

Task demands modulate pre-stimulus alpha frequency and sensory template during bistable apparent motion perception

Biao Han, PhD^{1,2,3,†}, Yanni Zhang, MSc^{2,†}, Lu Shen , PhD^{1,2,3,†}, Lei Mo, PhD^{1,2,3,*}, Qi Chen, PhD^{1,2,3,*}

¹Key Laboratory of Brain, Cognition and Education Sciences, Ministry of Education, Guangzhou 510631, China,

²School of Psychology, South China Normal University, Guangzhou 510631, China,

³Center for Studies of Psychological Application, and Guangdong Key Laboratory of Mental Health and Cognitive Science, South China Normal University, Guangzhou 510631, China

*Corresponding authors: Qi Chen, School of Psychology, South China Normal University, Guangzhou 510631, China. Email: qi.chen27@gmail.com; Lei Mo, School of Psychology, South China Normal University, Guangzhou 510631, China. Email: molei@scnu.edu.cn.

[†]Biao Han, Yanni Zhang, and Lu Shen contribute equally to this work.

Despite ambiguous environmental inputs, top-down attention biases our subjective perception toward the preferred percepts, via modulating prestimulus neural activity or inducing prestimulus sensory templates that carry concrete internal sensory representations of the preferred percepts. In contrast to frequent changes of behavioral goals in the typical cue-target paradigm, human beings are often engaged in a prolonged task state with only 1 specific behavioral goal. It remains unclear how prestimulus neural signals and sensory templates are modulated in the latter case. To answer this question in the present electroencephalogram study on human subjects, we manipulated sustained task demands toward one of the 2 possible percepts in the bistable Ternus display, emphasizing either temporal integration or segregation. First, the prestimulus peak alpha frequency, which gated the temporal window of temporal integration, was effectively modulated by task demands. Furthermore, time-resolved decoding analyses showed that task demands biased neural representations toward the preferred percepts after the full presentation of bottom-up stimuli. More importantly, sensory templates resembling the preferred percepts emerged even before the bottom-up sensory evidence were sufficient enough to induce explicit percepts. Taken together, task demands modulate both prestimulus alpha frequency and sensory templates, to eventually bias subjective perception toward the preferred percepts.

Key words: alpha frequency; sensory templates; Ternus display; temporal integration; temporal segregation.

Introduction

Despite the constantly noisy and ambiguous environmental inputs, top-down attention biases the perception system toward the percepts which are the most pertinent to the current behavioral goals (Moran and Desimone 1985; Corbetta and Shulman 2002; Reynolds and Chelazzi 2004; Maunsell and Treue 2006; Nobre and Kastner 2014). In particular, directing top-down attention toward a specific spatial or nonspatial object feature leads to prestimulus baseline activity increases both in the object/feature selective sensory cortex (Wyart et al. 2012; Kok et al. 2014, 2017) and the higher order frontoparietal cortex (Chelazzi et al. 1993; Buschman and Miller 2007; Szczepanski et al. 2010; Peelen and Kastner 2011; Baldauf and Desimone 2014; Eimer 2014). Moreover, top-down attention can induce sensory templates in the object/feature selective sensory cortex during the prestimulus phase as well (Eimer 2014; Peelen and Kastner 2014). These sensory templates carry neural representational information resembling the preferred stimuli, which biases the perception system toward percepts matching the templates (Stokes et al. 2009; Peelen and Kastner 2011, 2014;

Eimer 2014; Kok et al. 2014, 2017; Nobre and Stokes 2019). A majority of the above attentional effects are revealed via the typical cue-target paradigm, in which the behavioral goal frequently changes on a trial-by-trial base. However, rather than frequent moment-to-moment changes in transient behavioral goals, human beings often engage in an ongoing behavioral task for a prolonged period of time with only 1 specific behavioral goal, such as search for a target person or object (Duncan and Humphreys 1989; Wolfe et al. 1989; Wolfe 2015). It remains less known whether and how prestimulus neural signals and sensory templates are modulated by the sustained task state. By manipulating task demands in the bistable Ternus apparent motion paradigm (Pikler 1917; Ternus 1926), we aimed to investigate how the sustained task state biases subjective perception via modulation of prestimulus neural signals and sensory templates.

Human research has shown that the top-down modulation is closely related to variations in the prestimulus alpha-band power (Klimesch et al. 1998; Sauseng et al. 2005; Haegens et al. 2011; van Diepen et al. 2015),

phase (Samaha et al. 2015), and frequency (Haegens et al. 2014; Samuel 2018; Wutz et al. 2018). In one of our previous studies, we have shown that the prestimulus intrinsic alpha frequency gates the temporal window of perceptual grouping (Shen et al. 2019). When the inter-frame interval (IFI) between 2 consecutively presented frames reaches a certain psychophysical threshold, subjective perception spontaneously alternates between 2 interpretations of the same dynamic stimuli (Fig. 1A). The slower the prestimulus ongoing alpha frequency, the higher probability the 2 frames will fall in the same alpha cycle and be temporally grouped, resulting in element motion (EM) percepts. The faster the prestimulus alpha frequency, the higher possibility the 2 frames will fall in different alpha cycles, and be temporally segregated, resulting in group motion (GM) percepts. Here, we manipulated goal-directed attention toward either the EM or the GM percepts by introducing different task demands. In the element-motion task (EMT), participants were asked to judge whether or not they perceived the stimuli as EM. In this way, the EM percept was set as the current behavioral goal, and thus, temporal integration was emphasized. In the group-motion task (GMT), participants were asked to judge whether or not they perceived the stimuli as GM. In this way, the GM percept was set as the current behavioral goal, and thus, temporal segregation was emphasized (Fig. 1A and B). Therefore, both the GM and the EM percepts could be either preferred or nonpreferred by the current goal-directed top-down attention (Fig. 1C). If task demands indeed adopt the prestimulus alpha-band oscillations to facilitate behavioral goals, they should be able to modulate the prestimulus alpha frequency. Since temporal integration requires a longer alpha cycle, while temporal segregation requires a shorter alpha cycle (Cecere et al. 2015; Samaha and Postle 2015; Shen et al. 2019), the task demands emphasizing temporal integration in the EMT should cause significantly lower prestimulus alpha frequency than the task demands emphasizing temporal segregation in GMT. Moreover, this modulation effect should be larger for the preferred (i.e. GM percepts in the GMT, and EM percepts in the EMT) than nonpreferred (i.e. EM percepts in the GMT, and GM percepts in the EMT) percepts.

Moreover, it has been shown that, when attention was engaged for a prolonged period of time (e.g. with task demands), the sensory templates do not persistently exist during the prestimulus phase but rather are revealed only shortly around the presentation of the bottom-up stimuli (Myers et al. 2015). However, it remains unclear how such a short-lived sensory template around the stimulus onset facilitates the subsequent perceptual processing, especially in ambiguous situations. One possibility is that the sensory templates are induced by the very first but still ambiguous sign of sensory evidence, which biases subjective perception toward the matching percept. The bistable Ternus paradigm provides an opportunity to test this hypothesis: no explicit percepts could be formulated until the second frame. Therefore,

we predicted that sensory templates of the preferred percepts should be observed before the appearance of second frame. Particularly for the preferred percepts, neural signals immediately after the first frame and prior to the second frame should resemble the neural signals during the later stages of sensory processing after the second frame when explicitly perceived percepts have been formulated.

Materials and methods

Participants

Twenty-four healthy right-handed participants (6 females; age 22.7 ± 1.3 years, mean \pm SD) took part in this experiment. The sample size was decided a priori and ensured at least 80% power to detect experimental effects with at least moderate to large effect size (Cohen's $d > 0.6$). Two participants were excluded from the analysis because of excessive eye movement artifacts and 1 participant was excluded because of the excessive electroencephalogram (EEG) artifact. A total of 21 participants were included in the final analysis. All participants have a normal or corrected-to-normal vision and no history of neurological or psychiatric illness. All participants gave their informed consent prior to the experiment in accordance with the Declaration of Helsinki and the study was approved by the Ethics Committee of the School of Psychology, South China Normal University.

Data and code availability statement

All the data and code are available at <https://osf.io/t9bja/>. The data and code used in the study are available in the public domain for its sharing or reuse. The data and code sharing adopted by the authors comply with the requirements of the institute and comply with institutional ethics approval.

Stimuli and experimental design

Visual stimuli were presented using Presentation software (Neurobehavioral Systems; <http://www.neurobs.com>) on a high refresh-rate LCD monitor (resolution: 1366×768 , refresh rate: 100 Hz), which was located 57 cm in front of the participant and was calibrated using an oscilloscope before the experiment. During the experiment, a chin rest was used to stabilize the head position.

Visual stimuli consisted of 2 consecutively presented frames of stimuli (frame 1 and frame 2), and each frame was presented for 30 ms (Fig. 1B). There was a blank period between the 2 frames, i.e. the IFI. Each frame contained 2 horizontally arranged black disks on a gray background. The 2 black disks are both 3° of visual angle in diameter and the center-to-center distance between them is 4.5° of visual angle. The 2 frames shared 1 common disk location at the center of the display. The location of the lateral disk in frame 1, either on the left or the right side of the shared central disk, was always opposite to the lateral disk in frame 2 (Fig. 1B).

