 1 Hz oscillation cycle the better the behavioral performance. CDF: Cumulative distribution function.

Discussion

In natural vision, task-relevant information (the target) and task-irrelevant information (the distractor) often appear at the same time, and often overlap in visual space. The distractor information, upon entering the nervous system, interferes with the neural representations of task-relevant information, causing degraded task performance (Deweese et al., 2016). In this study, we examined the temporal dynamics of target and distractor processing during sustained visual attention by analyzing EEG data from a SSVEP paradigm in which random moving dots (target) flickered at one frequency were superimposed on IAPS pictures (distractor) flickered at another frequency. The results showed that (1) distractor information (i.e., IAPS pictures from different emotion categories) can be decoded from the distributed patterns of scalp EEG, (2) both the target and the distractor are sampled rhythmically with the same sampling frequency of ~1 Hz (delta frequency band), and (3) the more negative (i.e., closer to 180 degrees) the phase relationship between the sampling of the target and that of the distractor, i.e., the more temporally separated between target sampling and distractor sampling within a sampling cycle, the better the behavioral performance.

Rhythmic sampling of attended and ignored information

Previous studies have investigated how attended information is temporally sampled using the cue-target paradigm. In particular, if some behavioral measures such as the stimulus detection accuracy or reaction time are found to be a periodic function of the time between the cue and the target, i.e., the stimulus-onset asynchrony or SOA, then it is taken as evidence in support of rhythmic sampling. If there is only one attended target, the frequency of rhythmic sampling tends to fall in the upper end of the theta frequency band (4 to 8 Hz) (Busch & VanRullen, 2010; Q. Huang & Luo, 2020; Y. Huang et al., 2015; VanRullen, 2013). When there are more than one attended targets in the environment, each target is again sampled rhythmically but the sampling frequency is slower, often falling in the lower end of the theta frequency band (Fiebelkorn et al., 2013; Landau & Fries, 2012; Re et al., 2019). When the attended targets appear in different visual fields, an alternating sampling strategy was observed, evidenced by the 180 degree phase relation between the two behavioral time courses (Chota et al., 2022; Fiebelkorn et al., 2018)

How distractors are temporally sampled has not been investigated to date. One of the reasons is that because the distractor does not elicit behavioral responses, a purely behavioral approach is not able to address this question. We overcame the problem by recording EEG in a SSVEP paradigm in which the target and the distractor overlapped in space and time, and flickered at different frequencies, a method referred to as frequency tagging. Separately extracting the EEG signals underlying the neural response to the target and that to the distractor according to their flickering frequencies (4.29 Hz for target and 6 Hz for distractor), we found that at the whole trial level, target processing exhibited higher SSVEP amplitude than distractor processing and for both target and distractor processing, the signal power is maximal in the posterior channels. Cognizant of the possibility that the power at 4.29 Hz may leak into neighboring frequency bands where the power is weaker (See Figures S4 and S5), instead of using the 6 Hz SSVEP amplitude to quantify distractor processing, we adopted the MVPA decoding approach to quantify the distractor processing by leveraging the previous finding that different categories of emotional images evoked different patterns of neural responses in scalp EEG (Bo et al., 2022). This led us to construct classifiers that took the 6 Hz SSVEP amplitude across all electrodes as input to decode the spatial patterns evoked by different categories of emotional distractors, with higher classification or decoding accuracy taken to indicate stronger distractor processing. At the whole trial level, the observed above-chance decoding accuracy suggested that the distractor information is present in the brain, and could be revealed and quantified by combining machine learning with distractor-specific scalp EEG.

Prior studies of visual environmental sampling used the cue-target paradigm in which the cue serves both to instruct the participant on how the target should be responded to as well as to reset the brain oscillation mediating the rhythmic visual sampling (Kayser, 2009). In our paradigm, the resetting was prompted by the onset of the compound stimulus array. The time elapsed after the stimulus array onset, referred to as TFO (time from onset) here, plays the role of the SOA in the cue-target paradigm. To index the temporal dynamics of target and distractor processing, we applied a moving window approach, in which the window duration was 0.5 seconds, and the step size was 0.25 seconds. Within each window, the 4.29 Hz SSVEP amplitude was taken to index target processing and the accuracy of decoding different categories of emotional distractors based on the 6 Hz SSVEP amplitude pattern was taken to index distractor processing. Plotting these two indices as functions of TFO we assessed the temporal dynamics of target and distractor processing

and. The results revealed that both the target and the distractor were sampled rhythmically with the same sampling frequency of ~1 Hz (delta frequency band), which is considerably slower than those reported in previous studies (Re et al., 2019) in which the sampling frequency tends to fall in the theta frequency band (4 to 8 Hz).

Delta oscillations (0.5 to 3.5 Hz), traditionally associated with deep sleep and homeostatic processes (Amzica & Steriade, 1998; Franken et al., 2001; Franken & Dijk, 2024; Frohlich et al., 2021; Torres-Herraez et al., 2022), are being increasingly recognized for their role in a variety of cognitive functions (Basar et al., 1999; Başar et al., 2001; Başar-Eroglu et al., 1992). In the auditory domain, rhythmic sampling of an auditory scene is known to be mediated by delta oscillations (Kubetschek & Kayser, 2021; Morillon et al., 2019). Our findings suggest that similar mechanisms could also operate in the visual system. In a recent study, when observers directed temporal attention to one of two sequential grating targets with predictable timing, the steady-state visual evoked response of the flashing target was modulated at 2 Hz (Denison et al., 2022), which falls in the delta frequency band. In addition, extensive evidence has shown that expecting a stimulus, which is known to require the deployment of attentional resources, engages delta oscillations (Arnal et al., 2011, 2015; Breska & Deouell, 2017a; Cravo et al., 2013; Lakatos et al., 2008; Schroeder & Lakatos, 2009; Stefanics et al., 2010). Delta oscillations were also involved in mechanisms that synchronize distributed regions within functional neural networks in supporting cognitive control (Breska & Deouell, 2017b; Helfrich et al., 2017, 2019; Helfrich & Knight, 2016). The spatially overlapping target and distractor in our paradigm places high demand on the brain's cognitive control system which operates in the delta frequency band (Pagnotta et al., 2024). This could be another reason underlying the observed mediation by delta oscillations in the rhythmic sampling of the target-distractor environment.

