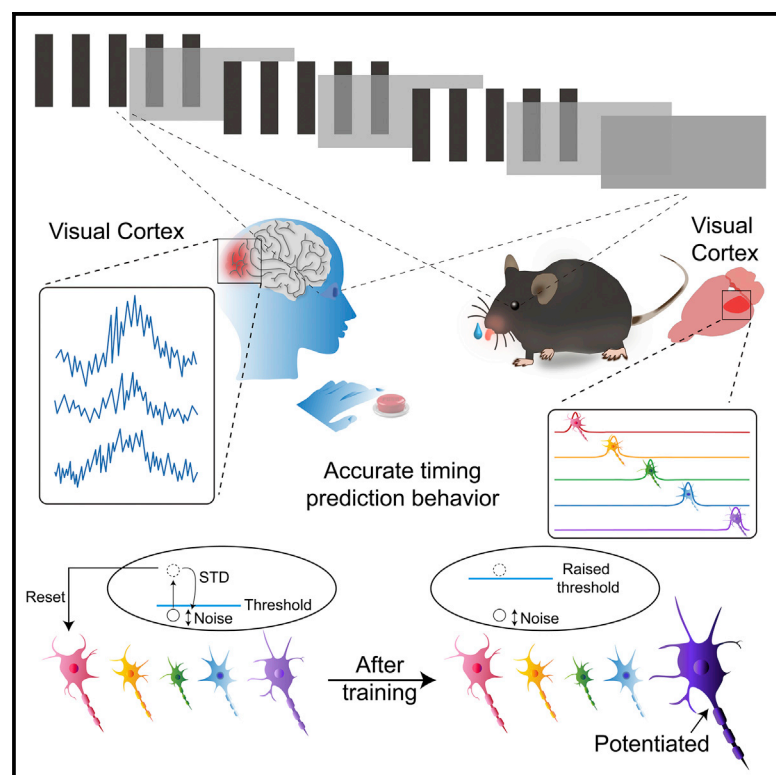


# Visual cortex encodes timing information in humans and mice

## Graphical abstract



## Authors

Qingpeng Yu, Zedong Bi,  
Shize Jiang, ..., Ying Mao, Liang Chen,  
Jiayi Zhang

## Correspondence

maoying@fudan.edu.cn (Y.M.),  
hschenliang@fudan.edu.cn (L.C.),  
jiayizhang@fudan.edu.cn (J.Z.)

## In brief

Qingpeng Yu et al. find that visual cortex acts as a key timing area in processing visually cued interval-timing information in the range of seconds to deca-seconds through intrinsic circuit dynamics in humans and mice.

## Highlights

- Visual cortex (VC) is a key interval-timing brain area in humans and mice
- Localized oscillations relate second-scale timing information to guide behavior
- Firing pattern and excitability of neurons in VC exhibited timed manner in mice
- Computational model revealed a self-correcting learning process for timing

Article

# Visual cortex encodes timing information in humans and mice

Qingpeng Yu,<sup>1,10</sup> Zedong Bi,<sup>2,3,4,5,10</sup> Shize Jiang,<sup>1,10</sup> Biao Yan,<sup>1,10</sup> Heming Chen,<sup>1,10</sup> Yiting Wang,<sup>1,10</sup> Yizhan Miao,<sup>1</sup> Kexin Li,<sup>1</sup> Zixuan Wei,<sup>1</sup> Yuanting Xie,<sup>1</sup> Xinrong Tan,<sup>1</sup> Xiaodi Liu,<sup>1</sup> Hang Fu,<sup>1</sup> Liyuan Cui,<sup>1</sup> Lu Xing,<sup>1</sup> Shijun Weng,<sup>1</sup> Xin Wang,<sup>6</sup> Yuanzhi Yuan,<sup>6</sup> Changsong Zhou,<sup>4,5</sup> Gang Wang,<sup>7</sup> Liang Li,<sup>7</sup> Lan Ma,<sup>1</sup> Ying Mao,<sup>1,\*</sup> Liang Chen,<sup>1,8,\*</sup> and Jiayi Zhang<sup>1,9,11,\*</sup>

<sup>1</sup>State Key Laboratory of Medical Neurobiology, MOE Frontiers Center for Brain Science and Institutes of Brain Science, Department of Neurosurgery, Huashan Hospital, Fudan University, Shanghai 200032, China

<sup>2</sup>Lingang Laboratory, Shanghai 200031, China

<sup>3</sup>Institute for Future, School of Automation, Qingdao University, Qingdao 266071, China

<sup>4</sup>Department of Physics, Centre for Nonlinear Studies and Institute of Computational and Theoretical Studies, Hong Kong Baptist University, Kowloon Tong, Hong Kong

<sup>5</sup>Research Centre, HKBU Institute of Research and Continuing Education, Shenzhen, China

<sup>6</sup>Department of Neurology and Ophthalmology, Zhongshan Hospital, Fudan University, Shanghai 200032, China

<sup>7</sup>Center of Brain Sciences, Beijing Institute of Basic Medical Sciences, Beijing 100850, China

<sup>8</sup>Tianqiao and Chrissy Chen Institute Clinical Translational Research Center, Shanghai 200040, China

<sup>9</sup>Institute for Medical and Engineering Innovation, Eye & ENT Hospital, Fudan University, Shanghai 200031, China

<sup>10</sup>These authors contributed equally

<sup>11</sup>Lead contact

\*Correspondence: [maoying@fudan.edu.cn](mailto:maoying@fudan.edu.cn) (Y.M.), [hschenliang@fudan.edu.cn](mailto:hschenliang@fudan.edu.cn) (L.C.), [jiayizhang@fudan.edu.cn](mailto:jiayizhang@fudan.edu.cn) (J.Z.)

