Using fast visual rhythmic stimulation to control inter-hemispheric phase offsets in visual areas

Qiaoyu Chen[1,2], Danying Wang[1,2], Kimron L. Shapiro[1,2], & Simon Hanslmayr[1,2,3]

1) School of Psychology, University of Birmingham, UK
2) Centre for Human Brain Health, University of Birmingham, UK
3) Institute for Neuroscience and Psychology, University of Glasgow, UK

Contact Information / Corresponding Author

Simon Hanslmayr

simon.hanslmayr@glasgow.ac.uk

Manuscript details

Number of pages:	24
Number of figures:	five figures and five supplementary figures
Abstract word count:	250
Introduction word count:	1256
Discussion word count:	1139

Conflict of interest

The authors declare no competing financial interests.

Abstract

Spike timing dependent plasticity (STDP) is believed to be important for neural communication and plasticity in human episodic memory, but causal evidence is lacking due to technical challenges. Rhythmic sensory stimulation that has been used to investigate causal relations between oscillations and cognition may be able to address this question. The challenge, however, is that the frequency corresponding to the critical time window for STDP is gamma (~40 Hz), yet the application of rhythmic sensory stimulation has been limited primarily to lower frequencies (<30 Hz). It remains unknown whether this method can be applied to precisely control the activation time delay between distant groups of neurons at a millisecond scale. To answer this question and examine the role of STDP in human episodic memory, we simulated the STDP function by controlling the activation time delay between the left and right visual cortices during memory encoding. This was achieved by presenting flickering (37.5 Hz) movie pairs in the left and right visual fields with a phase lag of either 0, 90, 180 or 270 degrees. Participants were asked to memorize the two movies within each pair and the association was later tested. Behavioral results revealed no significant difference in memory performance across conditions with different degrees of gamma phase synchrony. Yet importantly, our study showed for the first time, that oscillatory activity can be driven with a precision of 6.67 ms delay between neuronal groups. Our method hereby provides an approach to investigate relations between precise neuronal timing and cognitive functions.

Keywords: STDP | gamma | episodic memory | SSVEPs

1. Introduction

Brain oscillations have been shown to be crucial for efficient information transmission in neural networks (Buzsáki, 2010; Draguhn & Buzsáki, 2004; Fries, 2015). Via this central communication function oscillations are key to a host of cognitive functions, such as attention and memory. A currently important question is whether oscillations are causally important for cognition, or whether they merely are a non-causal by-product of cognitive processing (Hanslmayr et al., 2019). One way to answer this question is to perturb oscillations via rhythmic stimulation and test whether such perturbation induces a change in behavior. An efficient way to drive oscillatory activity is via rhythmic sensory stimulation (e.g., flickering a visual stimulus). In the current study, we present a novel application of rhythmic sensory stimulation in a memory experiment. We demonstrate that rhythmic visual stimulation in the gamma frequency range (37.5 Hz) is capable of controlling the phase delay between left and right visual cortex in a temporally finely-grained manner.

Traditionally, rhythmic visual stimulation has been used to tag cognitive processes, which would otherwise be difficult to observe via so-called steady-state visual evoked potentials (SSVEPs, Müller et al., 2003, 2006). SSVEPs have been particularly successfully applied in the study of attention and memory. This is because in comparison with transient evoked potentials, SSVEPs provide a substantially longer time window over which a cognitive process can be monitored (Capilla et al., 2011), allowing for the tracking of attention both temporally and spatially (Adamian et al., 2020; Müller et al., 2006; Störmer et al., 2014). Frequency tagging (or SSVEPs) has also been used in memory research to elucidate reactivation of early visual signatures of a specific memory. For instance, Wimber et al. (2012) found rapid reinstatement of tagged frequencies during successful memory retrieval (but see Lewis et al., 2018; Price & Johnson, 2018 for difficulties in replicating). Together, these studies demonstrate the power of using rhythmic sensory stimulation as an effective means to study neural correlates of cognitive processes.

Recent studies have gone beyond a mere correlational approach and used rhythmic sensory stimulation to drive a specific oscillation to induce a change in the associated behavior (Garcia-Argibay et al., 2019; Henry & Obleser, 2012; Mathewson et al., 2012; Papalambros et al., 2017; Spaak et al., 2014). For instance, interregional synchrony in the theta band is believed to be of crucial importance for memory formation. To test this hypothesis, two recent studies induced theta phase synchrony or asynchrony between visual and auditory cortices to examine if synchrony affects memory formation. Visual and auditory regions were driven at phase offsets of 0 (synchronous), 90, 180 or 270 (asynchronous) degrees. Importantly, both studies found better memory for multisensory (i.e. audio-visual) memories on trials in which auditory and visual cortices were stimulated synchronously compared to asynchronously stimulated trials (Clouter et al., 2017; Wang et al., 2018). These studies demonstrate a causal role of theta phase synchrony for memory formation. In particular, they show that subsequent remembering and forgetting depends on the relative timing of sensory information, and that this timing can be controlled through sensory stimulation. However, the extent to which this approach applies to faster frequencies (i.e., gamma), which have also been shown to play a role in synchronization, and to coordinating brain areas within the same sensory domain (i.e., visual cortex) is unclear.

Fine-grained timing of neural cell assemblies, as achieved by synchronization in the gamma frequency band, is important for effective information transmission between neurons (Fries, 2015). Precise timing is critical because neurons in the brain integrate input over time, with the rate of relaxation of the membrane potential dictating the length of the temporal window. For neocortical principal cells that time window is typically between 10 – 30 ms (100 – 33 Hz), therefore fine-grained

temporal synchronization is necessary for an upstream cell assembly to drive a down-stream neuron (Buzsáki, 2010). Equally important, synaptic plasticity has been shown to be critically dependent on the time delay between the firing of an up-stream and a down-stream neuron, which is termed 'spike timing dependent plasticity' or STDP. In STDP, the efficiency of synaptic modification declines exponentially as a function of time delay between the activation of a pre- and postsynaptic neuron. The optimal time window for synaptic modification is very narrow due to the exponential decay (Bi & Poo, 1998). In support of this contention, a rodent study found that spikes must co-occur in a time window of approximately 25 ms to facilitate synaptic modification, corresponding to gamma frequency at ~40 Hz (Wespatat et al., 2004). Indeed, phase synchronization at the gamma frequency range is optimal for the facilitation of STDP (Fell & Axmacher, 2011). Action potentials tend to appear in the depolarized phase of local field potential (Fries, 2005) which has been observed in the theta and gamma band (Jacobs et al., 2007; Vinck et al., 2010). As a result, synchronizing the phase of the local field potential from two distant brain regions would promote the induction of long-term potentiation (LTP; Axmacher et al., 2006; Jutras & Buffalo, 2010). Critically, higher frequencies mean tighter coupling, therefore gamma synchronization in particular leads to more precise coupling of action potentials which in turn facilitates STDP (Abbott & Nelson, 2000; Caporale & Dan, 2008; Fell & Axmacher, 2011; Jutras & Buffalo, 2010). Modulating the degree of synchrony in gamma phase between neuron assemblies arguably manipulates the time delay of action potentials and therefore, STDP between them. STDP supposedly underlies memory formation but its role in human episodic memory has been rarely studied because of the technical challenge to non-invasively control the timing of neural assemblies. Rhythmic sensory stimulation as employed in the present study may open a new approach to address these questions, thus providing a potential link between highly invasive research in animals and non-invasive human studies.