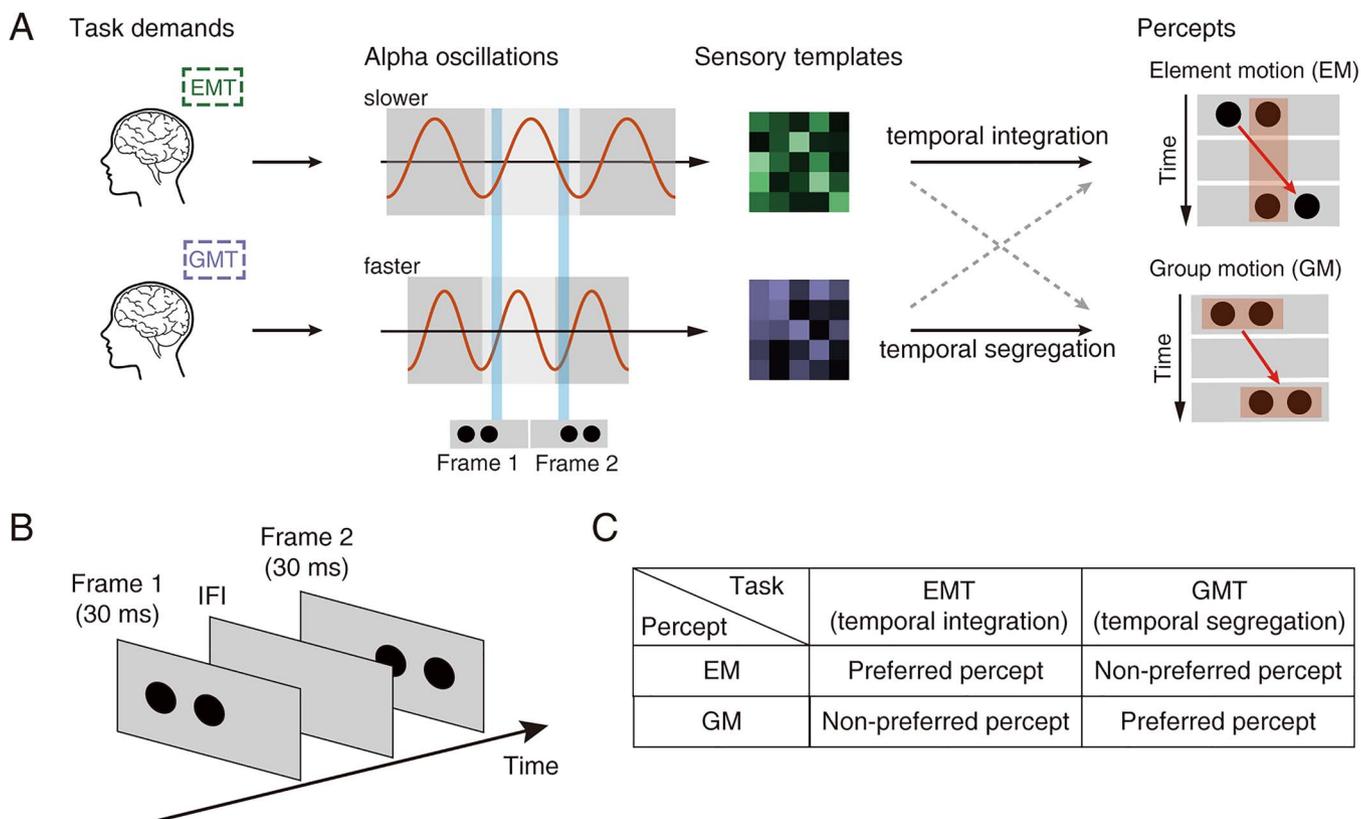


Fig. 1. Hypothesis and paradigm. A) In the bistable Ternus display, when the IFI between two consecutively presented frames reaches a certain psychophysical threshold, subjective perception spontaneously alternates between 2 interpretations of the same dynamic stimuli: when the 2 frames are temporally grouped, the central overlapping disks between the 2 frames are integrated as 1 single disk and the lateral disk appears to jump across the central disk, resulting in EM percept; when the 2 frames are temporally segregated, the 2 disks within each frame are spatially grouped and perceived as moving together, resulting in GM percepts. In the current experiment, participants were instructed to pay attention to either the EM percepts, i.e. the EMT, emphasizing temporal integration, or to the GM percepts, i.e. the GMT, emphasizing temporal segregation. We hypothesize that task demands slow down the alpha oscillations in EMT and accelerate the alpha oscillations in GMT, so that the 2 frames will be more likely to fall in the same alpha cycle and be temporally integrated in EMT, and will be more likely to fall in different alpha cycles and be temporally segregated in GMT. Meanwhile, the EMT and the GMT should also employ different sensory templates that contain neural representations similar to preferred percepts. B) The experimental stimuli consisted of 2 consecutively presented frames of stimuli and each frame was presented for 30 ms. There was a blank period between the 2 frames, i.e. the IFI, which was determined by a psychophysical pretest. Participants were asked to discriminate whether or not the stimuli were perceived as the EM in the EMT and whether or not the stimuli were perceived as the GM in the GMT. The intertrial interval was selected randomly between 1.5 and 2.5 s. C) Under different task demands, either the EM or the GM percepts could be perceived at the threshold IFI. If the perceived percept was favored by the task demands, it was defined as the preferred percept; otherwise, it was defined as the nonpreferred percept.

Specifically, if the lateral disk in frame 1 is on the left side and the lateral disk is on the right side in frame 2, a rightward apparent motion will be induced; vice versa, a leftward apparent motion will be induced. Despite the direction of the apparent motion but depend on the IFIs, 2 types of percepts could be induced: (i) with short IFIs, the central overlapping disks between the 2 frames were temporally integrated as 1 single disk, and the visual persistence of the central overlapping disk made the lateral disk in frame 1 appear to jump across the central disk, i.e. EM; (ii) with long IFIs, the 2 disks within each frame were spatially grouped and perceived as moving together as a group, i.e. GM. Participants were asked to perform a 2-alternative forced-choice to an affirmative-negative question to discriminate whether the stimuli were perceived as the EM in the element motion task (EMT) and whether the stimuli were perceived as the GM in the group motion task (GMT). In EMT, a “yes” response corresponded to an EM percept which was defined as

preferred percept, whereas a “no” response corresponded to a GM percept which was defined as the nonpreferred percept. Similarly, in GMT, GM was defined as the preferred percept and EM was defined as the nonpreferred percept (Fig. 1C).

In the experiment, each participant conducted a practice session, then a psychophysical pretest, and finally the main EEG experiment. During the practice session, participants were shown demos of the explicit EM and GM trials and practiced the experiment with only explicit EM and GM trials until the accuracy reached no less than 95%. During the psychophysical pretest, the first frame was presented for 30 ms, after a variable IFI (seven levels: 50, 80, 110, 140, 170, 200, or 230 ms), the second frame was presented for 30 ms as well. Participants were asked to perform the EMT and GMT in 2 different blocks and the order of the tasks was counter-balanced between participants. In EMT, the proportion of EM reports (i.e. “1- the proportion of GM reports”) was collapsed over

the leftward and rightward motion directions for each IFI condition and relabeled as the proportion of preferred percepts. Similarly, in GMT, the proportion of GM reports was collapsed and also relabeled as the proportion of preferred percepts. Two psychometric curves were separately fitted for the 2 tasks using logistic functions (Treutwein and Strasburger 1999), and the PSE for each task was obtained by estimating the 50% performance point on the fitted logistic function for each participant. In order to match the difficulty between the 2 tasks, the common IFI in the intersection of the 2 psychometric curves was assigned as the participant-specific task threshold IFI (Wutz et al. 2018), which was used for subsequent main EEG experiment.

In the main EEG experiment, the first frame was presented for 30 ms, and after a variable IFI (50 ms, 230 ms, or the task threshold IFI), the second frame was presented for another 30 ms. Participants were instructed to perform the EMT or GMT in 2 separate blocks using a response pad. The order of the tasks and the mapping between the 2 response buttons was counter-balanced between participants. For each task block, each participant completed 40 explicit EM trials (IFI 50 ms), 40 explicit GM trials (IFI 230 ms), 320 bistable trials (task threshold IFI), which were intermixed randomly for each participant, resulting in 800 experimental trials in total. For both tasks, each trial was followed by a time interval that was selected randomly between 1.5 and 2.5 s. The participants were asked to rest for a short period of time for every 100 trials.

Recording and preprocessing of the EEG data

EEG was recorded at 1024 Hz from Ag/AgCl electrodes using a Neuronscan SynAmps2 system using 64-channels Quick-Caps. Horizontal and vertical electrooculograms were recorded by 4 additional electrodes around the participants' eyes. All the electrode impedances were kept below 5 k Ω . Signals were referenced online to an electrode between Cz and CPz. Offline processing and analysis were performed using FieldTrip toolbox (Oostenveld et al. 2011) (www.fieldtriptoolbox.org) and customized MATLAB scripts. Data were down-sampled to 512 Hz, filtered to remove the 50-Hz power line interference, and rereferenced to the average reference. Data were epoched from -1000 to 1000 ms and baseline corrected by subtracting the mean amplitude during an interval between -500 and 0 ms relative to the presentation of frame 1. Please note that subtracting the average amplitude does not affect the prestimulus alpha oscillations, as it mainly affects only the zero-frequency component (Cohen 2014a). Trials with blinks, muscle movements, and other EEG artifacts were identified and removed based on visual inspection before dividing the trials into different conditions.