<https://doi.org/10.1016/j.neuron.2022.09.008>

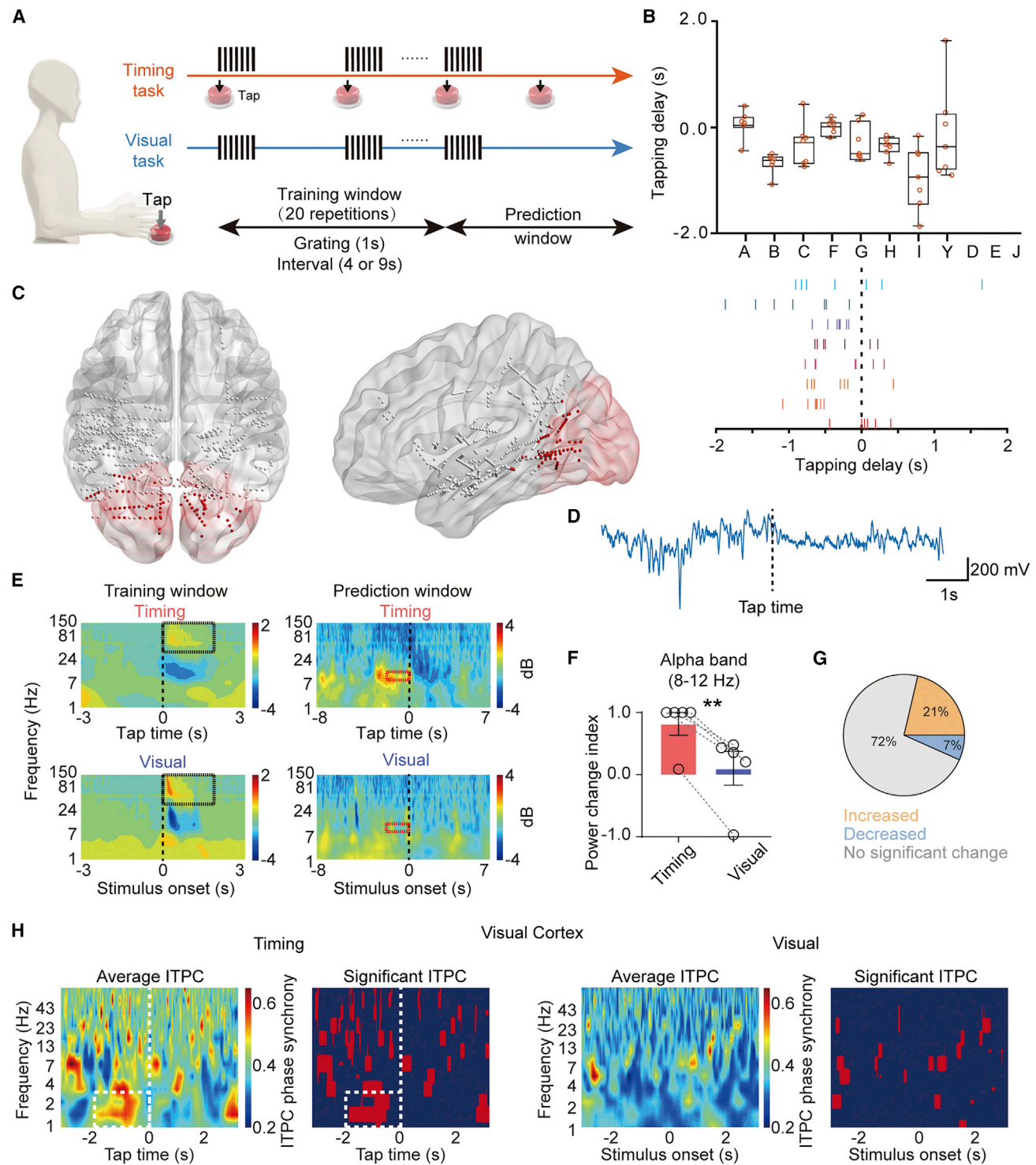
## SUMMARY

Despite the importance of timing in our daily lives, our understanding of how the human brain mediates second-scale time perception is limited. Here, we combined intracranial stereoelectroencephalography (SEEG) recordings in epileptic patients and circuit dissection in mice to show that visual cortex (VC) encodes timing information. We first asked human participants to perform an interval-timing task and found VC to be a key timing brain area. We then conducted optogenetic experiments in mice and showed that VC plays an important role in the interval-timing behavior. We further found that VC neurons fired in a time-keeping sequential manner and exhibited increased excitability in a timed manner. Finally, we used a computational model to illustrate a self-correcting learning process that generates interval-timed activities with scalar-timing property. Our work reveals how localized oscillations in VC occurring in the seconds to deca-seconds range relate timing information from the external world to guide behavior.

## INTRODUCTION

Most human behaviors involve the perception of both space and time. Human brains organize sensory projections, including auditory, visual, and place maps, to perceive space in the external world (Bushara et al., 1999; Ekstrom et al., 2003; Kravitz et al., 2011). The perception of time on the order of 24 h, namely circadian timing, is mainly driven by the suprachiasmatic nucleus of the hypothalamus (Hastings et al., 2018). The perception and prediction of temporal patterns on the order of seconds to deca-seconds are prerequisites for the estimation of predators looming over animals, as well as the performance of sports, dance, and music for humans (Merchant et al., 2013; Nobre and van Ede, 2018). However, the representation of timing in the human brain in the range of seconds to deca-seconds is much less understood.