In the current study, we aim to address two questions. The first is a technical question, namely whether it is possible that SSVEPs can be used to control neural activity in the left and right visual cortex with high temporal precision corresponding to a quarter cycle (i.e., 6.67 ms) of a gamma oscillation. Given the high temporal precision to which we aimed, we chose the hemifield approach, which controls neural activity between left and right visual hemifield. We chose this because the transduction time should vary little between the left and right visual cortex, while this is not the case between auditory and visual sensory regions. The second question is whether such stimulation has an impact on human episodic memory formation. To address these questions, we presented two sinusoidal flickering movies in the left and right visual hemifields (Fig. 1A). A sine wave of 37.5 Hz was used to modulate the luminance of the movies. To induce a phase offset of 0°, 90°, 180°or 270° (corresponding

to a time lag of 0, 6.67, 13. 33 and 20 ms) between the left and right visual cortices, sinusoidal flickering movies were presented at the left and right visual field with a phase lag of either 0°, 90°, 180° or 270° (Fig. 1B). We predicted that, if STDP plays a role in episodic memory, subsequent recall of the association of movie pairs should exponentially decrease with increasing phase lags. Although the 90° and 270° offset condition may look similar from a purely circular perspective, they are very different from each other from a temporal perspective, i.e., the difference in time which indeed is 6.67 ms vs. 20 ms. According to our STDP hypothesis, this time delay, although subtle, has dramatic consequences for plasticity due to the exponential decay ((Bi & Poo, 1998).

2. Materials and Methods

2.1 Participants

Thirty-eight participants were recruited (mean age = 22.11; age range = 18-32; 71.1% female; all right-handed). Participants received course credit or financial reimbursement in return for their participation. Two participants were excluded for excessive horizontal eye movements (see below). One participant was excluded due to EEG recording interruption. This left 35 participants for further analysis (mean age = 22.37; age range = 18-32; 68.6% female; all right-handed). Ethical approval was granted by the Research Ethics Committee at the University of Birmingham (ERN_15-0335), complying with the Declaration of Helsinki.

2.2 Stimuli

384 randomly paired three-second movie clips were employed as movie pairs (N = 192) for associative memory. Half of the movies were drawn from the same pool as those used in Experiment 3 of Clouter et al. (2017), while the remainder were trimmed documentaries downloaded from an online resource (https://www.youtube.com) with creative commons license. Four additional movie pairs were used in a practice block. Consistent with movies used in previous work (Clouter et al., 2017; Wang et al., 2018), all movies were emotionally neutral human/natural activities.

Movies were resized to 360(W) x 288(H) pixels, with the frame rate increased to 150 Hz from 25 Hz by replicating each frame 6 times with in-house scripts written in MATLAB (R2017a; The MathWorks, Inc., Natick, MA, USA). Movies were randomly paired once resulting in 192 unique associations. The same pairs of movies were used for each participant.

All movies were luminance-modulated from 0% to 100% by a 37.5 Hz sine wave, but with different onset depending on the condition. Specifically, within each pair of movies, one always (across

conditions and participants) served as the leading movie, with an onset of 0°, while the other one served as the trailing movie, with an onset of one of the following degrees (conditions evenly distributed): 0°, 90°, 180° or 270°. This results in four phase lag conditions, i.e., 0°, 90°, 180° or 270° (see Fig. 1B) between the leading movie and the trailing movie, with 48 trials in each of the condition. Phase lag conditions assigned to each group of 48 movies pairs were counterbalanced across participants.

All behavioral tasks were programmed using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007) running on MATLAB (R2015b; The Mathworks, Natick, MA, USA). To make sure that the stimulus was presented at the frequency we need, precise timing of the stimulus was verified with a photodiode before the experiment was run (see Fig. S5 and Supplemental Material).

2.3 Experimental procedure

Participants were seated in a testing room and requested to complete forms for safety screening and to provide consent after they were informed with the procedure of the study and prepared for EEG data collection. Details of the memory task were explained to participants and a practice block was used to ensure familiarity with the memory task. Participants were seated at a distance of 60 cm from the screen center with their heads resting on a chin support. A web camera was used by the experimenter to verify the subject's head position and compliance during the task. The experiment consisted of three tasks; (i) a memory task, (ii) a synchrony judgement task and (iii) an EOG calibration task with instructions provided before each task. At the end, to allow for source analysis and to provide more precise estimation for electrode interpolations, 3D geometric locations of each electrode were recorded using a Polhemus FASTRAK device (Colchester, Vermont, USA) and Brainstorm (Tadel, Baillet, Mosher, Pantazis, & Leahy, 2011) implemented in MATLAB (R2018a; The MathWorks, Inc., Natick, MA, USA).

Based on our previously published results (Clouter et al., 2017; Wang et al., 2018), the present experiment aimed at testing 24 participants with a minimum number of 32 trials per condition, which would lead to a power of 97.7% (alpha level=0.05, one-tailed paired sample t-test). However, given concerns on the low signal to noise ratio of gamma entrainment for this study, as compared to theta entrainment in previous studies (Clouter et al., 2017; Wang et al., 2018), we initially decided to double the sample size (48 participants), however, data collection had to stop because of the pandemic. A post-hoc power analysis performed on the null results presented here revealed that at least 219 subjects are needed to reveal a significant difference in memory performance. Therefore, we decided to write the present report based on the data so far collected (35 participants with a minimum of 48 trials per condition).

193 *2.3.1 Memory Task*

194 The memory task was comprised of six blocks, each containing an encoding phase, a distractor
195 phase and a retrieval phase. For each block, 32 pairs of movies were shown during the encoding phase
196 and participants were asked to remember the pairs as they would be tested in the retrieval phase. Block
197 order was fully randomized for each participant. During intervals between blocks, memory
198 performance was shown to the participant on the screen to motivate good performance. After blocks
199 of poor performance participants were encouraged to take a break.

200 During the encoding phase, for each trial, a pair of movies was displayed simultaneously on
201 the left and right (see Fig. 1A) of a 21-inch CRT screen (150 Hz refresh rate and a resolution of 800
202 width x 600 high pixels) through an nVidia Quadro K600 graphics card (875 MHz graphics clock,
203 1024 MB dedicated graphics memory; Nvidia, Santa Clara, CA, USA). Movie pairs were resized to
204 250(W) x 200(H) pixels and were aligned horizontally to the screen center, with a visual angle of 8.4°
205 (movie center to the screen center; see Fig. 1A). The background of the experimental interface was set
206 to grey. For trials with phase offsets of 90, 180 or 270 degrees each pair of movies consisted of one
207 leading movie and one trailing movie (for 0-degree phase offsets both movies flickered synchronously
208 hence neither movie was leading nor trailing). The left vs. right position of leading movies presented
209 on the screen was randomly assigned for each trial. Participants were instructed to maintain fixation at
210 the screen center and attend the movie pairs without shifting gaze. Participants were encouraged to
211 link the movies via stories or visual imagery to help them learn the associations. Participants were
212 informed that their memory of the associations would be tested in the later retrieval section. After
213 watching each pair of movies, participants responded on a scale from 1 to 5 to which extent that the
214 movies fit with each other, i.e., how easy it was for participant to link them (1: no fit, 5: perfect fit).
215 This judgement task was implemented to maintain the engagement of participants during encoding.