Alpha oscillations analysis

Instantaneous peak alpha frequency and alpha power were analyzed using the methods and codes developed

by Cohen (Cohen 2014b). To be specific, instantaneous peak alpha frequency analysis was performed using the data from -500 to 0 ms to avoid contaminations by the poststimulus signals. To reduce the edge artifacts caused by filtering, the prestimulus signals of each bistable trial were copied, flipped from left to right, and appended to the right side of the original data. Since no poststimulus data were entered into the analysis in this way, the analysis could not be contaminated by the poststimulus activity. These epochs were bandpass filtered between 8 and 13 Hz with a zero-phase, plateau-shaped filter with 15% transition width. A Hilbert transform was applied to obtain the phase angle and amplitude time series on each trial. The temporal derivative of the instantaneous Hilbert phase corresponds to the instantaneous frequency in hertz (when scaled by the sampling rate and 2π) and the square of the amplitude corresponds to the alpha power. Because the phase angle time series is prone to noise that can cause sharp, nonphysiological responses in its derivative, as previous studies (Cohen 2014b; Samaha and Postle 2015), the instantaneous frequency estimate was filtered 10 times with a median filter with an order of 10 and a maximum window size of 400 ms: data were median filtered 10 times with 10 time windows ranging from 10 to 400 ms prior to averaging across trials.

To control for the possible effect of $1/f$ slope, we reconstructed the demodulated time-domain signal with attenuated $1/f$ (Samaha and Cohen 2022). Specifically, after extracting the amplitude spectrum and the phase spectrum, respectively, from each trial of the original signal by fast Fourier transform (FFT), the polynomial fitting was performed on the amplitude spectrum to obtain the best fitting value; the reciprocal of which was point-wise multiplied with the amplitude spectrum to obtain the demodulated amplitude spectrum. The obtained amplitude spectrum was combined with the phase spectrum of the original signal into complex values, and the demodulated time-domain signal was recovered by the inverse FFT. We then performed the instantaneous frequency analysis with the demodulated signal through a filter with fixed frequency boundary of 8–13 Hz for all subjects. To further control for the instantaneous frequency bias due to variations in individual's alpha peak, after the removal of the $1/f$ component, we detected the alpha frequency peak in each subject and designed the individualized filter centered on the subject-specific alpha peak. The frequency boundaries of the individualized filter were designed as ± 2 Hz from the individual peak alpha frequency. We then performed the same instantaneous frequency analysis with the redesigned filter.

For the power analysis, we obtained the full power spectrum from 2 to 30 Hz in the region of interest for both conditions by performing a FFT on the prestimulus EEG data (-500 to 0 ms relative to the first frame) for all participants. We also extracted the instantaneous alpha power using the same filter, concatenation, and Hilbert

transform as those in the instantaneous alpha frequency (IAF) analysis.

Decoding analysis

Multivariate pattern analysis techniques were employed to investigate the neural representation of different percepts in different conditions with the progress of time. Classifications were based on the regularized linear discriminant analysis (Fisher 1936; Mostert et al. 2015; Kok et al. 2017; Shen et al. 2019) to identify a projection in the multidimensional EEG data, x , that maximally discriminated between the 2 subjective percepts (EM and GM) regardless of the task. Each projection is defined by a weight vector, w , which describes a representation profile y at time t using the EEG data x

$$y_t = \sum_i w_{i,t} x_{i,t} + c_t,$$

where i is the number of the EEG channel and c is a constant. For the sake of being representative, the representation profiles should be similar from each other across different trials in 1 condition, but they also should be dissimilar from each other across trials in different conditions. Therefore, in order to best characterize the 2 subjective percepts in our experiment, a score function can be created

$$S(W) = \frac{\text{dist}(y_t^{EM}, y_t^{GM})}{\text{var}(y_t^{EM}) + \text{var}(y_t^{GM})},$$

where W is a set of weight vectors, $\text{dist}(x, y) = (x - y)^2$ is the function for calculating the distance between the mean representation profiles of the 2 percepts, and $\text{var}(x)$ is the function to calculate the variance within one percept. The weight vectors were the optimal ones when the score function was maximized.

Before performing the decoding, in order to improve the signal-to-noise ratio, the data were first averaged within a window of 40 ms centered on the time point of interest. To avoid potential confounds caused by the different numbers of trials, the trial count was matched for all conditions by randomly selecting a subsample of trials from the conditions with more trials. Classifiers were trained and tested using a leave-one-out cross-validation method (Varoquaux et al. 2017). For the temporal generalization analysis, the classifier performance was assessed not only at the time point used for training (for example, classifier $w(t_1)$ was tested at t_1 , $w(t_2)$ was tested at t_2 , and so on) but also on data from all the other time points (for example, classifier $w(t_1)$ was tested on all the time points t_1, t_2, t_3 , and so on) which provided us with a (training time) \times (decoding time) temporal generalization matrix. The performance of the classifier was quantified using the representation profile y and area under receiver operator characteristic curve (AUC).

Statistics

Statistical analysis of the instantaneous peak alpha frequency was performed at the spatiotemporal level using a nonparametric, cluster-based permutation test, which controls the type I error rate in the context of multiple comparisons (Maris and Oostenveld 2007). Data from all participants were pooled together, and independent sample t -statistics were first calculated between the 2 conditions for all prestimulus time points (-500 to 0 ms) and all channels. Next, elements that passed a threshold value corresponding to a P -value of 0.05 (2-tailed) were marked, and spatiotemporal clusters (Fig. 3A) of marked elements were identified. The t -values within a connected cluster were summed up as a cluster-level statistic. Then, data were shuffled between 2 conditions for 1000 times. For each shuffle, the maximum cluster-level statistic was entered into a distribution of cluster-level statistics, which was expected under the null hypothesis. Last, cluster-level statistics were considered as significant only if they exceeded 95% of the null distribution of clusters, at $\alpha = 0.05$. The channels and time points in the cluster were selected for subsequent analyses as the region of interest and time of interest, respectively. Data from the region of interest were further averaged together for each condition, and the paired t -statistics were further performed only for display purposes (Fig. 3B). Statistical analysis for the peak alpha frequency in preferred and nonpreferred percepts (Fig. 3D and E) was performed within the region of interest using a similar cluster-based multiple-comparisons correction method: paired t -statistics were first computed between the EM and GM conditions across participants, and elements, which passed a threshold of $P < 0.05$ (2-tailed), were marked to identify the temporal clusters. The cluster-level statistics were obtained using a similar shuffling process within each participant. For the power analysis, paired t -statistics were then computed between the 2 conditions across participants, and frequency points, which passed a threshold value corresponding to a P -value of 0.05 (2-tailed), were marked as significant.

Statistical analysis of the time-resolved analysis was performed using a similar nonparametric cluster-based permutation approach as in the alpha oscillations analysis. For the temporal generalization analysis, instead of using a single-dimensional cluster-based permutation, a 2-dimensional cluster-based permutation was applied and clusters were defined as the neighboring elements which were cardinally adjacent.

Results

Behavioral results

For each participant, via psychophysical pretests prior to the main experiment, 2 psychometric curves between the proportions of EM/GM percept and the IFIs were first fitted for the EMT and GMT, respectively (Fig. 2A; see Materials and methods, the fitted curves and R^2 s for all