Much of what is known about timing-related human brain structures comes from studies using functional magnetic resonance imaging (fMRI), electroencephalography (EEG), and magnetoencephalography (MEG). These studies revealed various timing-related brain networks during timing tasks on the order of seconds to deca-seconds. For instance, basal ganglia, thalamus, insula, and cingulate-parietal cortical networks are often associated with interval-timing tasks (Bartolo and Merchant, 2015; Thaut, 2003; Peczenka et al., 2013; Teki et al., 2011). Phase analysis revealed that theta and alpha band oscillations in visual and auditory cortices play an important role in timing-learning tasks (Bueti et al., 2010; Cravo et al., 2013; Henry et al., 2014; Keitel et al., 2017; Kösem et al., 2018; Mathewson et al., 2009; Thut et al., 2006). This association was further confirmed by transcranial stimulation studies with local perturbation of V1 in humans (Helfrich et al., 2014; Salvioni et al., 2013). These studies



**Figure 1. Elevation of alpha band oscillation power and phase synchrony in visual cortex during a synchronization-and-continuation timing prediction task**

(A) Schematics of the synchronization-and-continuation experiment. In the timing task, the participants were instructed to tap the keyboard in both the training and the prediction windows, indicating the behavior of timing prediction. In the visual task, the participants were instructed to passively look at the screen. (B) Upper: tapping latency of each participant relative to the expected onset time of periodic visual stimuli in the prediction window. Lower: raster plots of the same data in the upper panel.

(legend continued on next page)

presented important evidence for the identification of timing-related brain areas. However, fMRI studies rely on indirect measurement of neural activities with limited temporal resolution. Although non-invasive EEG signals have millisecond temporal resolution, the signal quality deteriorates when penetrating through the scalp and the skull.

To address the question of which brain areas play a critical role in timing perception on the order of seconds to deca-seconds for humans, we took advantage of stereoelectroencephalography (SEEG) recordings (Kirkby et al., 2018) that were performed in eleven patients with epilepsy for seizure localization. SEEG allows intracranial measurement of electrical activities at millisecond temporal resolution in awake humans who are performing interval-timing tasks. The recordings were made over 1 week as participants performed interval-timing tasks daily. By conducting power and phase analysis in the 28 recorded brain areas, we proposed visual cortex (VC) to be a critical brain area for the interval-timing behavior.

To further tackle the problem of the cellular and circuit mechanism that mediates such timing behavior, we took advantage of circuit dissection tools in mice. Similar to humans, mice can also recognize and learn temporal sequences with visual stimuli (Gavornik and Bear, 2014; Xu et al., 2012, 2014; Yao et al., 2007), or visually cued reward timing (Chubykin et al., 2013; Gavornik et al., 2009; Liu et al., 2015; Makino and Komiyama, 2015; Nambodiri et al., 2015; Shuler, 2016; Shuler and Bear, 2006; Zold and Shuler, 2015). Similar to the timing prediction paradigm in humans, we conducted a synchronization-and-continuation behavioral experiment in mice. To explore the representation of timing information in the primary VC (V1), we conducted *in vivo* cell-attached recordings and population recordings in V1, revealing the plastic changes and time-keeping properties in V1 neurons. Computational modeling further illustrates the mechanism in V1 circuits that represents timing information. Our study reveals that the primary VC acts as a localized center to represent visually cued timing information through intrinsic circuit dynamics.

## RESULTS

### Increased alpha band power in visual cortex precedes the timing prediction behavior in humans

We recorded 11 presurgical patients diagnosed with epilepsy using SEEG to investigate the brain activity patterns during a

synchronization-and-continuation task that consists of a visually cued interval-prediction (“timing”) task and a control task (“visual”) (Figure 1A; see STAR Methods). The “timing” task consisted of 20 periodic drifting gratings (1-s stimulation with 5-s period, defined as the training window). Patients were instructed to tap the keyboard right before the onset of each drifting grating. After 20 stimulations, the patients tapped the keyboard one more time, anticipating an additional drifting grating that would not be presented (defined as the prediction window). In the “visual” task, the patients were instructed to passively look at 20 periodic drifting gratings without taking any action. Each session consisted of one “timing” task and one “visual” task. Each patient participated in seven sessions over 2–7 days. Among the eleven patients, three did not tap the keyboard in the prediction window in the “timing” task; the others tapped the keyboard around the anticipated time of drifting grating onset (Figure 1B; Video S1) with a standard deviation of 0.3 s. We tracked the eye movement during the timing prediction tasks and did not find an anticipatory eye movement toward the keyboard, excluding the possibility of preparatory activity in the motor system modulating VC (Figure S1).

We recorded intracranial local field potentials (LFPs) from the implanted electrodes in both “timing” and “visual” tasks. Based on MRI scans and anatomical atlas, the recording sites were located in 28 brain regions, namely culmen, putamen, claustrum, caudate, pulvinar, insula, retrosplenial cortex, cingulate gyrus, amygdala, hippocampus, perirhinal cortex, entorhinal cortex, inferior temporal gyrus, superior temporal gyrus, middle temporal gyrus, fusiform gyrus, precentral gyrus, superior frontal gyrus, middle frontal gyrus, inferior frontal gyrus, medial frontal gyrus, postcentral gyrus, inferior parietal lobule, superior parietal lobule, precuneus, and VC (primary VC, secondary VC and associative VC) (Figures 1C and 1D; Table S1). The power of delta (1–3 Hz), theta (3–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), low gamma (30–80 Hz), and high gamma (80–150 Hz) band oscillations from all recorded brain regions were calculated (see STAR Methods). In order to confirm the population activities in VC in response to visual stimuli (Buzsáki and Wang, 2012), we first examined the power of gamma band oscillations during drifting grating stimulation in the training window in both timing and visual tasks. Gamma band activities were found to be increased in VC and other brain regions after the onset of the stimulus (in visual task) or behavior (in timing task), such as the

(C) Electrode locations in a three-dimensional brain across 11 participants in horizontal (top) and sagittal (bottom) views. Red area: visual cortex. Red dots: channels of electrodes located in visual cortex. Gray dots: the other channels of electrodes.