216 Following the encoding phase, a 30-second distractor task was implemented to prevent active
217 rehearsal of the movie pairs. Participants were asked to overtly count backwards from a random 3-
218 digit number displayed on the screen by subtracting three each time.

219 In the final test phase, a full movie was shown which acted as a cue. Only old movies (i.e.,
220 movies that have been shown during encoding in this block) were presented. For phase offset
221 conditions 90, 180, and 270 the leading movie was always served as the cue. Below the cue, four
222 screenshots which were taken from old encoding movies (only trailing movies within the same block)
223 were shown as options. The participant selected which of the 4 screenshots they thought was presented

with the cue during encoding. Each trailing movie served as lures for exactly three times and as matched movie (correct answer) for one time.

2.3.2 Synchrony Judgement

The synchrony judgement task tested whether participants were able to perceive the phase offsets between the movies, i.e., whether they could tell the difference between synchronous (0 degree) and asynchronous pairs (90, 180 or 270 degrees). To this end, 60 movies pairs were drawn randomly from the memory encoding phases (see Fig. 1A). Participants were asked to judge in a two alternative forced choice procedure whether a given movie pair was flickering synchronously (0 phase lag), or not (90, 180 and 270 phase lags). Responses were requested after presentation of the movies.

2.3.3 EOG Calibration Task

The EOG calibration task provided a template for monitoring the occurrence of horizontal eye movements. The data from this task was used to exclude trials where overt eye movements were made to the movies. In the EOG calibration task participants were asked to visually track a black cross appearing randomly to the left or right of the fixation cross at a visual angle of 8.44◦. After each movement, the fixation cross maintained its position for a duration between 800 ms to 1200 ms before moving back to the center, which marked the onset of the next trial. There were 80 trials in total. For half of the trials the fixation cross moved to the left, and for the other half it moved to the right. 10% of trials served as catch trials during which participants were asked to report a color change of the cross (black to red) by pressing the space button as soon as possible. This ensured that participants maintained attention throughout the calibration task.

2.4 Data Analysis

2.4.1 Behavioral Data

To test whether the manipulation of phase offset had consequence on episodic memory, memory accuracy was compared between 0°, 90°, 180° and 270° phase offset conditions using a one-way repeated measurements ANOVA. Trials with exceeding HEOG was excluded for this analysis. Further, to rule out perceptual factors that might have affected episodic memory, a sensitivity index *d'* was calculated for each participant to test whether they were able to tell synchronous from asynchronous movie pairs. This was estimated as follows:

$d' = Z$ (hit rate) - Z (false alarm rate),

where the Z function was deployed by the normal inverse cumulative distribution function in MATLAB (R2017a; The MathWorks, Inc., Natick, MA, USA), considering synchronous movie pairs as signal and asynchronous pairs as noise. One-sample *t*-test was adopted to examine whether *d'* was statistically different from zero.

2.4.2 EOG calibration

The purpose of the EOG calibration was to exclude trials on which subjects made horizontal eye-movements, as such trials are detrimental to tracking the EEG response to the left/right visual hemifields independently. It is feasible to use EOG data for tracking eye movements as a result of the linear relationship between voltage and visual angle (Acuna et al., 2014). This procedure enabled a threshold for each participant to be calculated individually for the purpose of excluding trials as defined by the following procedure.

First, the EOG data from both the EOG calibration task and the encoding phase of the memory task was preprocessed using a procedure similar to that used for the EEG data. EOG data were epoched to 2000 ms before and after the onset of the fixation cross. For the memory encoding phase, the data were epoched to 2000 ms before and 5000 ms after onset of the movie pairs. Both datasets were then low-pass filtered at 30 Hz and resampled at 512 Hz.

Second, erroneous eye movements were manually excluded for further analysis via trial-by-trial visual inspection. Such errors included eye movements that were made too early (i.e., before the fixation cross jumped) or random saccades (i.e., eye movements not reflecting tracking of the fixation cross). To detect the voltage gradients generated by saccades from the center to the left/right center of movie position (at visual angle of 8.44◦), the first derivative was calculated using the 'diff' function in Matlab. The resulting data was then aligned according to the largest peaks appeared within 600 ms upon stimulus onset and then averaged. A threshold was calculated by taking 50% of the averaged peak value, thus corresponding to a saccade of 4.22◦. Peaks of EOG gradients were then calculated in the same way for the encoding phase of the memory task for each individual trial. Trials which exceeded the threshold (of 50%) were excluded from further behavioral and EEG analysis.

2.4.3 EEG recording and preprocessing

EEG data were collected via a 128-channel BioSemi ActiveTwo system. EOG recording was completed by one additional electrode placed 1 cm to each of the lateral canthus and 1 cm below the left eye, respectively. Online EEG data were sampled to 2048 Hz by the BioSemi ActiView software.

The Fieldtrip toolbox for EEG analysis (Oosenveld, Fries, Maris, & Schoffelen, 2011) was used for EEG data preprocessing. Data were first epoched from 2000 ms before and 5000 ms after stimulus onset and then bandpass filtered from 1 to 100 Hz. Line noise was removed from the raw data by bandstop filters between 48-52 Hz and 98-102 Hz. Before ICA (independent component analysis), data was downsampled to 512 Hz, followed by the removal of noisy EEG channels and trials with muscle artefacts by manual inspection. ICA components indicating horizontal and vertical eye movements and cardiac activity were removed from data. Triangulation of nearest neighbors calculated by individually recorded electrode positions were used for interpolation of rejected channels. Finally, data were re-referenced to the average reference and trials with artefacts were rejected by visual inspection.

2.4.3 Phase Offset Analysis

To confirm whether the phase stimulation is driving brain oscillations between left and right visual cortices at the corresponding degree of phase offset, phase comparisions were conducted by the following steps at both individual and group level.

2.4.3.1.ROI selection and ITPC calculation

First, for each participant, one electrode from left and right hemifield, respectively, was identified as ROI that responds strongest to the rhythmic visual stimulation. The basic rationale for the determination of these ROIs was that, in each of the left/right visual cortex, there existed one electrode that mostly responds to the visual stimulation from the contralateral visual field. Inter-trial phase coherence (ITPC) should be strongest across trials recorded from this electrode with the leading movies presented in the contralateral visual field. For example, across trials with leading movies presented at the right visual field, the ITPC was expected to be strongest in an electrode over the left visual cortex. This was because the leading movies were always starting with a same onset at 0°. On the contrary, the phases for the trailing movies (i.e. the associated stimuli) were equally distributed across 0°, 90°, 180° and 270° and thus lead to low ITPC. Since the leading movies were split between left and right hemifiled, this approach allowed us to determine one ROI (i.e. one electrode) for each hemisphere of the visual cortex. Details of this analysis is given below.

For each participant, we separated trials according to the location of leading movies. In the following statements, LeadLvsf and LeadRvsf are used to represent the conditions with the leading movies in the left and right visual fields, respectively.