Phase relation between target and distractor sampling and its functional significance

As mentioned earlier, when two attended objects are presented simultaneously in different visual hemifields, the visual system tends to sample them in a serial, alternating fashion, as evidenced by two rhythmic behavioral time series exhibiting a 180 degree relative phase (antiphase) (Denison et al., 2022; Fiebelkorn et al., 2013; Mo et al., 2019). When two attended objects overlap in space, however, this alternating sampling pattern is not observed, and the relative phase between the two rhythmic behavioral time series appears to be uniformly distributed across participants (Re

et al., 2019). In our experimental design, the target and the distractor overlapped in space, which is a configuration known to maximize the distraction effect, and the relative phase between the rhythmic samplings of the target and the distractor is also uniformly distributed across participants. Thus, regardless of the behavioral relevance of the two superimposed stimuli, there is no preferred phase relationship between their samplings at the population level.

Although a clear phase relation between the target sampling and the distractor sampling is absent at the population level, the relative phase between the two time series may nonetheless have functional significance. In particular, when the target sampling and the distractor sampling occur in opposite phases of a sampling cycle, i.e., when they are 180 degrees out of phase, the interference should be minimized and the task performance maximized. On the contrary, when the target sampling and the distractor sampling occur in the phase of a sampling cycle, i.e., when the target and the distractor are sampled at the same time, the interference should be maximized and the task performance minimized. Our results supported this hypothesis. Specifically, we showed that there was a positive correlation between the relative phase of the target and distractor sampling time series and the behavioral performance, namely, the greater the relative phase between the two time series, the higher the rate of correctly detecting the instances of coherent motion in the moving dots. The additional significance of this finding can be understood by considering the analysis results at the whole trial level. One may expect that at the whole trial level, the stronger the distractor representation indexed by higher decoding accuracy, the worse the task performance. This turned out to be not the case. As shown in Figure 3(D), the distractor decoding accuracy at the whole trial level was not correlated with task performance, nor was the overall power of the target evoked activity at the whole trial level. Instead, we found a new mechanism underlying the competition between distractor and target. In this mechanism, the key is not how well the target and the distractor are each represented but how their respective rhythmic sampling aligns over time: The more target sampling and distractor sampling are separated in time, the less direct competition between the two, the better the attended information is processed, and the better the behavioral performance.

Signal processing considerations

When the amplitude of a periodic signal with a frequency f is modulated at 1 Hz, we should observe sidebands at $f+1$ and $f-1$ Hz in the Fourier spectrum of the signal. These sidebands are not

clearly seen in the Fourier spectrum of the SSVEP time series (see Figure 2(B)). We investigated the underlying reason in the Supplementary Materials. Biological data is noisy. The SSVEP from the subjects contains a varying among of noise quantified by signal to noise ratio (SNR). We showed using both simulations and actual data that when the SNR is high, the sidebands are visible, whereas when the SNR is low, the sidebands are indistinguishable from the noise floor (Figure S1). The majority of our subjects have low SNR for observing sidebands. This is what the sidebands in the Fourier spectrum in Figure 2(B) are not readily identifiable. Another signal processing issue is that when a 4.29 Hz periodic component and a 6 Hz component are combined, one should observe a beating frequency at 1.71 Hz. This beating frequency is clearly seen in the Fourier spectra of the amplitude envelop of the SSVEP in Figure 4(A). However, this spectrum is secondary to a much stronger spectra peak occurring at ~1 Hz, which cannot be explained from a pure signal processing perspective (see Figure S2 in the Supplementary Materials for a further investigation). This suggests that the 1 Hz amplitude modulation of the SSVEP amplitude as well as decoding accuracy time series is of an endogenous origin and represents the frequency of the rhythmic sampling of the environment by the visual attention system in our paradigm.

Summary

In this work, we reported two main findings: (1) in sustained visual attention under distraction, the distractor as well as the target are sampled rhythmically with the sampling frequency being ~1 Hz (i.e., in the delta frequency band) and (2) the temporal relationship between the distractor sampling and the target sampling is a significant factor underlying task performance with a more antiphase relationship giving rise to better behavioral performance. These results provide further support for the important role of low-frequency brain oscillations in organizing cognitive operations. They demonstrate the utility of applying machine learning methods in uncovering the temporal dynamics of sustained attention in target-distractor scenarios.

Acknowledgements

This work was supported by NSF grants BCS2318886 and BCS2318984 and NIH grant R01 MH125615.

Reference

Allen, P. J., Josephs, O., & Turner, R. (2000). A Method for Removing Imaging Artifact from Continuous EEG Recorded during Functional MRI. *NeuroImage*, *12*(2), 230–239. https://doi.org/10.1006/nimg.2000.0599

Allen, P. J., Polizzi, G., Krakow, K., Fish, D. R., & Lemieux, L. (1998). Identification of EEG Events in the MR Scanner: The Problem of Pulse Artifact and a Method for Its Subtraction. *NeuroImage*, *8*(3), 229–239. https://doi.org/10.1006/nimg.1998.0361

Amzica, F., & Steriade, M. (1998). Electrophysiological correlates of sleep delta waves. *Electroencephalography and Clinical Neurophysiology*, *107*(2), 69–83. https://doi.org/10.1016/S0013-4694(98)00051-0

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015). Delta–Beta Coupled Oscillations Underlie Temporal Prediction Accuracy. *Cerebral Cortex*, *25*(9), 3077–3085. https://doi.org/10.1093/cercor/bhu103

Arnal, L. H., Wyart, V., & Giraud, A.-L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nature Neuroscience*, *14*(6), 797–801. https://doi.org/10.1038/nn.2810