(D) Example raw trace of SEEG signals in visual cortex.

(E) Left column: the heatmap of average power in visual cortex in the training window during the timing (top) or visual (bottom) task across 5 participants. The black rectangle indicates the gamma band frequency within 2 s after the onset of tap or visual stimuli. Right column: the heatmap of average power in visual cortex during the timing (top) or visual (bottom) task in the prediction window across 5 participants. The red dashed rectangle indicates the corresponding range of time and alpha band frequency analyzed in (F) and (G). Time 0 represents the tapping timing prediction behavior in the timing task and the expected onset time of periodic visual stimuli in the visual task.

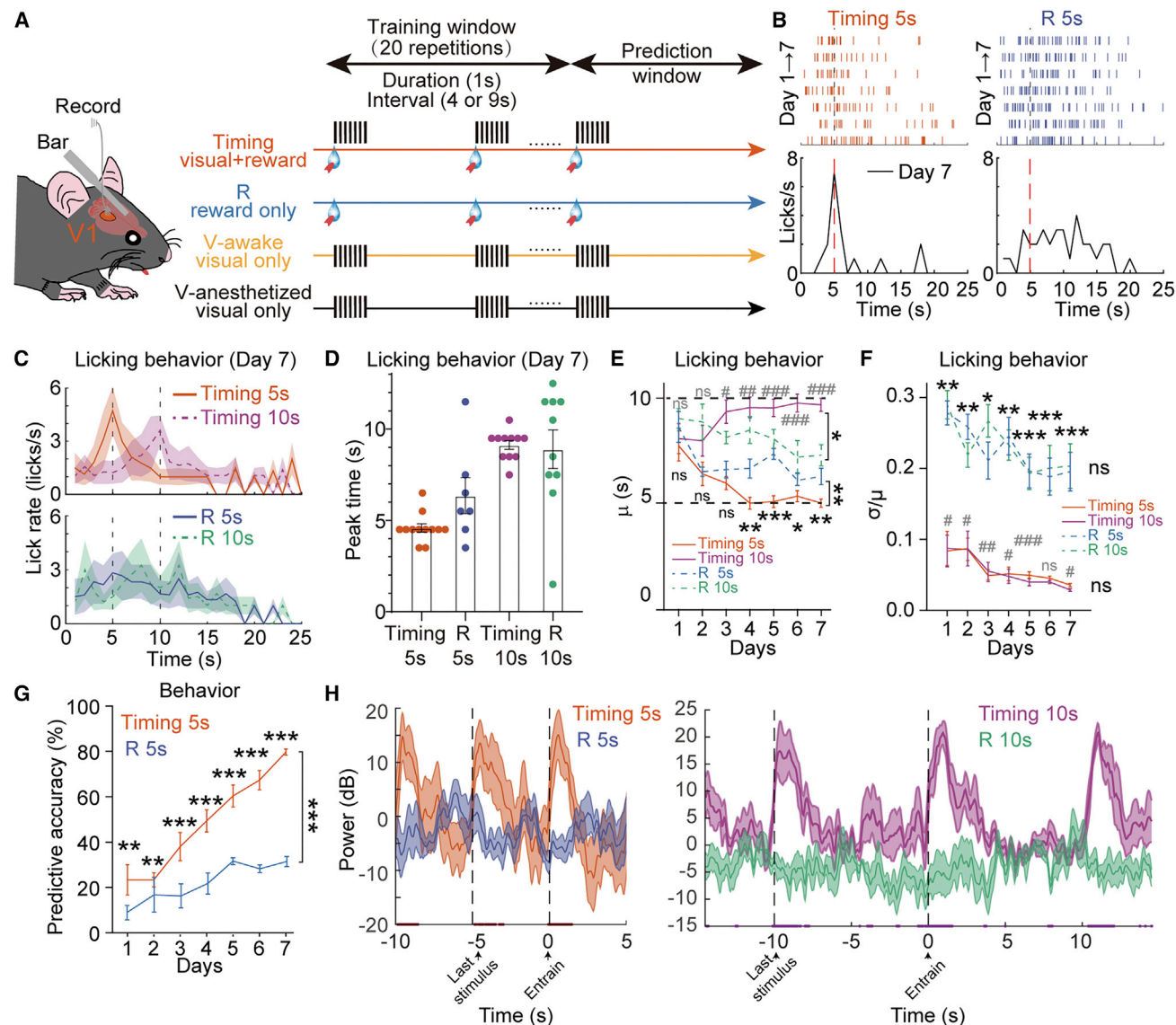
(F) Power change index of alpha oscillations in the prediction window of the timing or visual task across 5 participants.  $p = 0.0018$  and  $t = 7.387$ . Paired *t* test.

(G) Percentage of channels with significant power increase or decrease of alpha oscillations in visual cortex in 5 participants.

(H) Spectrotemporal heatmaps of averaged inter-trial phase clustering (ITPC) of electrodes in visual cortex across 4 participants during the prediction window. Left subplot of the timing or visual task: average ITPC in the prediction window. Right subplot of the timing or visual task: areas with significant difference compared with the null distribution (see STAR Methods) were filled with red, otherwise with blue. The white dashed rectangle indicates the delta band frequency. Error bars represent SEM over taps in (B), and represent SEM over 5 participants in (F).

See also Figures S1–S9, Tables S1 and S2, and Video S1.