ITPC from 20 to 50 Hz in steps of 2.5 Hz was calculated using a dpss multitaper (1 taper for each frequency of interest, from -1s to 3.75s upon stimulus onset, with a time window of 0.5s in width, 128 channels) time-frequency transformation based on multiplication in the frequency domain. A single subject example with ITPC averaged at the band of 37.5 ± 2.5 Hz for the LeadRvsf and LeadLvsf condition is displayed in Fig. 2A and B, respectively. The averaged ITPC was calculated for a selected time window of interest from 1-2s upon stimulus onset. The selection of this time window was to align with previous work (Clouter et al., 2017; Wang et al., 2018) with a similar paradigm in lower frequencies, but results are also shown for the entire epoch (see Fig. S2). The contrast between LeadRvsf and LeadLvsf (Rvsf-Lvsf) reveal the electrodes that responded most strongly to visual stimulations from the left and right visual field (see Fig. 2C). This analysis was implemented for all participants, such that one electrode (as highlighted in Fig. 2C as an example) from each side of the visual cortex was defined as a subject-specific ROI for later analysis (except for the analysis confirming the specificity of ITPC difference at 37.5 Hz, which is decribed in the next paragraph).

At the group level ITPC values were averaged across all subjects for LeadRvsf and LeadLvsf conditions, seperately. The difference between the two conditions was statistically assessed by means of a two-tailed (alpha = 0.025) paired-sample permutation test (number of randomizations = 2000) at the frequency of interest (37.5 Hz) across the time duration from 1s – 2s following stimulus onset. Specifically, the significance probability was performed by Monte Carlo method and multiple comparisons were corrected by “cluster”. To further confirm whether this difference was specific for the frequency of interest, we implemented another independent analysis based on posterior electrodes rather than the subject specific ROI mentioned above. We firstly averaged time-frequency structured ITPC values across 25 posterior electrodes (see electrodes highlighted in Fig. 3A for LeadRvsf and Fig.3B for LeadLvsf) within each hemifield. The same ROI was used for each subject. This was performed at single subject level for conditions with leading movies presented at the contralateral and ipsilateral visual fields, respectively. Then, the difference between the averaged ITPC values for the contralateral and ipsilateral conditions (contra - ipsilateral) were compared against zero, from 20-50 Hz (in steps of 2.5 Hz) and 1-2 second (in steps of 50 ms) upon stimulus onset, by using a one-tailed paired sample permutation test (number of randomizations = 2000). The p-value estimation and correction for multiple comparisons were Monte Carlo and “cluster”, same as the above analysis. Parameters for the calculation of the effect size was as follows. For each participant, ITPC difference was calculated by subtracting the ipsilateral (to the leading movie position) ITPC from the contralateral ITPC at the left and right posterior electrodes (25 electrodes at each hemifield, 50 electrodes in total, as highlighted in Fig.3A and B) at each time point of toi (1-2s upon stimulus onset, 0.05 s in steps).

Thereafter, ITPC difference averaged across 37.5 ± 2.5 Hz was selected for the calculation of effect size.

2.4.3.2. SSVEPS and Measuring Phase Offsets between Left and Right Visual Cortex

As reported above, one electrode corresponding to the left/right visual cortex was selected for each subject as a ROI, which is reflecting the strongest response to the rhythmic visual stimulation. To examine whether the phase lags between left and right ROI were consistent with the visual modulation (i.e., 0, 90, 180 and 270 phase lag conditions), we extracted the instantaneous phase for each phase lag condition. Since gamma is notorious for its susceptibility to noise, the extraction of instantaneous phases was based on SSVEPs rather than on single-trials (Fries et al., 2008). To increase trial number for the timelocked analysis, ROI data were swapped between left and right hemifield for the trials in the LeadLvsf condition. As a consequence, the LeadLvsf condition also has the leading movie stimulation projected to the left hemifield, similar to the LeadRvsf condition. This hemisphere swapping procedure allowed us to include trials in both LeadLvsf and LeadRvsf for the timelocked analysis, with the left and right side of visual cortex corresponding to stimulation from the leading movies and trailing movies, respectively. As a result, for each participant, there was one SSVEP for each combination of the following two factors: leading/trailing movies by 0/90/180/270 phase lags, resulting a total number of 8 SSVEPs. The number of trials for this timelocked analysis ranged between 32 and 47.

At the individual level, a bandpass filter from 35-40 Hz was then applied to the SSVEPs. To extract instaneous angles, Hilbert transformation was applied to the SSVEPs and the resulting angles were thereafter unwrapped. Instantaneous phase differences between left and right ROIs were calculated for each phase lag condition (0, 90, 180 and 270 degrees), with the time of interest (toi) from 1s to 2s upon stimulus onset, to avoid any possible contamination from stimulus onset and offset (see Clouter et al., 2017; Wang et al., 2018 for a similar rationale).

For group analysis, eight SSVEPs were calculated based on grand averaged SSVEPs across participants for each combination of the conditions (i.e. 2 by 4), leading/trailing movies by 0/90/180/270 phase lags. The instanteous angles and phase differences were generated with the same process as stated above for individuals but based on the grand averaged SSVEPs. The V test implemented in the Circular Statistics Toolbox (Berens, 2009) was used to test whether the instantaneous phase differences were uniformly distributed at the same phase as the visual stimulation conditions (0, 90, 180 and 270 phase lags).

3. Results

3.1 Behavioral performance

Recall accuracy did not differ significantly across the four phase offset conditions of 0°, 90°, 180° and 270° (N =35; one-way repeated measurements ANOVA; F (3,102) = 0.612, p = .609, ω^2 = 0, see Fig. 4 and Fig. S1). However, it should be noted that there was a weak trend for memory accuracy being slightly higher at the 0° phase lag condition as compared to the other three phase lag conditions (90, 180 and 270). This trend was more obvious visually when showing the accuracy for the four phase lag conditions by subtracting individual mean performance across conditions (see Fig. 4; and see Fig. S1 for data without the subtraction). Mean performance ratio for each condition ranges from 0.643 (180°) to 0.659 (0°), with *SD* ranging from 0.142 to 0.169, showing median difficulty of the task (M = 0.659, SD = 0.159; M = 0.646, SD = 0.143; M = 0.643, SD = 0.142; M = 0.652, SD = 0.169 for conditions of 0, 90, 180 and 270 degrees, respectively). The overall performance appears to be higher than previous work using a multisensory association paradigm (Clouter et al., 2017; Wang et al., 2018). Bayesian repeated measures ANOVA revealed a Bayes Factor (B10) of 0.076, suggesting that the null model outperformed our hypothesis. The largest effect size was found between 0° and 180° offset conditions (Cohen's d = 0.2331) and according to this effect size, at least 219 subjects are needed to reveal a significant difference in memory performance between 0° and 180°.

To determine if subjects were able to distinguish between synchronous and asynchronous movie pairs, *d'* was calculated for 34 participants out of 35. One subject was excluded for the analysis of the synchrony judgement task because the hit rate was zero. A one-sample *t*-test suggested that *d'* for the synchrony task was significantly different from zero, t (33) = 14.92, p < 0.1e-15, indicating that participants were well able to distinguish between synchronous and asynchronous trials.

3.2 ROI and ITPC

After trial rejection due to EOG thresholding and EEG preprocessing, the total number of trials at single subject level across all conditions survived ranges from 135-185 trials (out of 192), with a mean of 164.63 trials. The number of trials for LeadLvsf condition ranges from 68-95 trials (mean = 84.37), and that for LeadRvsf condition ranges from 67-96 trials (mean = 80.26). For each phase offset condition, the range of number of trials is 30 to 47, with a mean of 40.71, 40.97, 41.29 and 41.66 for 0°, 90°, 180° and 270° offset condition, respectively.