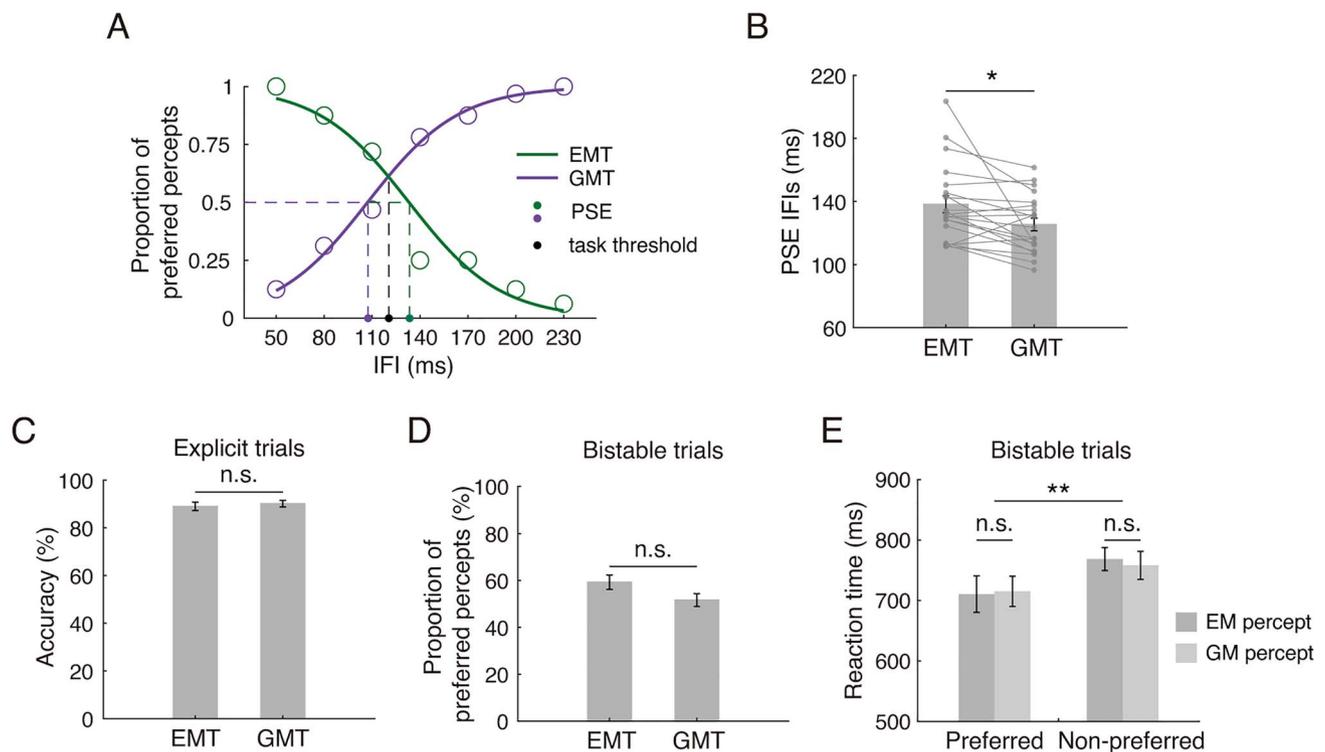


Fig. 2. Behavioral results. A) Behavioral data from the psychophysical pretest in a typical participant. Two fitted psychometric curves are shown for the EMT (green) and the GMT (purple). The green and purple dots on the horizontal axis show the IFIs for the PSE for EMT and GMT, respectively. The black dot shows the task threshold IFI which was determined as the intersection point of the 2 psychometric curves. At the task threshold IFI, equal proportions of the preferred percepts were perceived in the EMT and GMT. B) PSE IFI results in the psychophysical pretests for all the participants. The PSE IFI in GMT was significantly shorter than the EMT. C) Mean accuracy rates of the explicit trials (collapsed over the explicitly short and long IFI trials) in the EMT and the GMT of the main experiment. D) Mean proportions of the preferred percepts for the bistable trials in the EMT and GMT of the main experiment. E) Mean RT for the preferred and the nonpreferred percept in bistable trials in the EMT and GMT of the main experiment. n.s., not significant; * $P < 0.05$; ** $P < 0.01$.

the participants and both tasks are shown in Fig. S1, see online [Supplementary Material](#) for a color version of this figure). We first calculated the point of subjective equality (PSE) in the 2 tasks of the psychophysical pretests, respectively. The PSE was defined as the IFI at which equal proportions of EM and GM percepts were reported. A significant shift in the PSE was found between the 2 tasks: the PSE was significantly shorter in the GMT (125 ± 3.9 ms, mean \pm standard error) than the EMT (138 ± 5.2 ms), 2-tailed paired t -test: $t_{(20)} = 2.69$, $P = 0.01$ (Fig. 2B). The significant shift in PSE between the 2 tasks suggested that the different task demands successfully modulated the temporal window of perceptual grouping, which correspondingly biased subjective perception toward the percept matching the behavioral goal (i.e. EM in the EMT and GM in the GMT). Specifically speaking, when the task required temporal integration in the EMT, the 2 frames were biased to be more frequently integrated (i.e. perceived as EM). To reach a balance between the EM and GM percepts (i.e. comparable proportions), the IFI had to be further enlarged to allow the 2 frames to be more frequently segregated (i.e. perceived as GM). In contrast, when the task required temporal segregation in the GMT, the 2 frames were biased to be more frequently segregated (i.e. perceived as GM). To reach a proportional balance between the GM and EM percepts, the IFI had

to be further reduced, allowing the 2 frames to be more frequently integrated (i.e. perceived as EM).

To match the proportion of the preferred percepts and task difficulty, and to keep the bottom-up inputs constant between the GMT and the EMT in the main experiment, for each individual participant, we calculated the IFI at the intersection point of the 2 psychometric curves in the psychophysical pretests and used it as the same individual task threshold IFI for the GMT and the EMT in the main experiment. Therefore, for each individual participant, an identical stimulus set was adopted for the GMT and the EMT in the main EEG experiment. The identical stimulus set consisted of: (i) bistable trials with the individual task threshold IFI, to induce bistable percepts; and (ii) explicit trials with either short (50 ms) or long (230 ms) IFIs, to induce explicit EM or GM percepts. The order of the GMT and the EMT was counter-balanced across participants. The behavioral results in the main experiment further proved the effectiveness of the above manipulations. The mean accuracy rates in the explicit trials of the 2 tasks (collapsed over the short and long IFIs) were comparable and both above 85% (EMT: $89 \pm 7.9\%$, GMT: $90.1 \pm 6.2\%$, mean \pm SEM; 2-tailed paired t -test: $t_{(20)} = 0.71$, $P = 0.48$), indicating that participants could clearly discriminate the explicit percepts in both tasks (Fig. 2C). For the bistable trials with

the task threshold IFI, the proportion of the preferred percepts was comparable between the 2 tasks (EMT: $59.3 \pm 13.7\%$, GMT: $51.7 \pm 12.1\%$; 2-tailed paired t-test: $t_{(20)} = 1.72$, $P = 0.10$) (Fig. 2D). The mean reaction times (RTs) were submitted to a repeated-measures ANOVA, with the type of perceived percepts (EM vs. GM) and the percept preference (preferred vs. nonpreferred) as the 2 within-subject factors. The main effect of the percept preference was the only significant effect, $F_{(1,20)} = 16.23$, $P = 0.001$, indicating that the RTs were significantly faster for the preferred than the nonpreferred percepts in both tasks. Since the bottom-up stimuli were identical in the bistable trials of both tasks, this result suggested that the manipulation of task demands facilitated the processing of preferred percepts through goal-directed top-down attention. Neither the main effect of the type of perceived percepts, $F_{(1,20)} = 0.15$, $P = 0.70$, nor the 2-way interaction, $F_{(1,20)} = 0.09$, $P = 0.77$, was significant, indicating comparable task difficulties between the element-wise EM percept and group-wise GM percept (Fig. 2E).

Task demands modulated ongoing prestimulus alpha frequency

At the behavioral level, we showed that task demands toward the GM vs. EM percepts effectively modulated the temporal window of perceptual grouping (Fig. 2B). Next at the neurophysiological level, we further tested whether task demands employ the prestimulus alpha oscillations to bias subjective perception toward the preferred percepts. We hypothesized that task demand toward temporal integration in the EMT would expand the temporal window for perceptual grouping by slowing down the prestimulus alpha oscillations. It is thus more possible for the 2 frames to fall in the same alpha cycle and be temporally integrated, favoring the EM percepts. In contrast, task demand toward temporal segregation in the GMT would narrow the temporal window for perceptual grouping by accelerating alpha oscillations (Fig. 1A). It is thus more possible for the 2 frames to fall in different alpha cycles and be temporally segregated, favoring the GM percepts.

To test this hypothesis, we analyzed the whole-brain EEG data from 21 participants by comparing the prestimulus IAFs of the task threshold IFI trials between the 2 tasks. Specifically, the alpha-band phase angle time series were extracted via the Hilbert transform on the filtered prestimulus EEG data for both tasks and all channels, and then, the temporal derivative of the resulted time series was computed as the index of the IAFs (see Materials and methods). Consistent with our hypothesis, we found a significantly higher prestimulus alpha frequency in GMT than in EMT, $P < 0.001$, cluster-based correction (Fig. 3A). The significant frequency modulation emerged from ~ 200 ms prior to the stimulus onset of each trial (Fig. 3A and B), indicating that task demands emphasizing temporal integration (EMT) vs. segregation (GMT) indeed modulated the prestimulus ongoing alpha frequency profiles. Specifically speaking,

the task demands toward temporal integration in the EMT down-regulated the background alpha frequency, while task demand toward temporal segregation in the GMT up-regulated the prestimulus ongoing alpha frequency.

To rule out the possibility that the results of the fixed-effects analysis were driven by individual participants, we removed data from 1 participant at a time and recalculated the difference in prestimulus IAF between the EMT and GMT in the rest of the participants. The results showed that the effect persisted regardless of which participant was excluded (Fig. S2, see online Supplementary Material for a color version of this figure), indicating that the current results were unlikely to be driven by individual participants. Moreover, to rule out the possible biases in the estimation of IAFs due to the possible effect of $1/f$ slope, we reconstructed the demodulated time-domain signal with attenuated $1/f$ (Samaha and Cohen 2022) (Fig. S3A, see online Supplementary Material for a color version of this figure, see details in Materials and methods) and then performed the same instantaneous frequency analysis. The result pattern showed consistency (Fig. S3B, see online Supplementary Material for a color version of this figure). Last, to control for the instantaneous frequency bias due to variations in the individual's peak alpha frequency, we centered the filter on ± 2 Hz from the subject-specific alpha peak (Fig. S4, see online Supplementary Material for a color version of this figure, see details in Materials and methods), and the effect persisted (Fig. S5, see online Supplementary Material for a color version of this figure).