Bae, G.-Y., & Luck, S. J. (2019). Decoding motion direction using the topography of sustained ERPs and alpha oscillations. *NeuroImage*, *184*, 242–255. https://doi.org/10.1016/j.neuroimage.2018.09.029

Basar, E., Basar-Eroglu, C., Karakas, S., & Schurmann, M. (1999). Oscillatory brain theory: A new trend in neuroscience. *IEEE Engineering in Medicine and Biology Magazine*, *18*(3), 56–66. https://doi.org/10.1109/51.765190

Başar, E., Başar-Eroglu, C., Karakaş, S., & Schürmann, M. (2001). Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International Journal of Psychophysiology*, *39*(2–3), 241–248. https://doi.org/10.1016/S0167-8760(00)00145-8

Başar-Eroglu, C., Başar, E., Demiralp, T., & Schürmann, M. (1992). P300-response: Possible psychophysiological correlates in delta and theta frequency channels. A review. *International Journal of Psychophysiology*, *13*(2), 161–179. https://doi.org/10.1016/0167-8760(92)90055-G

Bo, K., Cui, L., Yin, S., Hu, Z., Hong, X., Kim, S., Keil, A., & Ding, M. (2022). Decoding the temporal dynamics of affective scene processing. *NeuroImage*, *261*, 119532. https://doi.org/10.1016/j.neuroimage.2022.119532

Breska, A., & Deouell, L. Y. (2017a). Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLOS Biology*, *15*(2), e2001665. https://doi.org/10.1371/journal.pbio.2001665

Breska, A., & Deouell, L. Y. (2017b). Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLOS Biology*, *15*(2), e2001665. https://doi.org/10.1371/journal.pbio.2001665

Busch, N. A., & VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proceedings of the National Academy of Sciences*, *107*(37), 16048–16053. https://doi.org/10.1073/pnas.1004801107

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3), 1–27. https://doi.org/10.1145/1961189.1961199

Chota, S., Leto, C., Van Zantwijk, L., & Van Der Stigchel, S. (2022). Attention rhythmically samples multi-feature objects in working memory. *Scientific Reports*, *12*(1), 14703. https://doi.org/10.1038/s41598-022-18819-z

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*(3), 332–361. https://doi.org/10.1037/0033-295X.97.3.332

Cravo, A. M., Rohenkohl, G., Wyart, V., & Nobre, A. C. (2013). Temporal Expectation Enhances Contrast Sensitivity by Phase Entrainment of Low-Frequency Oscillations in Visual Cortex. *The Journal of Neuroscience*, *33*(9), 4002–4010. https://doi.org/10.1523/JNEUROSCI.4675-12.2013

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Denison, R. N., Tian, K. J., Heeger, D. J., & Carrasco, M. (2022). *Anticipatory and evoked visual cortical dynamics of voluntary temporal attention*. https://doi.org/10.1101/2022.11.18.517084

Deweese, M. M., Müller, M., & Keil, A. (2016). Extent and time-course of competition in visual cortex between emotionally arousing distractors and a concurrent task. *European Journal of Neuroscience*, *43*(7), 961–970. https://doi.org/10.1111/ejn.13180

Dugué, L., Marque, P., & VanRullen, R. (2015). Theta Oscillations Modulate Attentional Search Performance Periodically. *Journal of Cognitive Neuroscience*, *27*(5), 945–958. https://doi.org/10.1162/jocn_a_00755

Fiebelkorn, I. C., & Kastner, S. (2019). A Rhythmic Theory of Attention. *Trends in Cognitive Sciences*, *23*(2), 87–101. https://doi.org/10.1016/j.tics.2018.11.009

Fiebelkorn, I. C., Pinsk, M. A., & Kastner, S. (2018). A Dynamic Interplay within the Frontoparietal Network Underlies Rhythmic Spatial Attention. *Neuron*, *99*(4), 842-853.e8. https://doi.org/10.1016/j.neuron.2018.07.038

Fiebelkorn, I. C., Saalmann, Y. B., & Kastner, S. (2013). Rhythmic Sampling within and between Objects despite Sustained Attention at a Cued Location. *Current Biology*, *23*(24), 2553–2558. https://doi.org/10.1016/j.cub.2013.10.063

Franken, P., Chollet, D., & Tafti, M. (2001). The Homeostatic Regulation of Sleep Need Is under Genetic Control. *The Journal of Neuroscience*, *21*(8), 2610–2621. https://doi.org/10.1523/JNEUROSCI.21-08-02610.2001

Franken, P., & Dijk, D.-J. (2024). Sleep and circadian rhythmicity as entangled processes serving homeostasis. *Nature Reviews Neuroscience*, *25*(1), 43–59. https://doi.org/10.1038/s41583-023-00764-z

Frohlich, J., Toker, D., & Monti, M. M. (2021). Consciousness among delta waves: A paradox? *Brain*, *144*(8), 2257–2277. https://doi.org/10.1093/brain/awab095

Haxby, J. V., Connolly, A. C., & Guntupalli, J. S. (2014). Decoding Neural Representational Spaces Using Multivariate Pattern Analysis. *Annual Review of Neuroscience*, *37*(1), 435–456. https://doi.org/10.1146/annurev-neuro-062012-170325

Helfrich, R. F., Breska, A., & Knight, R. T. (2019). Neural entrainment and network resonance in support of top-down guided attention. *Current Opinion in Psychology*, *29*, 82–89. https://doi.org/10.1016/j.copsyc.2018.12.016

Helfrich, R. F., Fiebelkorn, I. C., Szczepanski, S. M., Lin, J. J., Parvizi, J., Knight, R. T., & Kastner, S. (2018). Neural Mechanisms of Sustained Attention Are Rhythmic. *Neuron*, *99*(4), 854-865.e5. https://doi.org/10.1016/j.neuron.2018.07.032