ITPC in the frequency band from 35 to 40Hz was utilized to identify ROIs in each visual hemifield for each subject (see Methods section and Fig. 2). As expected, stimulation from leading

movies displayed in the right visual field (LeadRvsf) caused the strongest ITPC in the contralateral (left) hemifield, while at the ipsilateral (right) hemifield ITPC is lower for trailing movies with various phase onset asynchronies (Fig. 2A). A similar result occurred for the LeadLvsf condition (see Fig. 2B). To cancel out factors that may potentially affect ITPC beyond the conditioning, such as common responses to the fixation on the screen center, a contrast between the two conditions (LeadRvsf – LeadLvsf) was calculated which revealed the most responsive electrodes for phase modulation observed in each visual hemifield (See highlighted electrodes in Fig. 2C).

The averaged topographic pattern of ITPC across participants in the range of 37.5 ± 2.5 Hz in the left and right hemifields showed a highly consistent pattern across subjects (Fig. 3A and B). Cluster-based paired-sample permutation tests between LeadRvsf vs. LeadLvsf topographic distribution revealed a significant difference during the interval from 1s – 2s upon stimulus onset (Fig. 3C, $p_{corrected} < 0.001$). Grand averaged ITPC from 20-50 Hz across electrodes highlighted in Figure 3A and B indicate strongest ITPC at the frequency around 37.5 ± 2.5 Hz for both LeadRvsf and LeadLvsf conditions (Fig. 3D and E, see Fig. S2 for ITPC in other time windows). The specificity for 37.5 Hz was further confirmed by cluster-based paired-sample permutation test by comparing the resulting ITPC difference (contralateral - ipsilateral) with zero (see Fig. 3F). The Cohen's effect size of ITPC difference at posterior electrodes (as highlighted in Fig. 3A and B) across 37.5 ± 2.5 Hz at toi (1s-2s upon stimulus onset) indicated such difference was robust (d = 0.4145, M = 0.0162, SD = 0.0389). These results suggested that the aligned phase activity was specific to frequencies around the entrained frequency, 37.5 Hz, $p_{corrected} < 0.001$.

3.3 SSVEPs and entrained gamma phase offset

Band-pass filtered (37.5 ± 2.5 Hz) SSVEPs averaged across subjects are shown in figure 5. These results indeed show different phase lags between leading and trailing movies for the 4 different phase lag conditions (0, 90, 180 and 270 degrees) at the group level (Fig. 5A). Closer inspection of the SSVEPs (see Fig. 5B) suggests that the phase offsets between leading and trailing movies closely followed the phase lags induced by the different stimulation conditions (0, 90, 180 and 270 degrees).

In order to statistically test whether the phase offsets between the SSVEPs elicited by the leading/trailing movies were congruent with the phase offsets induced by our stimulation, the instantaneous phase differences between trailing and leading movies were calculated (see Fig. 5C; for phase offsets at individual level for each condition, see Fig. S3; for phase offsets at group level based on group ROI, see Fig. S4). The mean phase differences with 95% confidence interval are 10.88° ± 1.14°, 101.06° ± 0.84°, 176.09° ± 0.66° and 258.49° ± 1.60° for 0, 90, 180 and 270 phase offset

conditions, respectively. A V-test was performed on these phase differences, which tests the nonuniformity with a known mean direction of circular data. The V-test confirmed that the phases from 1s -2s upon stimulus onset were uniformly distributed around their entrained phase offset, with $p = 0$ for all conditions.

Together, the EEG results demonstrated the feasibility of inducing phase lags between the left and right visual cortex at a minimal time interval of 6.67 ms by hemispheric visual stimulation.

4. Discussion

Modulating brain oscillations via rhythmic sensory stimulation opened a new avenue to potentially draw causal links between oscillations and cognitive functions. However, such usage has so far been limited to low frequencies (Hanslmayr et al., 2019), while the feasibility of stimulating at higher frequencies (>30 Hz) has rarely been explored. STDP is a prominent theory of a synaptic plasticity mechanism by which synaptic efficiency declines exponentially with increasing time delay between pre- and postsynaptic neurons (Bi & Poo, 1998; Song, 2000; Caporale & Dan, 2008), and was observed within a critical time window corresponding to 40 Hz in an study using animal models (Wespatat et al., 2004). Similar observations of STDP in human participants has proved to be challenging. As a result, little is known about whether STDP has a similar function and time course in humans compared to animals (Mansvelder et al., 2019), and indeed whether it plays an important role in human episodic memory.

Using rhythmic sensory stimulation, we aimed to drive gamma oscillations at different phase offsets between left and right visual cortex at a high temporal precision (i.e., phase delays of 90 degrees or 6.67 ms) to investigate a potential role of STDP in human episodic memory. Despite observing no effects on memory performance, EEG results demonstrate the feasibility of inducing phase offsets at high frequencies, thus preserving the precise timing of the rhythmic sensory stimulation. Previous studies demonstrated that EEG/MEG signals show steady-state potentials up to 90 Hz (Herrmann, 2001; Zhigalov et al., 2019). However, our study shows, for the first time, that phase offsets between two regions in the brain can be controlled at a high temporal resolution. We believe this is important as it opens new avenues of investigating the neural and behavioral impact of subtle timing differences between large-scale neuronal assemblies in humans.

Despite the successful manipulation of phase lags between left and right visual cortex at 0, 90, 180 and 270 degree of 37.5 Hz, only a weak trend of enhanced memory by synchronizing gamma between hemifields at 0 degree (see Fig. 4) was observed. While the electrophysiological results look very clear,

the behavioral data are considerably noisier. Comparing with the effect size of 0.978 (24 participants and 32 trials per condition) based on our previous published results, the largest effect size reached was found between 0° and 180° offset conditions (Cohen's d = 0.2331, 35 participants and 48 trials per condition) in the current experiment. Based on the post-hoc effect size, at least 219 subjects are needed to reveal a significant difference in memory performance between 0° and 180°. Therefore, it is likely that more behavioral data is required to show an effect, while relatively fewer data are needed to show an electrophysiological effect.

In theory, one possible explanation for this null result is that the chosen frequency of stimulation may not be ideal for STDP in human visual cortex during memory encoding. Indeed, a wide STDP window determined by cellular mechanisms like NMDA receptors and voltage-dependent Ca2+ channels was revealed in vivo and vitro studies (Caporale & Dan, 2008), consistent with the observed time window for neocortical principal cells (10-30 ms, corresponding to 33-100 Hz). More importantly, fast and slow gamma oscillations are likely to serve distinct functions in the hippocampus. In rodents, the medial entorhinal cortex (for information inputs) and CA3 (essential for information storage) are phase-locked to CA1 at a fast (~65-140 Hz) and a slow gamma (~25-50 Hz) rhythm, respectively (Colgin, 2015; Colgin et al., 2009). A recent intracranial EEG study shows similar evidence in humans with increased fast (60-80 Hz) gamma power indicating successful episodic memory encoding, and enhanced slow (40-50 Hz) gamma power indicating successful memory retrieval in the hippocampus (Griffiths et al., 2019). Information encoding and retrieval therefore are likely to be implemented by different gamma frequency bands. Given the above evidence, it is conceivable that fast gamma around 60-80 Hz is critical for STDP in visual cortex for sending information to higher level structures (such as the hippocampus), which would explain why stimulating at slow gamma (37.5 Hz) did not modulate memory encoding in our study.