In addition, a between-subject correlation between the difference in the PSE IFIs between the 2 tasks and the difference in the IAFs between the 2 tasks was performed. The between-task difference in the PSE IFIs was calculated based on the psychophysical pretest data in each individual subject. The between-task difference in the IAFs in each participant was calculated as the maximal IAF difference between the 2 tasks across all the significant time points within the spatiotemporal cluster (time of interest, see Materials and methods). A marginally significant correlation was found between the 2 measures (Spearman's correlation: $r = -0.416$, $P = 0.060$; Fig. 3C), suggesting that there was a trend that the more the PSE IFIs were shifted between the 2 tasks in an individual, the larger the prestimulus frequency modulation effect between the 2 tasks.

To rule out the possibility that the prestimulus alpha frequency modulation is attributed to the oscillatory power, we performed a statistical test on the prestimulus oscillatory power within the time and channels of interest between the 2 tasks. No significant difference in prestimulus oscillatory power was found between the 2 tasks either for all the frequency bands (from 2 to 30 Hz) or across all the time points for the instantaneous alpha power, all $P_s > 0.05$ (2-tailed paired t-test, uncorrected, see

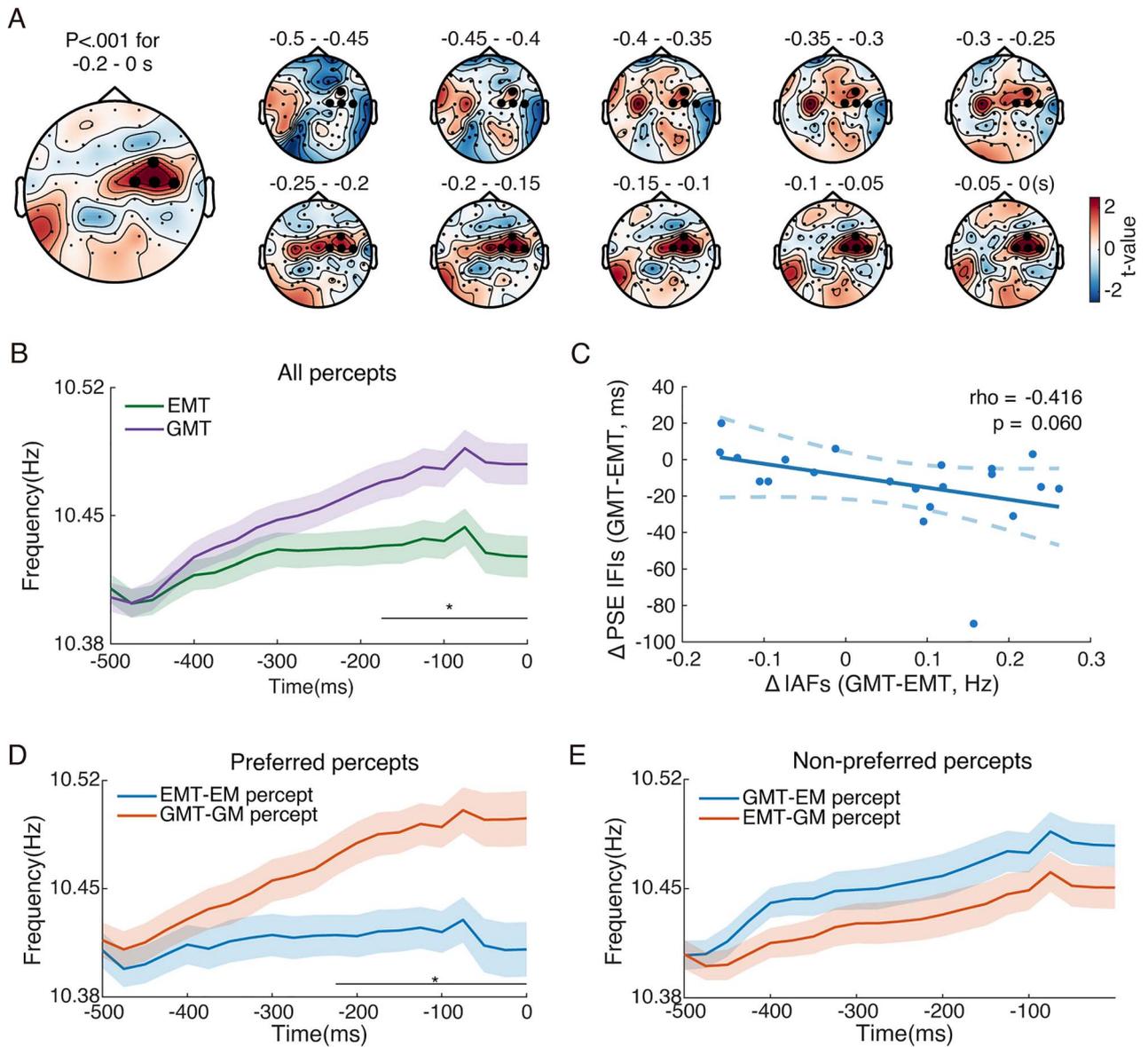


Fig. 3. Attentional modulation effect on the prestimulus IAF. A) Left: averaged EEG topography for the contrast of the instantaneous prestimulus alpha frequency between the EMT and GMT for the significant prestimulus spatiotemporal cluster (cluster-based correction, $P < 0.001$, all time points are relative to the first frame onset). Right: EEG topographies for the contrast of the instantaneous prestimulus alpha frequency between the EMT and GMT over time in steps of 50 ms. Significant channels are indicated by larger black dots. B) IAFs for EMT and GMT averaged over channels of interest (significant channels in A). C) Between-subject correlation between the difference in the instantaneous prestimulus alpha frequency in the main EEG experiment and the difference in PSE IAFs between the EMT and GMT in the psychophysical pretests. D) Instantaneous prestimulus alpha frequency for the preferred percept trials averaged over the region of interest (significant channels in A). E) Same as (D), but for the nonpreferred percept trials. Significant time points are indicated by the horizontal black bar (cluster-based correction, $P < 0.05$). Shaded areas denote ± 1 within-subjects SEM.

Fig. S6, see online [Supplementary Material](#) for a color version of this figure).

The above results suggested that task demands toward temporal integration in the EMT vs. temporal segregation in the GMT effectively modulated the prestimulus alpha frequency. If task demands indeed employed the prestimulus alpha frequency to bias the subjective perception toward the preferred percepts, we further predicted that the prestimulus alpha frequency modulation should be significantly larger for the preferred than nonpreferred

percepts. To test this hypothesis, we further performed a repeated-measures ANOVA on the IAFs within the time of interest by first averaging the IAFs over channel and time of interest, with the type of perceived percepts (EM percept vs. GM percept) and the percept preference (preferred vs. nonpreferred) as the 2 within-subject factors. Neither the main effect of the type of perceived percepts ($F_{(1,20)} = 1.78$, $P = 0.20$) nor the main effect of percept preference ($F_{(1,20)} = 0.09$, $P = 0.77$) was significant. Importantly, a significant interaction effect was revealed,

$F_{(1,20)} = 5.76, P = 0.026$. To further illustrate this interaction effect, we compared the IAFs of the EM and GM percepts across time in the preferred and the nonpreferred conditions, respectively. Consistent with our hypothesis, a significant difference from 225 ms before the stimulus onset was found across participants for the preferred percept trials ($P = 0.017$, cluster-based correction; Fig. 3D), while no significant difference was found for the nonpreferred percept trials (all $P_s > 0.05$, cluster-based correction; Fig. 3E). These results thus suggested that the prestimulus alpha frequency modulation was more pronounced for the preferred than nonpreferred percepts. In other words, when the prestimulus alpha frequency modulation was greater, the perceptual outcome of the bistable stimuli was more likely to be the preferred percepts.

Task demands biased poststimulus neural representations toward preferred percepts

To investigate whether and how task demands modulated the fidelity of neural representations, we employed the time-resolved decoding technique based on regularized linear discriminant analysis to investigate the neural representations in different conditions (Fisher 1936; Mostert et al. 2015; Kok et al. 2017; Shen et al. 2019). The classifier was trained for each time point between all the bistable EM and GM percept trials collapsed over the GMT and EMT (the trial number was matched between all conditions, see [Materials and methods](#)). The classification relied on the projections of neural activity in the high-dimensional activation space into a single-dimensional space of representation profiles, which could maximally separate the neural patterns between the EM and GM percept regardless of the task types (Fig. 4A). If the neural representation was similar to the GM percept, the representation profile was set to be a positive value; vice versa, if the neural representation was similar to the EM percept, the representation profile was set to be a negative value. The larger the difference between the 2 representation profiles, the clearer the separation between the underlying neural representations, and the better fidelity of the neural representations. To first test whether the neural representation in the EM and the GM trials was separable, we categorized all the bistable trials (collapsed over the EMT and GMT trials) into the EM vs. GM percept groups and tested the 2 groups using the trained classifier. The results showed that the neural representations of the EM vs. GM percepts per se could be successfully discriminated, $P = 0.003$, cluster-based correction (Fig. S7, see online [Supplementary Material](#) for a color version of this figure).

We then investigated the separation in neural representations between the GM and EM percepts as a function of the 2 tasks. The working hypothesis was 2-fold: even with the identical bottom-up inputs in the task threshold IFI trials, (i) task demands toward temporal segregation (GMT) vs. integration (EMT) will result in a general bias of neural representations toward the GM vs.