**Figure 2. Visual stimuli enhanced timing prediction performance in mice**

(A) Schematics of the synchronization-and-continuation experiment in head-fixed mice. Timing task: water rewards were delivered with periodic drifting gratings in the training window followed by a 60-s gray background as the prediction window. Reward-only (R) task: periodic water rewards were delivered without visual stimuli. V-awake visual only: periodic drifting gratings were presented to awake-behaving mice without water rewards. V-anesthetized visual only: periodic drifting gratings were presented to anesthetized mice.

(B) Upper: example raster plots of licks of one mouse in the prediction window of timing and reward-only tasks with 5-s interval from day 1 to day 7. Lower: corresponding example lick times over 1-s time bins in the prediction window on day 7. Vertical dashed lines represent the expected onset of the 5-s stimulus.

(C) Mean lick times for all mice over 1-s time bins in the prediction window on day 7. Shaded areas represent SEM across mice. Vertical dashed lines represent the expected onset of the 5- or 10-s stimulus.

(D) Peak times in (C) in 5-s timing (orange), 5-s reward-only (blue), 10-s timing (magenta), and 10-s reward-only (green) tasks in the prediction window on day 7 with 1-s time bin. Each data point represented one mouse. If a mouse licked the most at two different time bins, it was counted twice. The data of both peaks were included in the graphs.  $p < 0.001$  and  $t = 5.261$  for timing tasks with 5- and 10-s intervals;  $p = 0.018$  and  $t = 2.489$  for reward-only tasks with 5- and 10-s intervals;  $p = 0.08$  and  $t = 1.799$  for timing tasks and reward-only tasks with 5-s interval;  $p = 0.796$  and  $t = 0.261$  for timing tasks and reward-only tasks with 10-s interval.

(E) Average intervals of licks across mice in the prediction window. Horizontal dashed lines represent the expected onset of 5- or 10-s stimulus. Black asterisks represent statistical significance between timing and reward-only tasks with 5-s intervals.  $p = 0.0064$  and  $t = 3.38$  for 5-s interval.  $p = 0.012$  and  $t = 3.00$  for 10-s interval. Paired  $t$  test.

(F) Scalar property ( $\sigma/\mu$ ) of predictive licking intervals in the prediction window during 7 training days. Black asterisks represent statistical significance between timing and reward-only tasks with 5-s intervals. Gray pound signs represent statistical significance between timing and reward-only tasks with 10-s intervals.  $p = 0.73$  and  $t = 0.36$  for timing task;  $p = 0.57$  and  $t = 0.59$  for R task. Paired  $t$  test. The values of  $\sigma/\mu$  of 5-s V-awake, 5-s V-anesthetized, 10-s V-awake, and 10-s

(legend continued on next page)

precuneus, amygdala, hippocampus, and middle temporal gyrus (Figures 1E, see black rectangle and S2). We did not observe any change in gamma band oscillations in any brain regions (including VC) prior to the timing prediction behavior (tapping) in the prediction window in both timing and visual tasks (Figures 1E and S3; Table S2). These data suggest that gamma band oscillations in the recorded areas might be induced by visual stimuli rather than timing prediction.

Five patients were implanted with electrodes through VC. In Figure 1E, the average power of alpha band oscillations in VC exhibited a significant increase prior to the timing prediction behavior (tapping) in the prediction window, indicated by the large positive power change index (see STAR Methods; Figure 1F). 21% (16 channels) of the 75 recording channels in VC from 5 patients showed a significant alpha band power increase, while 7% of the 75 recording channels exhibited a significant power decrease before the timing prediction (Figure 1G). No obvious activation prior to tapping was observed in the training window (Figure 1E, upper left panel). Both the fraction of trials with elevated alpha oscillations and the fraction of predictive tapping were higher in the prediction window than in the training window (Figure S4). There was no significant increase in LFP when the patient conducted random taps without visual stimuli, indicating that tapping itself does not induce activity elevation in VC (Figure S5).

To explore whether the elevated alpha band oscillations appeared in the first session or gradually emerged as the patient got familiar with the timing prediction task, we further analyzed the power change index of alpha band oscillations in VC in the seven sessions. As a result, we found that the power change index decreased gradually with the number of sessions in the visual task, whereas the index remained similar in the timing task (Figure S6). These results indicate that the elevation of alpha oscillation power in VC was directly associated with the timing behavior. Interestingly, we found an increase in the power of alpha oscillations in the precuneus, a visuospatial-related brain area, and in that of theta oscillations in the hippocampus (Figure S3; Table S2). 10% of all the channels in hippocampus (2 out of 21) showed a significant increase in the theta band, while 5% of all the channels (1 out of 21) significantly decreased in the theta band, showing that the fraction of channels with a significant increase in the hippocampus is smaller than that in VC. Elevation in the beta band power prior to tapping was also observed in precentral gyrus of patient Y (Figure S7). These results collectively indicate that increased alpha band power in VC plays a role in the timing prediction behavior in humans.

### Enhanced low-frequency phase synchrony in visual cortex during timing prediction

The phase of intrinsic oscillations could be correlated with stimuli-induced activities across brain regions (Besle et al., 2011; Calderone et al., 2014; Helfrich et al., 2019; Nobre and van Ede, 2018). To study the phase synchrony between the oscillatory activities and the prediction behavior, we analyzed inter-trial phase clustering (ITPC) in the prediction window (see STAR Methods). Except for a scattered increase of ITPC in VC (Figure 1H), no other brain regions displayed significant enhancement of ITPC (Figure S8). The phase synchrony of delta band oscillations in VC showed significantly higher values prior to predictive taps on average (Figure 1H). While alpha band oscillations exhibited elevated power prior to prediction, Figure 1H showed that delta band oscillations in VC exhibited higher phase synchronization before the onset of the timing prediction behavior.