An alternative explanation of the absence of behavioral effects could be that gamma phase modulation affects implicit memory, rather than episodic memory (which is inherently explicit; Tulving, 1972). Critically, the hippocampus has long been believed to be pivotal for episodic memory (Milner, 1966), given that patients with bilateral damage to the hippocampus show intact implicit memory but impaired episodic memory (Bechara et al., 1995). While EEG results suggest successful gamma phase modulation at visual cortex, it remains unclear if and to what extent the hippocampus is affected by rhythmic stimulation. Indeed, SSVEPs appear to become more focal as the frequency increases (Zhigalov et al., 2019). If this is the case, implicit memory, which is not dependent on hippocampal function, would be more likely to be modulated by gamma. Moreover, given that the stimulation in our study is limited to a single modality, it is possible that hippocampus, as the binding

center of affluent information, contributes less to unimodal binding. Consistent with this notion, evidence from TMS and fMRI studies strongly suggests that priming (a form of implicit memory) is mediated by sensory areas that process the primed feature, such as color and location (Kristjánsson & Campana, 2010). More interestingly, uni-sensory entrainment at 40 Hz was found to exert alternation on visual perception (Elliott & Müller, 1998; Helfrich et al., 2014), suggesting that gamma synchrony plays a role in visual binding. In summary, rhythmic visual stimulation at gamma may have influenced unimodal implicit memory as it relies on sensory cortex and is independent from hippocampal involvement.

There is evidence, however, that run counter to the above assertion. Neurons in the hippocampus of mice fire when auditorily stimulated at 40 Hz (Martorell et al., 2019), indicating that sensory stimulation is able to modulate hippocampus activity; whether this is applicable to humans, however, remains unknown. It is also worth noting that Martorell et al.'s study only found multisensory stimulation to improve cognition. To further our knowledge regarding non-invasive protocols for probing neuronal activation, a better understanding of how rhythmic sensory stimulation affects sensory cortex and downstream neural assemblies needs to be developed.

In summary, although no memory effect was found by gamma phase stimulation, the EEG results strongly suggest the capability of gamma entrainment to modulate brain oscillations among sensory areas with high temporal precision. Our paradigm provides a non-invasive way to manipulate neural synchrony at high temporal resolution between two brain regions. We believe this method makes it possible for future studies to investigate the role of timing at high frequencies for behavioral and neural processes in human subjects. Although the present application of this method was to investigate the role of STDP on human memory formation, other applications in other cognitive domains (i.e., attention) are conceivable.

Acknowledgements

Authors declare that they have no conflict of interest. This research was supported by grants from the European Research Council (https://erc.europa.eu/, Consolidator Grant 647954; to S.H.); Wolfson Foundation and Royal Society (https://royalsociety. org/grants-schemes-awards/grants/wolfson-researchmerit/; to S.H.); The Economic Social Sciences Research Council (https://esrc.ukri.org/, ES/R010072/1; to S.H.). The first author is funded by the China Scholarship Council (CSC).

References

Abbott, L. F., & Nelson, S. B. (2000). Synaptic plasticity: taming the beast. *Nature neuroscience*, 3(11), 1178-1183.

Acuna, O. V., Aqueveque, P., & Pino, E. J. (2014). Eye-tracking capabilities of low-cost EOG system. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, 610–613. https://doi.org/10.1109/EMBC.2014.6943665

Adamian, N., Andersen, S. K., & Hillyard, S. A. (2020). Parallel attentional facilitation of features and objects in early visual cortex. *Psychophysiology*, *57*(3), 1–11. https://doi.org/10.1111/psyp.13498

Axmacher, N., Mormann, F., Fernández, G., Elger, C. E., & Fell, J. (2006). Memory formation by neuronal synchronization. *Brain Research Reviews*, *52*(1), 170–182. https://doi.org/10.1016/j.brainresrev.2006.01.007

Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., & Damasio, A. R. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science*, 269(5227), 1115-1118.

Berens, P. (2009). CircStat: a MATLAB toolbox for circular statistics. *J Stat Softw, 31*(10), 1-21.

Bi, G. Q., & Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of Neuroscience*, *18*(24), 10464–10472. https://doi.org/10.1523/jneurosci.18-24-10464.1998

Brainard, D. H. (1997) The Psychophysics Toolbox, *Spatial Vision 10*:433-436.

Buzsáki, G. (2010). Neural Syntax: Cell Assemblies, Synapsembles, and Readers. *Neuron*, *68*(3), 362–385. https://doi.org/10.1016/j.neuron.2010.09.023

Capilla, A., Pazo-Alvarez, P., Darriba, A., Campo, P., & Gross, J. (2011). Steady-state visual evoked potentials can be explained by temporal superposition of transient event-related responses. *PLoS ONE*, *6*(1). https://doi.org/10.1371/journal.pone.0014543

Caporale, N., & Dan, Y. (2008). Spike Timing–Dependent Plasticity: A Hebbian Learning Rule. *Annual Review of Neuroscience*, *31*(1), 25–46. https://doi.org/10.1146/annurev.neuro.31.060407.125639

Clouter, A., Shapiro, K. L., & Hanslmayr, S. (2017). Theta Phase Synchronization Is the Glue that Binds Human Associative Memory. *Current Biology*, *27*(20), 3143-3148.e6. https://doi.org/10.1016/j.cub.2017.09.001

Colgin, L. L. (2015). Do slow and fast gamma rhythms correspond to distinct functional states in the hippocampal network? *Brain Research*, *1621*, 309–315. https://doi.org/10.1016/j.brainres.2015.01.005

Colgin, L. L., Denninger, T., Fyhn, M., Hafting, T., Bonnevie, T., Jensen, O., Moser, M. B., & Moser, E. I. (2009). Frequency of gamma oscillations routes flow of information in the hippocampus. *Nature*, *462*(7271), 353–357. https://doi.org/10.1038/nature08573

Draguhn, A., & Buzsáki, G. (2004). Neuronal Oscillations in Cortical Networks. *Science*, *304*(June), 1926–1930.

Elliott, M. A., & Müller, H. J. (1998). Synchronous information presented in 40-Hz flicker enhances visual feature binding. *Psychological Science*, *9*(4), 277–283. https://doi.org/10.1111/1467-9280.00055

Fell, J., & Axmacher, N. (2011). The role of phase synchronization in memory processes. *Nature Reviews Neuroscience*, *12*(2), 105–118. https://doi.org/10.1038/nrn2979

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence.

Trends in Cognitive Sciences, *9*(10), 474–480. https://doi.org/10.1016/j.tics.2005.08.011

Fries, P., Scheeringa, R., & Oostenveld, R. (2008). Finding gamma. *Neuron*, 58(3), 303-305.

Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron*, *88*(1), 220–235. https://doi.org/10.1016/j.neuron.2015.09.034

Garcia-Argibay, M., Santed, M. A., & Reales, J. M. (2019). Binaural auditory beats affect long-term memory. *Psychological Research*, *83*(6), 1124–1136. https://doi.org/10.1007/s00426-017-0959-2

Griffiths, B. J., Parish, G., Roux, F., Michelmann, S., van der Plas, M., Kolibius, L. D., Chelvarajah, R., Rollings, D. T., Sawlani, V., Hamer, H., Gollwitzer, S., Kreiselmeyer, G., Staresina, B., Wimber, M., & Hanslmayr, S. (2019). Directional coupling of slow and fast hippocampal gamma with neocortical alpha/beta oscillations in human episodic memory. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(43), 21834–21842. https://doi.org/10.1073/pnas.1914180116

Hanslmayr, S., Axmacher, N., & Inman, C. S. (2019). Modulating Human Memory via Entrainment of Brain Oscillations. *Trends in Neurosciences*, *42*(7), 485–499. https://doi.org/10.1016/j.tins.2019.04.004

Helfrich, R. F., Knepper, H., Nolte, G., Strüber, D., Rach, S., Herrmann, C. S., Schneider, T. R., & Engel, A. K. (2014). Selective Modulation of Interhemispheric Functional Connectivity by HD-tACS Shapes Perception. *PLoS Biology*, *12*(12). https://doi.org/10.1371/journal.pbio.1002031

Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(49), 20095–20100. https://doi.org/10.1073/pnas.1213390109

Herrmann, C. S. (2001). Human EEG responses to 1-100 Hz flicker: Resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental Brain Research*, *137*(3–4), 346–353. https://doi.org/10.1007/s002210100682

Jacobs, J., Kahana, M. J., Ekstrom, A. D., & Fried, I. (2007). Brain oscillations control timing of single-neuron activity in humans. *Journal of Neuroscience*, *27*(14), 3839–3844. https://doi.org/10.1523/JNEUROSCI.4636-06.2007

Jutras, M. J., & Buffalo, E. A. (2010). Synchronous neural activity and memory formation. *Current Opinion in Neurobiology*, *20*(2), 150–155. https://doi.org/10.1016/j.conb.2010.02.006

Kleiner, M., Brainard, D., & Pelli, D. (2007). *What's new in Psychtoolbox-3?*.

Kristjánsson, Á., & Campana, G. (2010). Where perception meets memory: A review of repetition priming in visual search tasks. *Attention, Perception, & Psychophysics, 72*(1), 5-18.

Lewis, A. G., Schriefers, H., Bastiaansen, M., & Schoffelen, J. M. (2018). Assessing the utility of frequency tagging for tracking memory-based reactivation of word representations. *Scientific Reports*, *8*(1), 1–12. https://doi.org/10.1038/s41598-018-26091-3

Mansvelder, H. D., Verhoog, M. B., & Goriounova, N. A. (2019). Synaptic plasticity in human cortical circuits: cellular mechanisms of learning and memory in the human brain? *Current Opinion in Neurobiology*, *54*, 186–193. https://doi.org/10.1016/j.conb.2018.06.013

Martorell, A. J., Paulson, A. L., Suk, H. J., Abdurrob, F., Drummond, G. T., Guan, W., Young, J. Z., Kim, D. N. W., Kritskiy, O., Barker, S. J., Mangena, V., Prince, S. M., Brown, E. N., Chung, K., Boyden, E. S., Singer, A. C., & Tsai, L. H. (2019). Multi-sensory Gamma Stimulation Ameliorates Alzheimer's-Associated Pathology and Improves Cognition. *Cell*, *177*(2), 256-271.e22.

https://doi.org/10.1016/j.cell.2019.02.014

Mathewson, K. E., Prudhomme, C., Fabiani, M., Beck, D. M., Lleras, A., & Gratton, G. (2012). Making waves in the stream of consciousness: Entraining oscillations in EEG alpha and fluctuations in visual awareness with rhythmic visual stimulation. *Journal of Cognitive Neuroscience*, *24*(12), 2321–2333. https://doi.org/10.1162/jocn_a_00288

Milner, B. (1966). Amnesia following operation on the temporal lobes. *Amnesia*.

Müller, M. M., Andersen, S., Trujillo, N. J., Valdés-Sosa, P., Malinowski, P., & Hillyard, S. A. (2006). Feature-selective attention enhances color signals in early visual areas of the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(38), 14250–14254. https://doi.org/10.1073/pnas.0606668103

Müller, M. M., Malinowski, P., Gruber, T., & Hillyard, S. A. (2003). Sustained division of the attentional spotlight. *Nature*, *424*(6946), 309–312. https://doi.org/10.1038/nature01812

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience, 2011*.

Papalambros, N. A., Santostasi, G., Malkani, R. G., Braun, R., Weintraub, S., Paller, K. A., & Zee, P. C. (2017). Acoustic enhancement of sleep slow oscillations and concomitant memory improvement in older adults. *Frontiers in Human Neuroscience*, *11*(March), 1–14. https://doi.org/10.3389/fnhum.2017.00109

Pelli, D. G. (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies, *Spatial Vision* 10:437-442.

Price, M. H., & Johnson, J. D. (2018). Failure to reactivate salient episodic information during indirect and direct tests of memory retrieval. *Brain Research*, *1699*, 9–18. https://doi.org/10.1016/j.brainres.2018.06.031

Song, S., Miller, K. D., & Abbott, L. F. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature neuroscience, 3*(9), 919-926.

Spaak, E., de Lange, F. P., & Jensen, O. (2014). Local entrainment of alpha oscillations by visual stimuli causes cyclic modulation of perception. *Journal of Neuroscience*, *34*(10), 3536–3544. https://doi.org/10.1523/JNEUROSCI.4385-13.2014

Störmer, V. S., Alvarez, G. A., & Cavanagh, P. (2014). Within-hemifield competition in early visual areas limits the ability to track multiple objects with attention. *Journal of Neuroscience*, *34*(35), 11526–11533. https://doi.org/10.1523/JNEUROSCI.0980-14.2014

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Computational intelligence and neuroscience, 2011*.

Tulving, E. (1972). Organization of memory. *Episodic and semantic memory*.

Vinck, M., Lima, B., Womelsdorf, T., Oostenveld, R., Singer, W., Neuenschwander, S., & Fries, P. (2010). Gamma-phase shifting in awake monkey visual cortex. *Journal of Neuroscience*, *30*(4), 1250–1257. https://doi.org/10.1523/JNEUROSCI.1623-09.2010

Wang, D., Clouter, A., Chen, Q., Shapiro, K. L., & Hanslmayr, S. (2018). Single-trial phase entrainment of theta oscillations in sensory regions predicts human associative memory performance. *Journal of*

Neuroscience, *38*(28), 6299–6309. https://doi.org/10.1523/JNEUROSCI.0349-18.2018

Wespatat, V., Tennigkeit, F., & Singer, W. (2004). Phase sensitivity of synaptic modifications in oscillating cells of rat visual cortex. *Journal of Neuroscience*, *24*(41), 9067–9075. https://doi.org/10.1523/JNEUROSCI.2221-04.2004

Wimber, M., Maaß, A., Staudigl, T., Richardson-Klavehn, A., & Hanslmayr, S. (2012). Rapid memory reactivation revealed by oscillatory entrainment. *Current Biology*, *22*(16), 1482–1486. https://doi.org/10.1016/j.cub.2012.05.054

Zhigalov, A., Herring, J. D., Herpers, J., Bergmann, T. O., & Jensen, O. (2019). Probing cortical excitability using rapid frequency tagging. *NeuroImage*, *195*(December 2018), 59–66. https://doi.org/10.1016/j.neuroimage.2019.03.056

722 ## Appendices

723 *Verification of precise timing of the stimulus*

724 A photodiode was attached to the top left of the CRT screen. Each movie (3-second) was
725 displayed at the top left of the screen with an interval of 1 second. Data was recording at a sample rate
726 of 1000 and digitized by the National Instruments USB-6343 multifunction I/O device (see
727 https://www.ni.com/pdf/manuals/377874a.pdf). Since no trigger was implemented during the
728 recording, the trials were defined using the following process. A threshold of 0.7 was used to defined
729 37.5 Hz peaks. Since a quarter cycle of 37.5 Hz is around 6.67 ms, considering the potential noise (e.g.,
730 from the sampling position), at least 4 ms of peaks were expected in each cycle of 37.5 Hz. Hence, the
731 point, from which was followed by at least 150 (37.5 times 4) peaks detected in every second of three
732 continuous seconds, was defined as the stimulus onset time. 286 out of 384 trials were detected using
733 this criterium. Time frequency analysis (Morlet wavelet) with a time window from -0.5 to 3.5 upon
734 stimulus onset, in step of 0.02 second, foi from 1 to 200 Hz, in step of 1 Hz revealed a peak at 37.5 Hz
735 (see Fig. S5).