EM percepts after the second frame onset (Fig. 4B); and (ii) since the preferred percepts benefit more from the task demands than the nonpreferred percepts, we predicted a larger separation of neural representations between the GM and EM percepts for the preferred than nonpreferred percepts (Fig. 4C). To test the first hypothesis, we calculated the separation of neural representations in all the task threshold IFI trials between the EMT and GMT. Significant differences in the representation profiles of all the task threshold IFI trials were found between the EMT and the GMT, from 273 to 539 ms after the first frame onset of the Ternus display, $P = 0.03$, cluster-based correction (Fig. 4D). For reference, the onset of the second frame in the task threshold IFI trials varied from 100 to 170 ms after the first frame onset across participants. Therefore, consistent with the first hypothesis, the current results showed that despite the identical bottom-up inputs in the task threshold IFI trials between the GMT and EMT, the poststimulus neural representations of these trials were significantly biased toward the preferred percepts from 103 to 173 ms after the second frame onset. To test the second hypothesis, we subsequently investigated the differential representation profiles between the GM and the EM percepts, depending on whether they were the preferred or nonpreferred percepts. For both the preferred and the nonpreferred percepts, significant differences in the neural representation profile were found between the EM and GM percepts. For the preferred percepts, the significant difference started from 211 ms after the first frame onset, i.e. 41–111 ms after the second frame onset ($P = 0.01$, cluster-based correction; Fig. 4E). For the nonpreferred percepts, the significant difference started from 428 ms after the first frame onset, i.e. 258–328 ms after the second frame onset ($P = 0.02$, cluster-based correction; Fig. 4F). To further test whether there was a significant difference in the GM vs. EM representation profile between the preferred and nonpreferred percept trials, a repeated-measures ANOVA was performed on the representation profile averaged from 100 to 600 ms, with the type of perceived percepts (EM percept vs. GM percept) and the percept preference (preferred vs. nonpreferred) as the 2 within-subject factors. The main effect of the type of perceived percepts was significant, $F_{(1,20)} = 25.37, P = 6.31 \times 10^{-5}$, but the main effect of percept preference was not significant, $F_{(1,20)} = 0.23, P = 0.64$. Critically, the interaction effect was significant, $F_{(1,20)} = 5.08, P = 0.036$, indicating that the separation in the neural representations between the EM and GM percepts was significantly larger for the preferred than nonpreferred percepts. Therefore, consistent with the second hypothesis, the current results showed that task demands modulated the fidelity of neural representations.

Sensory templates were induced before the second frame onset

To directly investigate the temporal dynamics of sensory templates, we further performed the temporal

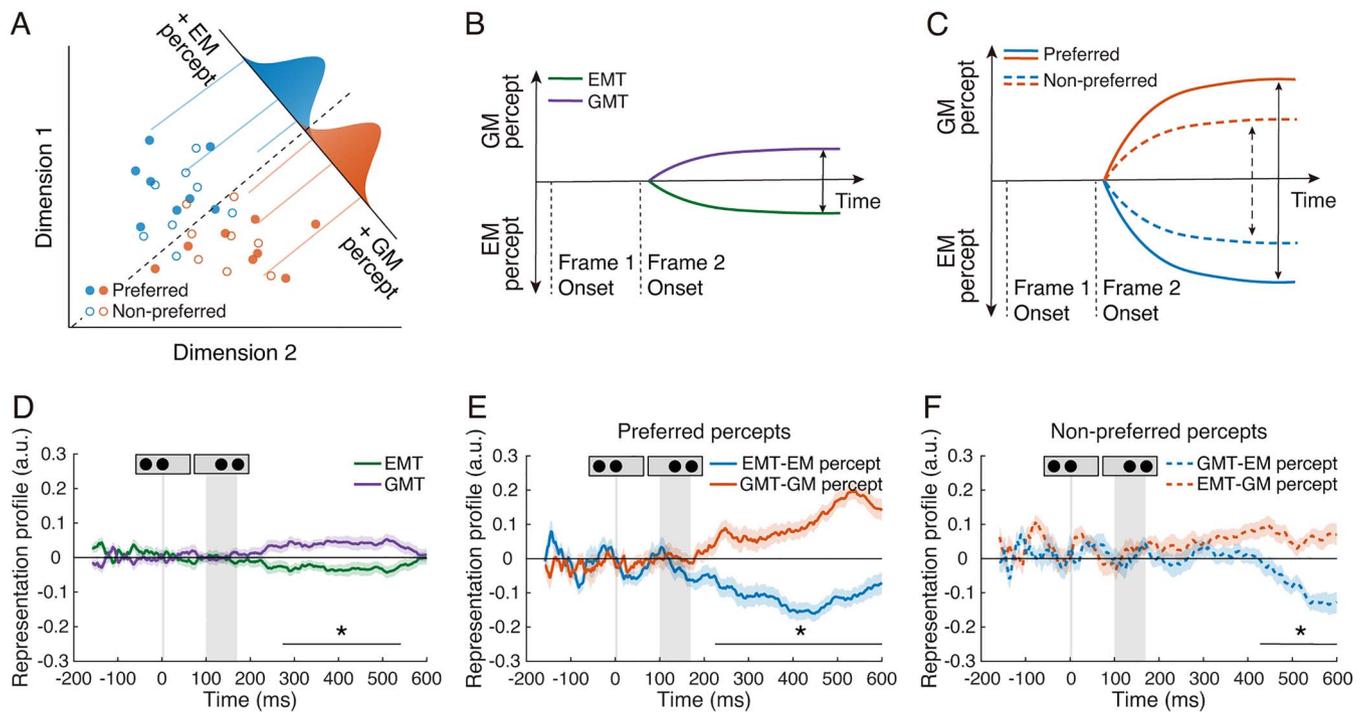


Fig. 4. Hypothesis and results of the time-resolved decoding analysis. A) Hypothetical 2D activation space for the EEG signals representing the EM and GM percepts. The decoding relies on the projections of neural activity in the high-dimensional activation space into a single-dimensional space of representation profiles regardless of the type of the task. A larger difference in representation profile indicates a clearer separation in neural representations. B) Hypothesized effect of attentional modulation on the neural representation profiles for the preferred and the nonpreferred percept trials. C) Hypothesized effect of attentional modulation on the neural representation profiles for the preferred and the nonpreferred percept trials. D) Comparison of the representational profiles in all the bistable trials (collapsed over the preferred and the nonpreferred percepts) between the EMT and GMT. E) Comparison of the representational profiles between the EM and GM percepts for the preferred percept trials. F) Same as (E), but for the nonpreferred percept trials. Significant time points are indicated by the horizontal black bar (cluster-based correction, $P < 0.05$). Shaded areas denote ± 1 within-subjects SEM.

generalization decoding analysis on the GM vs. EM percepts for the preferred and the nonpreferred percept trials, respectively. Instead of training a classifier for each time point and applying the classifier for the same time point in the time-resolved analysis, the temporal generalization analysis trains the classifier on 1 time point and applies the classifier on another time point, which allows direct comparisons of the neural representations between different time points (King and Dehaene 2014). For example, it is possible to train the classifiers at the time points of 100, 200, 300, 400, and 500 ms after the first frame onset of the Ternus display and then apply the trained classifiers on all the time points (Fig. 5A). If the neural representations at 1 time point were similar to those at the trained time point, the trained classifier should perform well; vice versa, if the neural representations were dissimilar to those at the trained time point, the trained classifier should perform poorly.

The current results showed that the classifiers trained on 300, 400, and 500 ms in the preferred percept trials performed well on the time points immediately after the first frame onset and prior to the second frame onset. The classifier trained in nonpreferred percept trials, however, performed poorly over similar time periods (Fig. 5A). Specifically speaking, by systematically adopting this approach on the preferred and the nonpreferred percept

trials, respectively, 2 temporal generalization matrices were obtained (Fig. 5B and C), in which each row corresponded to the time at which the classifier was trained and each column corresponded to the time at which the classifier was tested. Similar to the results in the time-resolved analysis (Fig. 4E and F), the current results showed that the EM and GM percepts could be successfully discriminated for both the preferred, $P = 0.006$, cluster-based correction (Fig. 5B), and the nonpreferred percept trials, $P = 0.049$, cluster-based correction (Fig. 5C). Critically, for the preferred percept, significant separation of neural representations was found from 45 to 115 ms before the second frame onset for classifiers trained on data around 290–360 ms after the second frame onset (Fig. 5B). Since a clearly reportable percept arises only after the presentation of the second frame in the Ternus display, the observed early separation of neural representations for the classifiers trained on the poststimulus data could only be attributed to the fact that a sensory template resembling the neural signals in the later stages of sensory processing was induced immediately after the presentation of the first frame. For the nonpreferred percepts, no significant separation of neural representations was found before the second frame onset, providing no evidence for the existence of sensory templates for the nonpreferred percepts (Fig. 5C). Direct comparisons between the preferred and the