Helfrich, R. F., Huang, M., Wilson, G., & Knight, R. T. (2017). Prefrontal cortex modulates posterior alpha oscillations during top-down guided visual perception. *Proceedings of the National Academy of Sciences*, *114*(35), 9457–9462. https://doi.org/10.1073/pnas.1705965114

Helfrich, R. F., & Knight, R. T. (2016). Oscillatory Dynamics of Prefrontal Cognitive Control. *Trends in Cognitive Sciences*, *20*(12), 916–930. https://doi.org/10.1016/j.tics.2016.09.007

Hindi Attar, C., & Müller, M. M. (2012). Selective Attention to Task-Irrelevant Emotional Distractors Is Unaffected by the Perceptual Load Associated with a Foreground Task. *PLoS ONE*, *7*(5), e37186. https://doi.org/10.1371/journal.pone.0037186

Huang, Q., & Luo, H. (2020). Saliency-based Rhythmic Coordination of Perceptual Predictions. *Journal of Cognitive Neuroscience*, *32*(2), 201–211. https://doi.org/10.1162/jocn_a_01371

Huang, Y., Chen, L., & Luo, H. (2015). Behavioral Oscillation in Priming: Competing Perceptual Predictions Conveyed in Alternating Theta-Band Rhythms. *The Journal of Neuroscience*, *35*(6), 2830–2837. https://doi.org/10.1523/JNEUROSCI.4294-14.2015

Jiang, Y., He, S., & Zhang, J. (2024). The adaptive flexibility of rhythmic attentional sampling in attending to multiple targets. *Journal of Experimental Psychology: General*, *153*(1), 26–37. https://doi.org/10.1037/xge0001468

Kastner, S., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Mechanisms of Directed Attention in the Human Extrastriate Cortex as Revealed by Functional MRI. *Science*, *282*(5386), 108–111. https://doi.org/10.1126/science.282.5386.108

Kastner, S., & Pinsk, M. A. (2004). Visual attention as a multilevel selection process. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(4), 483–500. https://doi.org/10.3758/CABN.4.4.483

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased Activity in Human Visual Cortex during Directed Attention in the Absence of Visual Stimulation. *Neuron*, *22*(4), 751–761. https://doi.org/10.1016/S0896-6273(00)80734-5

Kayser, C. (2009). Phase Resetting as a Mechanism for Supramodal Attentional Control. *Neuron*, *64*(3), 300–302. https://doi.org/10.1016/j.neuron.2009.10.022

Kubetschek, C., & Kayser, C. (2021). Delta/Theta band EEG activity shapes the rhythmic perceptual sampling of auditory scenes. *Scientific Reports*, *11*(1), 2370. https://doi.org/10.1038/s41598-021-82008-7

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of Neuronal Oscillations as a Mechanism of Attentional Selection. *Science*, *320*(5872), 110–113. https://doi.org/10.1126/science.1154735

Landau, A. N., & Fries, P. (2012). Attention Samples Stimuli Rhythmically. *Current Biology*, *22*(11), 1000–1004. https://doi.org/10.1016/j.cub.2012.03.054

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. https://doi.org/10.1016/j.tics.2004.12.004

Michel, R., Dugué, L., & Busch, N. A. (2022). Distinct contributions of alpha and theta rhythms to perceptual and attentional sampling. *European Journal of Neuroscience*, *55*(11–12), 3025–3039. https://doi.org/10.1111/ejn.15154

Mo, C., Lu, J., Wu, B., Jia, J., Luo, H., & Fang, F. (2019). Competing rhythmic neural representations of orientations during concurrent attention to multiple orientation features. *Nature Communications*, *10*(1), 5264. https://doi.org/10.1038/s41467-019-13282-3

Morillon, B., Arnal, L. H., Schroeder, C. E., & Keitel, A. (2019). Prominence of delta oscillatory rhythms in the motor cortex and their relevance for auditory and speech perception. *Neuroscience & Biobehavioral Reviews*, *107*, 136–142. https://doi.org/10.1016/j.neubiorev.2019.09.012

Muller, M. M., Andersen, S. K., & Keil, A. (2008). Time Course of Competition for Visual Processing Resources between Emotional Pictures and Foreground Task. *Cerebral Cortex*, *18*(8), 1892–1899. https://doi.org/10.1093/cercor/bhm215

Murphy, G., Groeger, J. A., & Greene, C. M. (2016). Twenty years of load theory—Where are we now, and where should we go next? *Psychonomic Bulletin & Review*, *23*(5), 1316–1340. https://doi.org/10.3758/s13423-015-0982-5

Pagnotta, M. F., Riddle, J., & D'Esposito, M. (2024). Multiplexed Levels of Cognitive Control through Delta and Theta Neural Oscillations. *Journal of Cognitive Neuroscience*, *36*(5), 916–935. https://doi.org/10.1162/jocn_a_02124

Posner, M. I. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

Posner, M. I., Inhoff, A. W., Friedrich, F. J., & Cohen, A. (1987). Isolating attentional systems: A cognitive-anatomical analysis. *Psychobiology*, *15*(2), 107–121. https://doi.org/10.3758/BF03333099

Posner, M. I., Petersen, S. E., Fox, P. T., & Raichle, M. E. (1988). Localization of Cognitive Operations in the Human Brain. *Science*, *240*(4859), 1627–1631. https://doi.org/10.1126/science.3289116

Re, D., Inbar, M., Richter, C. G., & Landau, A. N. (2019). Feature-Based Attention Samples Stimuli Rhythmically. *Current Biology*, *29*(4), 693-699.e4. https://doi.org/10.1016/j.cub.2019.01.010

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. https://doi.org/10.1016/j.tins.2008.09.012

Seidl, K. N., Peelen, M. V., & Kastner, S. (2012). Neural Evidence for Distracter Suppression during Visual Search in Real-World Scenes. *The Journal of Neuroscience*, *32*(34), 11812–11819. https://doi.org/10.1523/JNEUROSCI.1693-12.2012