Previous studies suggest that synchronized low-frequency oscillations between different brain regions indicate communication among them (Fries, 2005). To further confirm whether the correlation of low-frequency oscillations in VC with timing behaviors involves other brain areas, we calculated inter-site phase clustering (ISPC) of the low-frequency band between VC and other synchronously recorded brain areas, namely precuneus, hippocampus, amygdala, entorhinal cortex, and middle temporal gyrus. We found that VC did not show consistent phase synchronization with the other brain areas (Figure S9). In summary, prior to the timing prediction behavior, the power of alpha oscillations and the phase synchrony increased significantly only in VC but not in other recorded brain areas. These results indicated that VC plays a critical role in visually cued timing perception.

### Visual stimuli enhance timing prediction performance in mice

Previous studies and our results suggested that oscillations and the phase synchrony in VC were strongly correlated with the timing prediction behavior in humans (Carrillo-Reid et al., 2016; Salvioni et al., 2013). To study the underlying mechanism, we designed a synchronization-and-prediction behavioral paradigm with visual stimuli (Figure 2A; see STAR Methods) to experiment on mice, similar to that studied with human participants (Figure 1A). Meanwhile, we recorded multisite LFPs in layer 2/3 of V1. The behavioral experiment consisted of two different tasks. In the timing task, 20 periodic drifting gratings (10- or 5-s interval, 1-s duration) were displayed on a monitor with simultaneous water reward, followed by a 60-s gray background that served as a prediction window (timing, Figure 2A). To explore whether mice

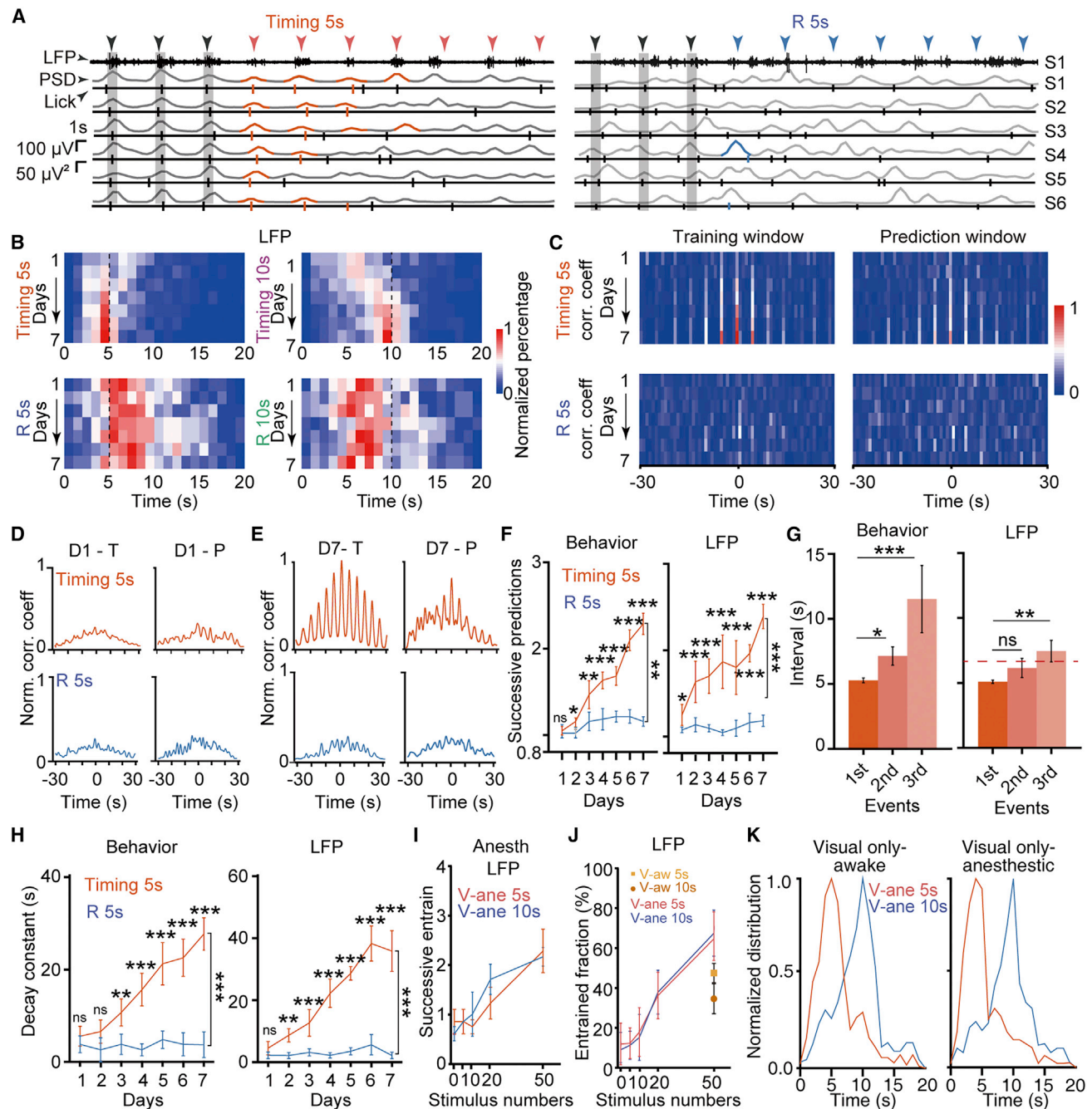
V-anesthetized are 0.194, 0.161, 0.177, and 0.190, respectively. There was no statistical significance between 5 and 10-s in either timing or reward-only conditions.

(G) Predictive accuracy of predictive lickings in the prediction window.  $p = 3.47 \times 10^{-4}$  and  $t = 5.10$ . Paired t test.