736 ## Legends of figures and supplemental figures

Fig. 1. Experimental procedure. The experiment contained 192 trials, evenly divided into 6 blocks. Each block (32 trials) consisted of an encoding phase, a distractor task and a retrieval task. (A) Encoding phase. A pair of movies was displayed simultaneously on the left and right of 8.44° off to the center of a 21-inch CRT screen with refresh rate of 150 Hz, for a duration of 3 second. Participants were encouraged to make associations between the two movies as their memory of associations would be tested. **(B)** Modulation of pairs of movies. Movies were luminance-modulated by a 37.5 Hz sine wave but with various onset depending on the phase lag conditions (0, 90, 180 or 270 degrees). One of the movies within each pair was defined as leading movie, because it always began with onset at 0° and the other was trailing movie, beginning with onset evenly distributed at 0°, 90°, 180° and 270°. Sine waves utilized for luminance modulation for leading (blue) and trailing movies (red) showed a phase lag of 0, 90, 180 and 270 degree for the corresponding experimental condition. Time lags between leading and trailing movies were 0, 6.67, 13.33 and 20 ms for 0, 90, 180 and 270 phase lag conditions, simulating the activation time delay between pre- and postsynaptic in STDP. The minimal time lag between leading and trailing movies for asynchrony conditions was only 6.67 ms. (**C**) Retrieval phase. The memory test was carried out after a 30-second distractor task. All 32 pairs of movies showed at encoding phase were tested for once. During the test, participants were asked to recall the associated movie cued by leading movies. Four options as a static frame taken from trailing movies were displayed below the cue. All cues and options were stimuli from the encoding phase within the same block.

737

738

Fig. 2. A subject example for definition of ROI for later analysis. (A) Topography of ITPC values in LeadRvsf. ITPC across trials with leading movies presented at the right visual field were calculated. Strongest ITPC was clustered at the contralateral (left) visual hemifield. **(B)** Topography of ITPC values in LeadLvsf. Similar to (A) but for LeadLvsf with opposite pattern. **(C)** Topographic distribution of ITPC difference between LeadRvsf and LeadLvsf (LeadRvsf – LeadLvsf). One electrode showing strongest positive and negative ITPC difference was identified as ROI for the left and right hemifield, respectively, as highlighted in this example.

Fig. 3. Group ITPC distribution. (A) Topography of grand averaged ITPC values at 37.5 ± 2.5 Hz in LeadRvsf. ITPC across trials with leading movies presented at the right visual field were calculated between 1s and 2s upon stimulus onset from 35 to 40 Hz. Strongest ITPC was clustered at the contralateral (left) visual hemifield. **(B)** Same as (A), but for LeadLvsf with opposite pattern. Highlighted electrodes in (A) and (B) were channels pre-selected for averaging ITPC from 20-50 Hz as shown in (D) and (E), respectively. **(C)** Topographic distribution of grand averaged ITPC difference between LeadRvsf and LeadLvsf (LeadRvsf – LeadLvsf). Cluster-based permutation test indicated statistically significant difference ($p_{corrected} < 0.001$) between LeadRvsf and LeadLvsf. **(D)** and **(E)** Time-frequency representation of ITPC value from 20-50 Hz averaged across electrodes highlighted in (A) and (B), respectively. The specificity of high ITPC to frequencies around 37.5 Hz can be visually seen. **(F)** The specificity was statically confirmed by comparing the difference (between ITPC values contra- and ipsilateral to the leading movies presented field) vs. zeros. A cluster of significant t-values emerged in the frequency range around 37.5 Hz ($p_{corrected} <0.001$).

Fig. 4. Memory performance across conditions. Memory accuracy is shown after subtraction of individual mean performance (N =35). Although one-way ANOVA revealed no significant difference across conditions, there appeared to be a weak trend that memory performance was slightly better at 0°. Each dot represents one single subject memory accuracy in one corresponding phase lag condition. The dashed line represents zero. The thick line represents mean performance, the shaded area shows standard error of the mean, and the boxes show 95% confidence level within condition. Distributions within condition is also available. For memory performance without subtraction of individual mean, see Fig. S1.

739

Fig. 5. Group SSVEPs and instantaneous phase offset between leading movies and trailing movies. (A) SSVEPs band-pass filtered at 37.5 ± 2.5 Hz from -1 to 4 second for leading (blue) and trailing movies (red) in 0°, 90°, 180° and 270° appeared to have different patterns. **(B)** A closer inspection of shaded areas in (A). For demonstration purpose, a randomly selected area was zoomed in at the right bottom. **(C)** Instantaneous phase difference between SSVEPs for leading and trailing movies for shaded areas in (A) was binned in a circular histogram. The mean resultant vector of the instantaneous phase difference in 0°, 90°, 180° and 270° phase lag conditions were represented as dark bars (resultant vector lengths are 0.98, 0.99, 0.99 and 0.95, respectively). A V-test confirmed that the phase differences were uniformly distributed around the entrained phase offset, with $p = 0$ for all conditions.

740

741

Fig. S1. Memory performance across conditions. Except for the dash line, the element representation is the same as Fig. 4. The dash line denotes the mean performance across conditions and participants. Without subtraction of individual mean performance across conditions, the trend became less salient.

742

Fig. S2. ITPC in other time windows. ITPC from 1s before and 3.75s after stimulus onset in the LeadRvsf (A) and LeadLvsf (B) conditions. ITPC was averaged across electrodes highlighted in Fig. 3A and B for LeadRvsf and LeadLvsf, respectively.

743

Fig. S3. Phase offsets at individual level. Each dot represents a mean phase offset for each participant at the corresponding modulation condition. The calculation of mean phase offset was based on SSVEPs (similar to Fig. 5 but at individual level), averaged across the 1s-time interval from 1s to 2s upon stimulus onset.

744

Fig. S4. Phase offsets at group level based on group ROI. Group ROI (one electrode) at the left and right hemifield was identified using similar procedure as at the individual level (see Fig. 2). The calculation of mean phase offset was based on SSVEPs (similar to Fig. 5 but based on group ROI), averaged across the 1s-time interval from 1s to 2s upon stimulus onset. The mean phase differences with 95% confidence interval are 15.75° ± 0.42°, 93.95° ± 0.64°, 180.04° ± 0.22° and 269.67° ± 0.63° for 0, 90, 180 and 270 phase offset conditions, respectively. The V-test confirmed that the phases from 1s -2s upon stimulus onset were uniformly distributed around their entrained phase offset, with $p = 0$ for all conditions.

745

Fig. S5. Verification of precise timing of the stimulus. Time frequency analysis revealed a peak at 37.5 Hz.