Senoussi, M., Moreland, J. C., Busch, N. A., & Dugué, L. (2019). Attention explores space periodically at the theta frequency. *Journal of Vision*, *19*(5), 22. https://doi.org/10.1167/19.5.22

Stefanics, G., Hangya, B., Hernádi, I., Winkler, I., Lakatos, P., & Ulbert, I. (2010). Phase Entrainment of Human Delta Oscillations Can Mediate the Effects of Expectation on Reaction Speed. *The Journal of Neuroscience*, *30*(41), 13578–13585. https://doi.org/10.1523/JNEUROSCI.0703-10.2010

Thigpen, N., Petro, N. M., Oschwald, J., Oberauer, K., & Keil, A. (2019). Selection of Visual Objects in Perception and Working Memory One at a Time. *Psychological Science*, *30*(9), 1259–1272. https://doi.org/10.1177/0956797619854067

Torres-Herraez, A., Watson, T. C., & Rondi-Reig, L. (2022). Delta Oscillations Coordinate Intracerebellar and Cerebello-Hippocampal Network Dynamics during Sleep. *The Journal of Neuroscience*, *42*(11), 2268–2281. https://doi.org/10.1523/JNEUROSCI.1479-21.2021

Ungerleider, S. K. A. L. G. (2000). Mechanisms of Visual Attention in the Human Cortex. *Annual Review of Neuroscience*, *23*(1), 315–341. https://doi.org/10.1146/annurev.neuro.23.1.315

Van Der Werf, O. J., Schuhmann, T., De Graaf, T., Ten Oever, S., & Sack, A. T. (2023). Investigating the role of task relevance during rhythmic sampling of spatial locations. *Scientific Reports*, *13*(1), 12707. https://doi.org/10.1038/s41598-023-38968-z

VanRullen, R. (2013). Visual Attention: A Rhythmic Process? *Current Biology*, *23*(24), R1110–R1112. https://doi.org/10.1016/j.cub.2013.11.006

Wyart, V., de Gardelle, V., Scholl, J., & Summerfield, C. (2012). Rhythmic Fluctuations in Evidence Accumulation during Decision Making in the Human Brain. *Neuron*, *76*(4), 847–858. https://doi.org/10.1016/j.neuron.2012.09.015

Zalta, A., Petkoski, S., & Morillon, B. (2020). Natural rhythms of periodic temporal attention. *Nature Communications*, *11*(1), 1051. https://doi.org/10.1038/s41467-020-14888-8

Zhang, G., Carrasco, C. D., Winsler, K., Bahle, B., Cong, F., & Luck, S. J. (2024). Assessing the effectiveness of spatial PCA on SVM-based decoding of EEG data. *NeuroImage*, *293*, 120625. https://doi.org/10.1016/j.neuroimage.2024.120625

Supplementary Materials

In this section, we address several issues related to the work reported in the manuscript. The first two concern signal processing: (1) the observability of the sidebands in the Fourier spectrum of the SSVEP time series and (2) the beating phenomenon in the Fourier spectrum of the amplitude envelope of the SSVEP. The third one concerns the best way for quantifying distractor processing at the whole trial level and at the moving window level.

Observability of the sidebands

When the amplitude of a periodic signal with a frequency *f* is modulated at 1 Hz, we should observe sidebands at *f+1* and *f-1* Hz in the Fourier spectrum of the signal. These sidebands are not clearly seen in the Fourier spectrum of the SSVEP time series (see Figure 2(B)). We investigated the underlying reason below. Our conclusion is that for clean sinusoidal signals with periodic amplitude modulation, the sidebands are observed. However, when noise is added to the signal, the observability of the sidebands diminished increasing noise level. Biological data is noisy, and the SSVEP from each subject shows significant variability in signal-to-noise ratio (SNR) (see definition below). SNR determines whether we can observe sideband frequencies or not. We demonstrate this point first through simulation and then on our data.

(a) Simulation. Simulated signals were generated by adding a 4.29 Hz component and a 6 Hz component together. These were the same frequency components as in our experiment and also, similar to what we observed in the data, the magnitude of the 6 Hz component was made ½ that of the 4.29 Hz component. We then modulated the amplitude of these components at 1 Hz (the same as that observed in our data). The time course is shown in the left panel of Figure S1(A). In the right panel, we showed the Fourier spectrum, where the sidebands for both the 4.29 Hz component (at 3.29 Hz and 5.29 Hz) and the 6 Hz component (at 5 Hz and 7 Hz) are clearly seen. Note that there was no noise in this case.

Next, we added noise to the same signal. The signal-noise-ratio is defined as:

$$SNR_{dB} = 10 \cdot log_{10} \frac{P_{signal}}{P_{noise}}$$

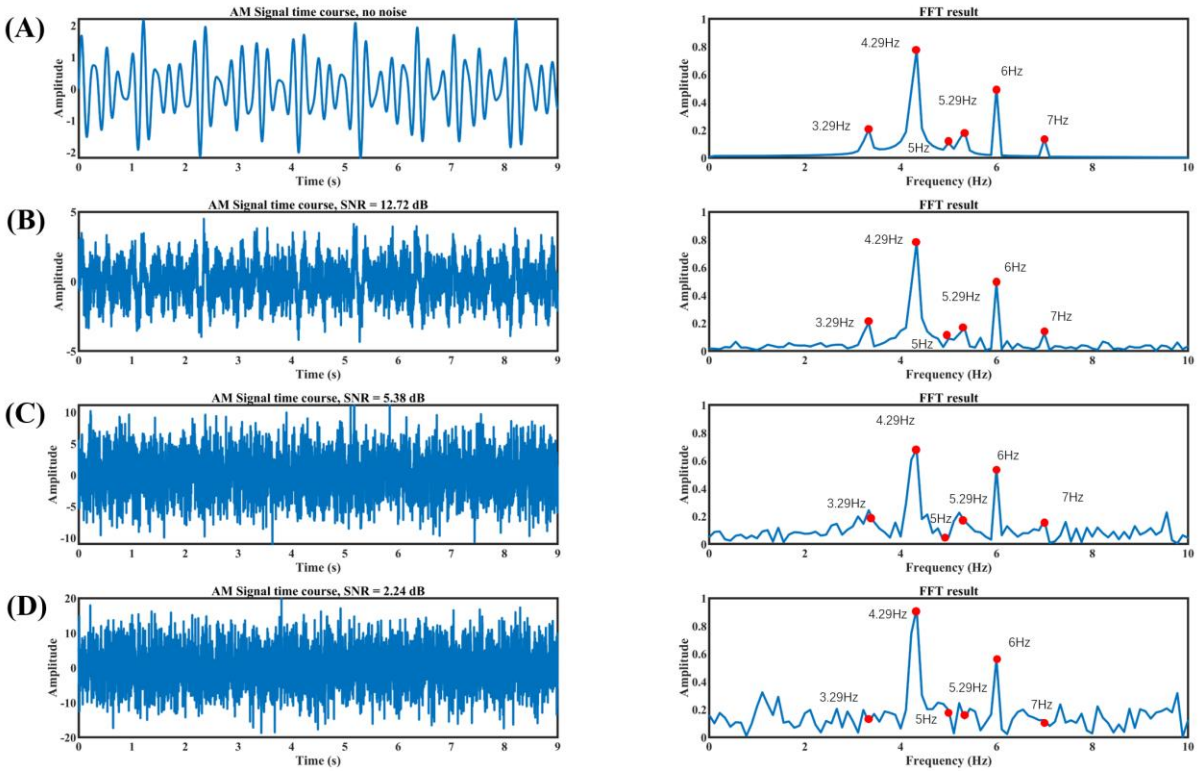

Figure S1. Simulation results. (A) The signal containing a 4.29 Hz component and a 6 Hz component where the 6Hz signal's magnitude is about half that of the 4.29Hz signal. The amplitude is modulated at 1 Hz. No noise is added. (B) Low level of noise is added to the signal in Figure S1(A) where the SNR = 12.72 dB. Sidebands are still seen. (C) Middle level of noise is added to the signal in Figure S1(A) where the SNR = 5.38 dB. Sidebands become difficult to see. (D) High level of noise is added to the signal in Figure S1(A) where the SNR = 2.24 dB, sidebands become more indistinguishable from the noise floor. Red dots indicate the location of the main frequency components and the locations where the sidebands should appear.

where the signal power is defined to be the average Fourier power within 3.8 to 4.8Hz and 5.5 to 6.5Hz and the noise power is the average Fourier power within 0 to 3Hz and 7 to 10Hz. Figure S1(B), Figure S1(C), and Figure S1(D) show the results after adding progressively more noise to the simulated signal. When the noise level is low, e.g., SNR=12.72 dB (Figure S1(B)), the sidebands are still clearly visible in the right panel of Figure S1(B) although they are not as prominent as in the right panel of Figure S1(A). When more noise is added, as shown in Figure S1(C) where SNR=5.38 dB, which is similar to what we see in a typical high SNR subject in our