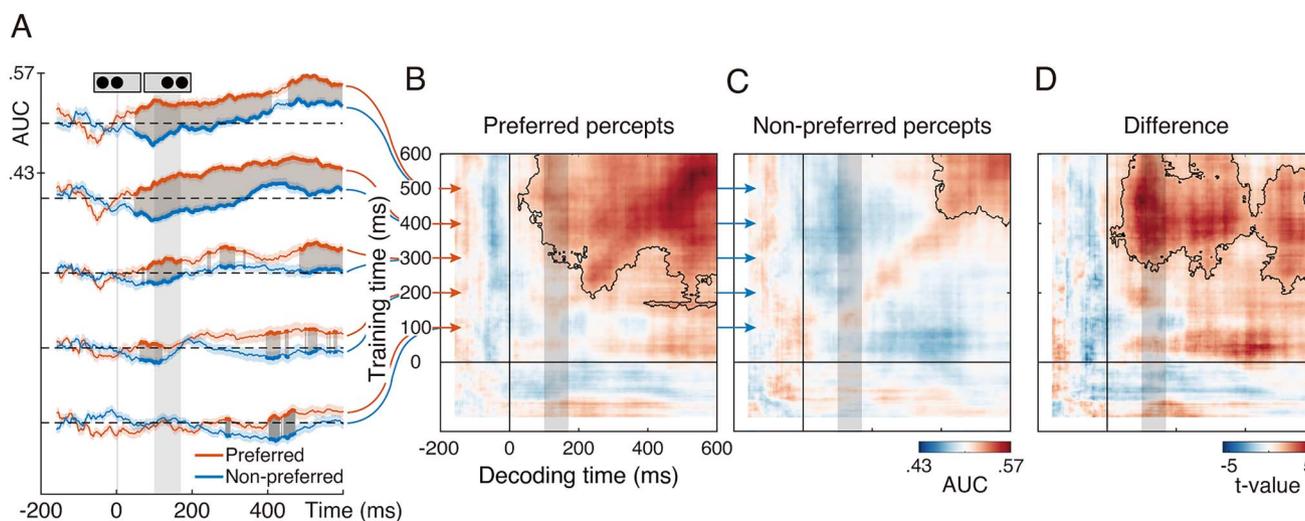


Fig. 5. Results of the temporal generalization analysis. A) Temporal generalized decoding performance for the preferred (red) and the nonpreferred (blue) percept trials. Classifiers were trained on the time points of 100, 200, 300, 400, and 500 ms, and tested on all the time points. Significant time points for the difference between the preferred and the nonpreferred percepts are indicated by gray shadings (uncorrected, $P < 0.05$). B) Temporal generalization matrices for preferred percept trials. C) Temporal generalization matrices for nonpreferred percept trials. D) Temporal generalization matrices for the differential contrast between the preferred and the nonpreferred percept trials. Rows in the images are the time points at which the classifier was trained, and columns are the time points at which the classifier was tested. Color values indicate decoding performance in terms of AUC (B, C) or t-statistics between the AUCs in different conditions (D). Significant differences are indicated by the black contours (cluster-based correction, $P < 0.05$) and the time range for the second frame onset is indicated by the gray rectangle.

nonpreferred conditions further confirmed significant differences from 90 to 160 ms before the second frame onset for the classifiers trained on data around 350 ms after the first frame onset (i.e. 180–250 ms after the second frame onset) ($P = 0.009$, cluster-based correction; Fig. 5D).

Taken together, although explicit EM or GM percepts can be formulated only after the presentation of the second frame in the Ternus display, sensory templates resembling the neural signals of the later sensory processing were induced before the second frame onset (Fig. 5B and D).

Discussion

By manipulating the task demands in 2 types of difficulty-matched tasks with identical bistable Ternus displays, we investigated how the task demands bias subjective perception. Specifically, we asked the following 2 questions in the present study: (i) whether the intrinsic prestimulus alpha-band oscillations were effectively modulated by the 2 types of task demands; and more importantly (ii) whether sensory templates, which carried representational information of the preferred percepts, were induced by different task demands before explicit percepts were formulated.

At the behavioral level, we found that task demands toward the EM (temporal integration) vs. GM (temporal segregation) percept modulated the temporal window for perceptual grouping in the visual system (Fig. 2B). Correspondingly at the neural level, we found that task demands effectively modulated the prestimulus alpha frequency especially for the preferred percepts, with significantly higher prestimulus alpha frequency for the GM

percepts in the GMT than the EM percepts in the EMT (Fig. 3D). The perceptual cycle theory suggests that our perception samples the outside world in a discrete and cyclic manner (VanRullen 2016; White 2018): 2 sequential events that fall in a single perceptual cycle of a critical rhythm would be grouped together, while the same 2 events falling in different perceptual cycles would be perceived as 2 separate percepts. Therefore, a higher frequency of a critical rhythm would result in a shorter temporal window for information integration, while a lower frequency would result in a longer temporal window. Accumulating evidence from EEG, magnetoencephalography, and transcranial alternating current stimulation studies consistently suggested a causal role of the intrinsic alpha frequency in gating the temporal window of temporal integration in the visual system (Kristoffer-son 1967; Cecere et al. 2015; Samaha and Postle 2015; Minami and Amano 2017; Shen et al. 2019; Zhang et al. 2019). The current results further suggested that top-down attention was able to voluntarily modulate the prestimulus intrinsic alpha frequency, to bias the subsequent perception toward the preferred percepts (see also in Wutz et al. (2018)). Specifically speaking, task demands toward temporal separation in the GMT increased the prestimulus alpha frequency so that the 2 frames had a higher probability to fall in different alpha cycles, and the subjective perception was biased toward the GM percepts. On the other hand, task demands toward temporal integration in the EMT decreased the prestimulus alpha frequency so that it was more possible for the 2 frames to fall in the same alpha cycle, rendering subjective perception toward the EM percepts. Also, the potential confounds by the different physical properties of the stimuli (Cohen 2014a) and task difficulty (Haegens et al. 2014;

Babu Henry Samuel et al. 2018) were ruled out since the 2 tasks were matched in task difficulty and identical stimuli were adopted.

Please note, consistent with previous studies using similar IAF analysis (Cohen 2014b; Samaha and Postle 2015; Wutz et al. 2018; Shen et al. 2019), we observed small frequency changes between the 2 tasks (~ 0.1 Hz in all the trials, Fig. 3B, and 0.2 Hz in the preferred percept trials, Fig. 3D). Technically, this small effect may result from the fact that the EEG alpha frequency signal is the average alpha frequency signal from all neural populations, including both task-relevant and task-irrelevant populations. Thus, the observed effect may be attenuated by the noise from neuronal populations not involved in the task (Samaha and Postle 2015; Shen et al. 2019). Moreover, in terms of the location of the effect, the EEG topographies showed that the effect centered around the central electrodes (Fig. 3A) rather than the occipital electrodes as in our previous study (Shen et al. 2019). There are 2 possible explanations for this difference. The first possibility is that the effect could be actually located in the occipital region, similar to the previous study (Wutz et al. 2018), since EEG topographies do not accurately reflect the location of the source. The second possibility is that the observed frequency modulation effect indeed occurs in the higher level regions (e.g. parietal regions), or even a combination of the higher level (e.g. frontoparietal regions) and the lower level (e.g. occipital regions) regional effects. However, due to the poor spatial resolution of the EEG, it is difficult to tell where the effect actually originates, and imaging techniques with higher spatial and temporal resolution (e.g. intracranial recordings) will be necessary for future studies.

By performing multivariate pattern analysis on neural activity in the identical bistable Ternus displays between the 2 difficulty-matched tasks, we further investigated the temporal dynamics of the sensory template. The results showed that although the poststimulus neural representations of the subjectively perceived EM vs. GM percepts could be efficiently decoded (Fig. 4), the 2 types of percepts could not be clearly decoded during the prestimulus phase (Figs. 4 and 5). By using the temporal generalization technique, we further showed that the sensory templates, which resembled the neural signals evoked by the later explicitly reported percepts, were induced prior to the presentation of the second frame (Fig. 5). Since the mere presentation of the first frame does not provide sufficient sensory evidence to generate explicit percepts, the results suggested that the sensory templates are generated the very first (still ambiguous) evidence of sensory inputs. The induced sensory templates are then adopted to bias the subsequent perception of ambiguous stimuli toward the preferred percepts.

The biased competition model of attention proposes that persistent pre-activation of sensory-specific neural coding raises the baseline for subsequent processing of relevant inputs, facilitating target processing (Desimone and Duncan 1995). Accordingly, it has been

documented, in the typical cue-target paradigm, that top-down attention toward a particular spatial or nonspatial object feature leads to prestimulus baseline increases both in the sensory cortex (Chelazzi et al. 1993; Luck et al. 1997; Kok et al. 2017) and the higher order frontoparietal cortex (Buschman and Miller 2007; Peelen and Kastner 2014). Moreover, the sensory templates, which carry representational information of the attended stimuli, can be decoded out during the prestimulus phase in the object/feature selective sensory cortex (Stokes et al. 2009; Kok et al. 2014, 2017). Note, all the previous relevant studies adopted the classical cue-target paradigm, in which: (i) the prestimulus cue period is relatively short (from hundreds of milliseconds to seconds), and (ii) behavioral goals vary frequently on a trial-by-trial base. In this case, it might be the most optimal way for the short prestimulus attentional signals to carry representational information of the attended stimuli, i.e. the sensory template, which in turn facilitates the subsequent attentional selection.