(H) Left panel: alpha band power with time in tasks with 5-s intervals. Moments when the power in timing task (orange line) was significantly higher than that in R task (blue line) are indicated by red dots along the x axis. Right panel: alpha band power with time in tasks with 10-s interval. Moments when the power in timing task (purple line) was significantly higher than that in R task (green line) are indicated by purple dots along the x axis. Two-way repeated measured ANOVA, with Bonferroni correction. Shaded areas represent SEM across mice.

$n = 9$  mice for 10-s interval and  $n = 11$  mice for 5-s interval in timing task;  $n = 5$  mice for 10-s interval and  $n = 6$  mice for 5-s interval in R task.  $*p < 0.05$ ,  $**p < 0.01$ ,  $***p < 0.001$ . Error bars represent SEM across mice in each task.

See also Figures S10 and S11 and Video S2.



**Figure 3. Timing prediction behavior correlated with V1 LFP activities in mice**

(A) Examples of predictive lickings and entrained LFP activities on day 7 in timing task with 5-s intervals (S: session). Gray areas: periodic drifting gratings. Black traces: raw LFP data. Gray traces: power spectrum density (PSD). Vertical ticks: licking. Colored triangles, traces and vertical ticks indicate expected stimuli, entrained LFPs and predictive lickings, respectively.

(B) Normalized percentage of intervals of LFP events in the prediction window. Vertical dashed lines represent the expected onset of either 5 or 10-s stimulus.

(C) Normalized cross-correlation coefficients between LFPs and lickings from day 1 to 7 in timing task with 5-s intervals.

(D and E) Examples of normalized cross-correlation coefficients between LFPs and lickings in the training window (T) and the prediction window (P) on day 1 (D1) and day 7 (D7) in timing task with 5-s intervals.

(F) Successive predictive lickings (left) and entrained LFP activities (right) in the prediction window.  $p = 0.020$  and  $t = 2.72$  for behavior;  $p = 0.0014$  and  $t = 2.90$  for LFP. Paired  $t$  test.

(legend continued on next page)



utilize the visual stimuli for timing prediction, we designed the reward-only task in which mice received 20 periodic water reward (10- or 5-s intervals) without visual stimuli (reward-only, R; [Figure 2A](#)).

On training day 7, head-fixed mice in the timing group licked around the stimulus interval in the prediction window ([Figures 2B–2D](#); [Video S2](#)). The lick time in the prediction window in the timing task showed more distinct peaks at the stimulus interval than that in the reward-only task for both 5- and 10-s intervals ([Figures 2B–2D](#), [S10A](#), [S10B](#), and [S11](#)), demonstrating that mice made predictions more accurately at the predictive window in the timing task than in the reward-only task. The average intervals of licking behaviors ( $\mu$ ) were also around 5 and 10 s ([Figure 2E](#)). These behavioral data together demonstrated that visual stimuli significantly enhanced the timing prediction behavior in mice, indicating that visual input plays an important role in the timing prediction behavior.

Psychophysics experiments on the perception and estimation of time showed that the variability of interval timing is proportional to the duration of the interval, known as Weber's law ([Cheng et al., 2008](#); [Meck et al., 2008](#); [Penney et al., 2008](#); [Toda et al., 2017](#)). The standard deviation over mean ( $\sigma/\mu$ ) of the licking behaviors in timing tasks was similar for both 10- and 5-s intervals ([Figure 2F](#)). Interestingly,  $\sigma/\mu$  in timing tasks was significantly smaller than that in reward-only tasks ([Figure 2F](#)), and the observed  $\sigma/\mu$  in both tasks was smaller than the case when the time stamps of licking behaviors were randomly shuffled ([Figure S12](#)). These results suggest that mice were mainly using the visual cue to time and that the reward also contributed to their behaviors. Note that  $\sigma/\mu$  in timing tasks decreased from day 1 to day 7 and were much smaller than those in reward-only tasks ([Figure 2F](#)), suggesting that the mice improved their timing capabilities using the visual cue during the 7-day training.

When the latency of licking to the presumed stimulus in the prediction window was smaller than  $\pm 20\%$  of stimuli intervals, we deemed that the mice exhibited timing prediction behaviors, and “entrained LFPs” were LFP up states that happened during those intervals. We then defined the predictive accuracy as the percentage of timing prediction lickings over total lickings during the prediction window. Consistent with the decreasing  $\sigma/\mu$  value ([Figure 2F](#)), mice exhibited progressively more timed licking behaviors in the prediction windows in the 7-day training period ([Figures 2G](#), [S10C](#), and [S10D](#)).

Visual stimuli also enhanced the timing of LFP during timing prediction. The power of alpha band in mouse V1 increased during prediction in the timing task but not in the reward-only task ([Figure 2H](#)), which shared the similar change found in human data ([Figures 1E and 1F](#)). Similar to that observed in human data, ITPC in mice appeared to be higher prior to the timing prediction. However, no significant difference between timing and reward-only tasks was observed ([Figure S13](#)).

### Interval-timing prediction is correlated with entrained activity in V1 in mice

Intriguingly, a peak in LFP power in V1 occurred around the time of licking behaviors ([Figures 3A and S10E](#)). These results indicated a potential correlation between licking behaviors and LFPs in V1. LFP events became more concentrated around 5 or 10 s in the timing task from day 1 to day 7 ([Figure 3B](#)). The cross-correlation curve between LFPs and licking behaviors in the prediction window exhibited 10- or 5-s periodicity, confirming that the activity in V1 is correlated with the interval-timing predictive behavior ([Figures 3C–3E](#), [S10F](#), and [S10G](#)).