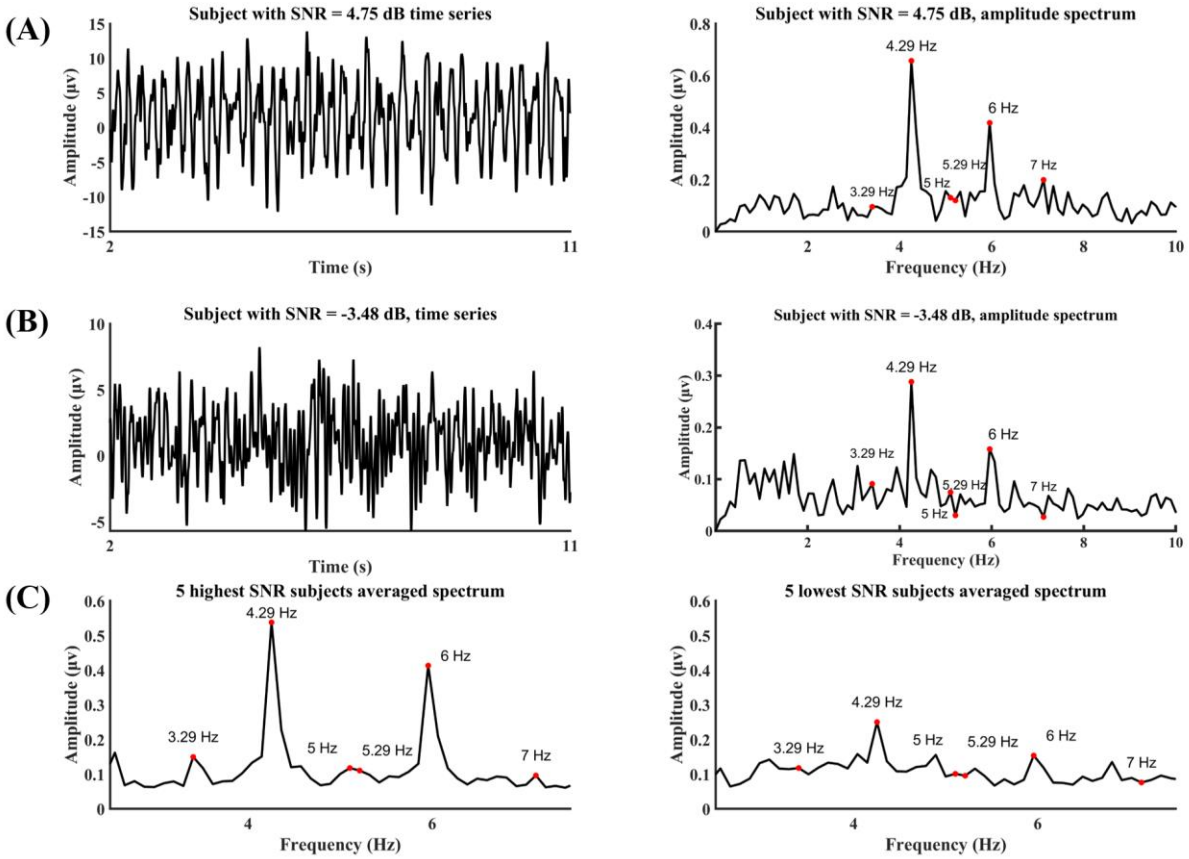

Figure S2. Experimental data. (A) The time course of the SSVEP and its Fourier spectrum from a subject with high SNR. The sidebands can be observed. (B) The time course and its Fourier spectrum from a subject with low SNR. The sidebands are indistinguishable from the noise floor. (C) The averaged Fourier spectrum from 5 highest SNR subjects and 5 lowest SNR subjects. Again, for subjects with high SNR, the sidebands are identifiable, whereas for subjects with low SNR, the sidebands are not identifiable.

data, the sidebands are beginning to become indistinguishable from the noise floor. Figure S1(D) shows a case when an even higher level of noise is added, e.g.=2.24 dB, the sidebands become even more indistinguishable from the noise floor.

(b) Experimental data. Now we demonstrate the impact of SNR on sidebands in experimental data. Figure S2(A) and (B) compare two subjects, one with relatively high and the other relatively low SNRs, respectively. For the subject with high SNR, the sidebands are still somewhat distinguishable from the noise floor, whereas for the subject with lower SNR, the sidebands are no longer visible. In Figure S2(C), we averaged the Fourier spectra of the five

subjects with the highest SNR and that of the five subjects with the lowest SNR, and the results again indicate that SNR plays a major role in determining whether the sidebands can be seen or not. For the Fourier spectrum averaged across all subjects, which is the figure shown in the manuscript, because of the influence of low SNR subjects, the sidebands are not clearly visible.

Beating phenomenon analysis

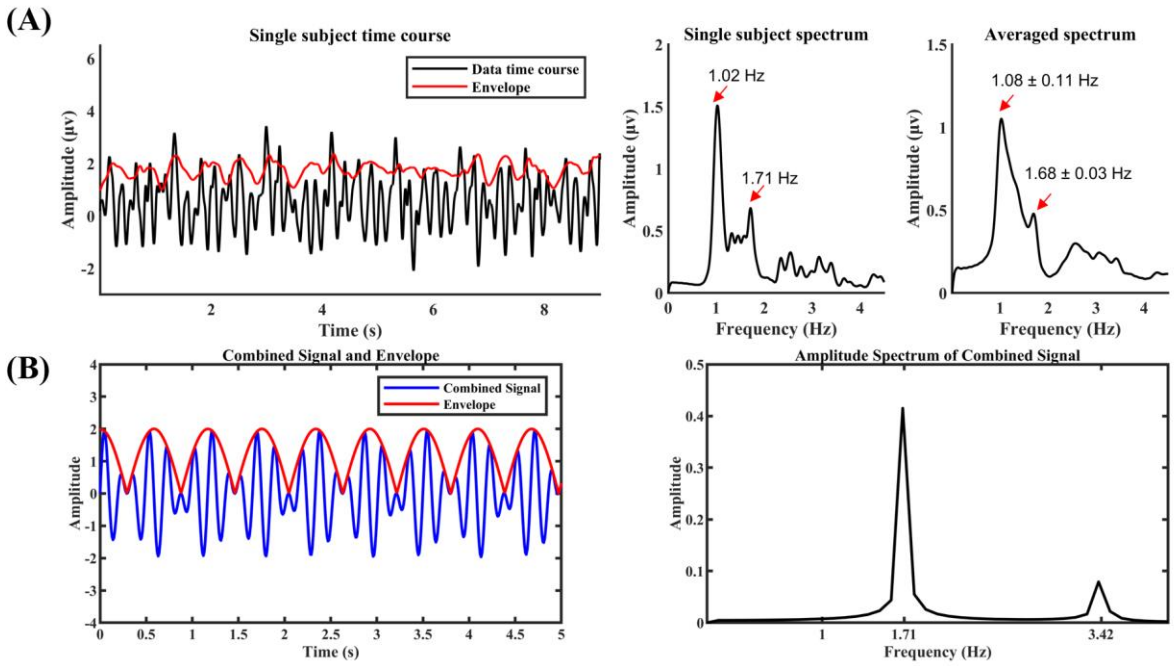

Figure S3. Beating frequency analysis. (A) The SSVEP time course from one subject (left), the Fourier spectrum of the amplitude envelope from the subject (middle), and average Fourier spectrum of the amplitude envelops from all subjects (right). In both spectra we observed the spectral peal at the beating frequency 1.71 Hz. Importantly, we also observed a much stronger spectral peak at around 1 Hz which cannot be explained by the beating phenomenon. (B) Simulated signal by combining a 4.29 Hz component and a 6 Hz component (left). The spectral peak corresponding to the beating frequency 1.71 Hz is clearly seen but there is no spectral peak at 1 Hz. This simulation further demonstrated that the 1 Hz spectral peak in (A) does not arise from a pure signal processing perspective.

From a pure signal processing point of view, the combination of the 4.29 Hz and the 6 Hz components will generate a signal whose amplitude is modulated at 1.71 Hz (the beating frequency). Thus, when the spectrum of the amplitude envelope is computed, a peak at 1.71 Hz should be observed. We see this peak in our data, as shown in Figure S3(A), where the left panel shows the amplitude envelop of a subject, the middle panel shows the spectrum of the Fourier

spectrum of the amplitude envelope from the subject in the left panel, and the right panel shows the average Fourier spectrum of the amplitude envelopes from all subjects. Both spectra exhibit a clear peak at around the beating frequency (1.71 Hz). Importantly, a much stronger spectral peak occurred at around 1Hz, which cannot be explained from a pure signal processing point of view, suggesting that it is an endogenously generated rhythmic activity.