In our daily life, however, human beings often engage in an ongoing behavioral task (e.g. search for a target person or object) for a prolonged period of time (Duncan and Humphreys 1989; Wolfe et al. 1989; Wolfe 2015). In the latter case, it has been shown that the sensory templates do not persistently exist during the prestimulus phase but rather is revealed only shortly after the presentation of the bottom-up stimuli (Myers et al. 2015). Therefore, the top-down attentional signals do not necessarily always carry representational information of the attended stimuli especially when persistently maintaining a representational attentional template throughout a prolonged attentional status is computationally expensive. Moreover, since the sensory templates exist to facilitate attentional selection of the bottom-up sensory inputs, an sensory template after the stimulus onset does not justify such a facilitatory role. One critical possibility is: the sensory template could be induced after the very first and still ambiguous sign of sensory evidence, which then guides the subsequent sensory processing. Our current results supported this notion by showing that the sensory template is induced after the first frame onset, i.e. after the initial sensory evidence has been presented, and before the second frame onset, i.e. before the explicit percepts are formulated.

In summary, our results showed that both the prestimulus frequency of alpha-band oscillations and the sensory templates were modulated according to task demands during the bistable apparent motion perception. The prestimulus alpha frequency provides the optimal temporal window for temporal integration vs. segregation, based on the current task demands. Once the very first sensory evidence falls in the temporal window predefined by the frequency of prestimulus alpha-band oscillations, concrete sensory templates of the preferred percepts are induced, biasing subjective perception toward the preferred percepts.

Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

Funding

This work was supported by grants from the National Natural Science Foundation of China (31871138, 32071052, 32000741, 32000785), the Guangdong Natural Science Foundation (2021A1515011100, 2020A1515110223, 2021A1515011185), and the Guangzhou Basic Research Program (202102020761, 202102020274).

Conflict of interest statement: The authors declare no competing financial interests.

References

- Babu Henry Samuel I, Wang C, Hu Z, Ding M. The frequency of alpha oscillations: task-dependent modulation and its functional significance. *NeuroImage*. 2018;183:897–906.
- Baldauf D, Desimone R. Neural mechanisms of object-based attention. *Science*. 2014;344(6182):424–427.
- Buschman TJ, Miller EK. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*. 2007;315(5820):1860–1862.
- Cecere R, Rees G, Romei V. Individual differences in alpha frequency drive crossmodal illusory perception. *Curr Biol*. 2015;25(2):231–235.
- Chelazzi L, Miller EK, Duncan J, Desimone R. A neural basis for visual search in inferior temporal cortex. *Nature*. 1993;363(6427):345–347.
- Cohen MX. *Analyzing neural time series data: theory and practice*. MIT Press; 2014a
- Cohen MX. Fluctuations in oscillation frequency control spike timing and coordinate neural networks. *J Neurosci*. 2014b;34(27):8988–8998.
- Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci*. 2002;3(3):201–215.
- Desimone R, Duncan J. Neural mechanisms of selective visual attention. *Annu Rev Neurosci*. 1995;18(1):193–222.
- Duncan J, Humphreys GW. Visual search and stimulus similarity. *Psychol Rev*. 1989;96(3):433–458.
- Eimer M. The neural basis of attentional control in visual search. 2014;18(10):10.
- Fisher RA. The use of multiple measurements in taxonomic problems. *Ann Eugenics*. 1936;7(2):179–188.
- Haegens S, Händel BF, Jensen O. Top-down controlled alpha band activity in somatosensory areas determines behavioral performance in a discrimination task. *J Neurosci*. 2011;31(14):5197–5204.
- Haegens S, Cousijn H, Wallis G, Harrison PJ, Nobre AC. Inter- and intra-individual variability in alpha peak frequency. *NeuroImage*. 2014;92:46–55.
- King J-R, Dehaene S. Characterizing the dynamics of mental representations: the temporal generalization method. *Trends Cogn Sci*. 2014;18(4):203–210.
- Klimesch W, Doppelmayr M, Russegger H, Pachinger T, Schwaiger J. Induced alpha band power changes in the human EEG and attention. *Neurosci Lett*. 1998;244(2):73–76.
- Kok P, Failing MF, de Lange FP. Prior expectations evoke stimulus templates in the primary visual cortex. *J Cogn Neurosci*. 2014;26(7):1546–1554.
- Kok P, Mostert P, de Lange FP. Prior expectations induce prestimulus sensory templates. *Proc Natl Acad Sci U S A*. 2017;114(39):10473–10478.
- Kristofferson AB. Successiveness discrimination as a two-state, quantal process. *Science*. 1967;158(3806):1337–1339.
- Luck SJ, Chelazzi L, Hillyard SA, Desimone R. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J Neurophysiol*. 1997;77(1):24–42.
- Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods*. 2007;164(1):177–190.
- Maunsell JHR, Treue S. Feature-based attention in visual cortex. *Trends Neurosci*. 2006;29(6):317–322.
- Minami S, Amano K. Illusory jitter perceived at the frequency of alpha oscillations. *Curr Biol*. 2017;27(15):2344–2351.e4.
- Moran J, Desimone R. Selective attention gates visual processing in the extrastriate cortex. *Science*. 1985;229(4715):782–784.
- Mostert P, Kok P, de Lange FP. Dissociating sensory from decision processes in human perceptual decision making. *Sci Rep*. 2015;5:18253.
- Myers NE, Rohenkohl G, Wyart V, Woolrich MW, Nobre AC, Stokes MG. Testing sensory evidence against mnemonic templates. *elife*. 2015;4:e09000.
- Nobre AC, Kastner S. Attention: time capsule 2013. In: *The Oxford handbook of attention*; 2014. pp. 1201–1222
- Nobre AC, Stokes MG. Remembering experience: a hierarchy of time-scales for proactive attention. *Neuron*. 2019;104(1):132–146.
- Oostenveld R, Fries P, Maris E, Schoffelen J-M. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci*. 2011;2011:156869.
- Peelen MV, Kastner S. A neural basis for real-world visual search in human occipitotemporal cortex. *Proc Natl Acad Sci U S A*. 2011;108(29):12125–12130.
- Peelen MV, Kastner S. Attention in the real world: toward understanding its neural basis. *Trends Cogn Sci*. 2014;18(5):242–250.
- Pikler J. *Sinnesphysiologische Untersuchungen*. Johann Ambrosius Barth; 1917.
- Reynolds JH, Chelazzi L. Attentional modulation of visual processing. *Annu Rev Neurosci*. 2004;27(1):611–647.
- Samaha J, Cohen MX. Power spectrum slope confounds estimation of instantaneous oscillatory frequency. *NeuroImage*. 2022;250:118929.
- Samaha J, Postle BR. The speed of alpha-band oscillations predicts the temporal resolution of visual perception. *Curr Biol*. 2015;25:2985–2990.
- 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 U S A*. 2015;112(27):8439–8444.
- Samuel IBH, Wang C, Hu Z, Ding M. The frequency of alpha oscillations: task-dependent modulation and its functional significance. 2018;10.
- Sauseng P, Klimesch W, Stadler W, Schabus M, Doppelmayr M, Hanslmayr S, Gruber WR, Birbaumer N. A shift of visual spatial attention is selectively associated with human EEG alpha activity. *Eur J Neurosci*. 2005;22(11):2917–2926.
- Shen L, Han B, Chen L, Chen Q. Perceptual inference employs intrinsic alpha frequency to resolve perceptual ambiguity. *PLoS Biol*. 2019;17(3):e3000025.
- Stokes M, Thompson R, Nobre AC, Duncan J. Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proc Natl Acad Sci*. 2009;106(46):19569–19574.

- Szczepanski SM, Konen CS, Kastner S. Mechanisms of spatial attention control in frontal and parietal cortex. *J Neurosci*. 2010;30(1):148–160.
- Ternus J. Experimentelle Untersuchungen über phänomenale Identität. *Psychol Forsch*. 1926;7(1):81–136.
- Treutwein B, Strasburger H. Fitting the psychometric function. *Percept Psychophys*. 1999;61(1):87–106.
- 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.
- VanRullen R. Perceptual cycles. *Trends Cogn Sci*. 2016;20(10):723–735.
- Varoquaux G, Raamana PR, Engemann DA, Hoyos-Idrobo A, Schwartz Y, Thirion B. Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. *NeuroImage*. 2017;145:166–179.
- White PA. Is conscious perception a series of discrete temporal frames? *Conscious Cogn*. 2018;60:98–126.
- Wolfe JM. Visual search. In: *The handbook of attention*. Cambridge (MA): Boston Review; 2015. pp. 27–56.
- Wolfe JM, Cave KR, Franzel SL. Guided search: an alternative to the feature integration model for visual search. *J Exp Psychol Hum Percept Perform*. 1989;15(3):419–433.
- Wutz A, Melcher D, Samaha J. Frequency modulation of neural oscillations according to visual task demands. *Proc Natl Acad Sci U S A*. 2018;115(6):1346–1351.
- Wyart V, Nobre AC, Summerfield C. Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proc Natl Acad Sci U S A*. 2012;109(9):3593–3598.
- Zhang Y, Zhang Y, Cai P, Luo H, Fang F. The causal role of α -oscillations in feature binding. *PNAS*. 2019;116(34):17023–17028.