To further investigate this, we performed a simulation in which a 4.29 Hz component and a 6 Hz component were combined. As Figure S3(B) shows, the Fourier spectrum of the amplitude envelop revealed a distinct 1.71 Hz peak and its harmonics, but there was no spectral peak at the 1 Hz peak, again suggesting that the presence of the 1 Hz rhythmic activity in the experimental data is not a signal processing property.

These results strongly support the conclusion that the 1 Hz amplitude modulation is not a byproduct of signal processing but is of an endogenous origin. The work reported in the manuscript established that this 1 Hz modulation reflects a rhythmic sampling process intrinsic to the attention system.

Quantifying distractor processing

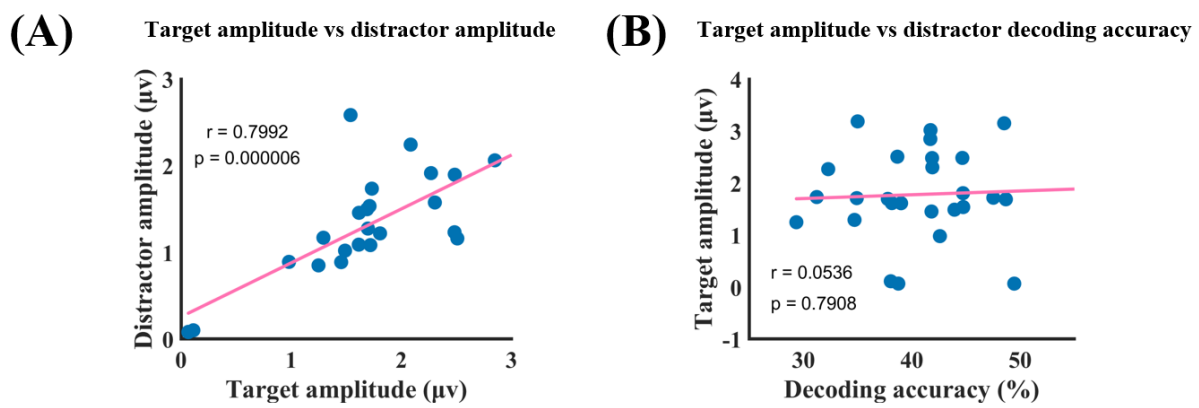

Figure S4. SSVEP amplitude analysis at the whole trial level. (A) Target amplitude vs distractor amplitude, where the correlation is r = 0.7992 (p = 0.000006), suggesting the 6 Hz signal amplitude is strongly influenced by the 4.29 Hz signal amplitude. (B) Target amplitude vs distractor decoding accuracy, where the correlation is r = 0.0536 (p = 0.7908), suggesting that the decoding accuracy as an index of distractor processing is not influenced by the 4.29 Hz target amplitude.

In the SSVEP literature, signal amplitude at the flicker frequency is the main variable for quantifying the processing of the flickering stimulus. In this work, we treated the scalp pattern of the signal amplitude at the distractor flicker frequency as input features and subjected them to decoding analysis. Below we show that the decoding accuracy is a more suitable variable for quantifying distractor processing, together with some signal processing concerns about the analysis.

Whole trial level. The target amplitude (4.29 Hz) and the distractor amplitude (6 Hz) were extracted from each subject and displayed in Figure S4(A). The strong correlation suggests that there is significant power leakage from the stronger target frequency into the weaker distractor frequency. However, when the target amplitude and distractor decoding accuracy are plotted, no correlation was found, as shown in Figure S4(B), suggesting that the target amplitude has no influence on distractor decoding accuracy.

Moving window level. Similar results can be observed at the moving window level. The target amplitude time series and the distractor amplitude time series were extracted from each subject and the relation between the two time series assessed. The relative phase is also correlated with behavior. Figure S5(A) shows that the relative phase is narrowly distributed around 0.23 ± 0.05π, which is confirmed by the K-S static of 0.46 in Figure S5(B), demonstrating that the relative phase distribution is significantly different from the uniform distribution at p=0.00001. This means

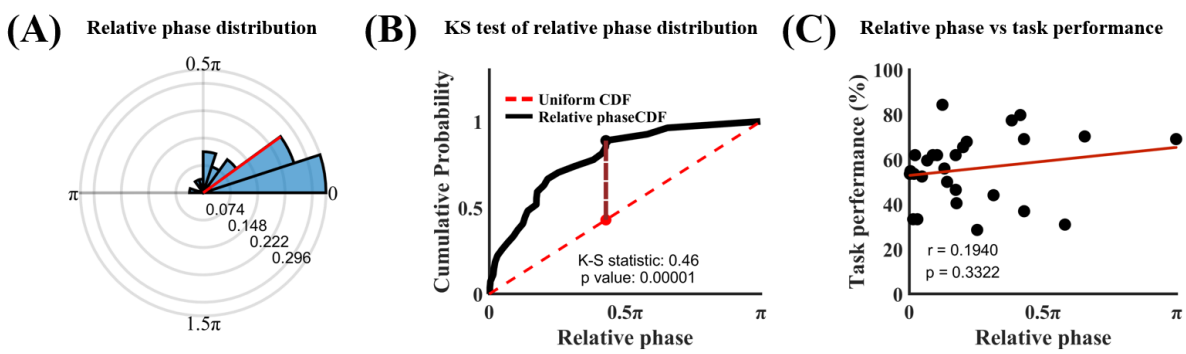

Figure S5. Moving window analysis. (A) The relative phase between the target amplitude time series and the distractor amplitude time series. (B) Kolmogorov-Smirnov test showed that the relative phase distribution is significantly different from the uniform distribution. (C) Relative phase vs task performance. r=0.1940 (p=0.3322) means that there is no significant correlation between amplitude relative phase and task performance.

that the power leakage also impacts the distractor time series, making it nearly phase locked to the target time series. It is not surprising that the relative phase between the two amplitude time series does not predict task performance (Figure S5(C)).