By day 7, mice had on average two successive predictive lickings as well as entrained LFPs in the prediction window in the timing tasks but not in the reward-only tasks ([Figures 3F and S10H](#)). From the first to the third predictive lickings for 5-s interval tasks, the average interval of the behavior increased significantly from around 5 s to over 10 s, indicating that the animal may no longer be anticipating the occurrence of a reward by the third lick ([Figures 3G and S10I](#)). Meanwhile, the LFP peak gradually approached the average spontaneous interval (about 7 s) in [Figures 3G and S10I](#). Moreover, mice retained the predictive licking behavior as well as the entrained LFPs much longer in the timing task than that in the reward-only task ([Figures 3H and S10J](#)). These results further showed that the V1 activity exhibited similar properties to the visually cued interval-timing-predictive behavior.

Both the presence of reward and the anesthetic status of mice were likely to influence the emergence of entrained activities in V1, so we investigated if these two factors contributed to the entrained activities. We presented 50 drifting gratings (with 10- or 5-s interval) to one eye while recording LFPs in contralateral V1 in both awake-behaving (without reward) and anesthetized C57 mice ([Figure 2A](#)). Periodic stimulations induced entrained LFPs in V1 under both conditions, with increasing (successive) entrainment when the number of stimuli became larger ([Figures 3I and](#)

(G) The average interval of the first, second, and third licking behavior or LFP activity in the prediction window for 5-s interval tasks. For behavior events,  $p = 0.016$  and  $t = 2.83$  for 1st and 2nd and  $p = 2.32 \times 10^{-3}$  and  $t = 3.35$  for 1st and 3rd; for LFP events,  $p = 0.0038$  and  $t = 3.69$  for 1st and 3rd. Student  $t$  test. The horizontal dashed line represents the average spontaneous interval of LFP.

(H) Decay time constants of licking behaviors (upper) and entrained LFP activities (lower) in timing (red) and R (blue) task. The decay time constant is the characteristic time of successive predictive lickings or entrained LFP activities in the prediction window. Details can be seen in the [STAR Methods](#). Comparison of decay time constants of licking behaviors:  $p = 2.79 \times 10^{-5}$  and  $t = 6.84$ . Comparison of decay time constant of entrained LFP activities:  $p = 1.40 \times 10^{-5}$  and  $t = 7.38$  for LFP. Paired  $t$  test.

(I) Successive entrained LFP activities in the prediction window depend on increasing stimulus numbers in visual-only-anesthetic trials.

(J) Entrained fractions of LFP in the prediction window of visual-only-awake and visual-only-anesthetic trials.

(K) Normalized interval distributions of LFP activities in visual-only-awake and visual-only-anesthetic tasks.

$n = 9$  mice for 10-s interval and  $n = 11$  mice for 5-s interval in timing task;  $n = 5$  mice for 10-s interval and  $n = 6$  mice for 5-s interval in R task.

$n = 5$  mice for 10-s interval and  $n = 5$  mice for 5-s interval in awake experiments;  $n = 10$  mice for 10-s interval and  $n = 6$  mice for 5-s interval in anesthetized experiments. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . Error bars represent SEM across mice in each task.

See also [Figures S10–S14](#).





3J). For mice both awake and anesthetized, the normalized percentage of entrained LFPs peaked around the corresponding intervals (Figure 3K). These results showed that the emergence of entrained activities in V1 is independent of reward and cognitive status.

Since the experimental tasks involved movements, similar timing experiments were also conducted in the M2 area. Increasing predictive accuracy of licking behaviors was observed from day 1 to day 6, while the predictive accuracy of LFP remained almost the same (Figure S14).

### Entrained activities in V1 contribute to the interval-timing predictive behavior

Nambodiri and colleagues previously demonstrated that the optogenetic perturbation of V1 during inter-trial interval caused a consistent shift in the time of the reward-seeking behavior (licking), which nonetheless abided by temporal scaling (Nambodiri et al., 2015). In our study, the correlation between timing-predictive lickings and LFP activities in V1 suggested that V1 activation in the prediction window also contributes to the timing process. Optogenetic activation of ChR2-expressing PV<sup>+</sup> neurons across V1 in PV-ChR2 mice (PV-Cre mice crossed with loxP-flanked-hChR2-tdTomato mice) (Cang and Feldheim, 2013) largely suppressed the spontaneous activities and visual responses in V1 (Figures S15A and S15B). While PV-ChR2 mice were normally trained to lick water with visual cues in the training window, we suppressed V1 of the PV-ChR2 mice with 30-s continuous laser stimuli during the prediction window (timing-L1 experiments, Figure 4A). The accuracy of predictive lickings increased over 7 days in the training window (Figures S15C and S15D), with the performance of timing prediction similar to that of the control group, indicating the successful conditioning of the licking behavior with visual cues for the PV-ChR2 group. On the other hand, optogenetic manipulation markedly reduced the occurrence of both entrained LFP activities and predictive licking behaviors in the prediction window for

both 5- and 10-s intervals (Figures 4B and S16A). The interval distributions did not show distinct peaks around the stimulus intervals (Figures 4C, S15E, and S15F), consistent with the larger standard deviation over mean ( $\sigma/\mu$ ) in the timing-L1 group than that in the control group (Figure S15G). The fraction of predictive licking behaviors and entrained LFPs in V1 almost reached the baseline (Figures 4D and 4E). More importantly, there was no significant cross-correlation between LFPs in V1 and licking behaviors in the timing-L1 experiments (Figures S15H–S15K). In a word, the suppression of entrained activities in V1 stops the visually cued interval-timing behavior from being accurately performed, indicating that V1 activities in the prediction window play an important role in accurate timing prediction.

Visual signals from the retina were delivered not only to V1 but also to various brain regions including the superior colliculus. To explore whether visual pathway other than retino-thalamic-cortical contributes to the timing behavior, we optogenetically suppressed V1 activities during visual stimuli in the training window in timing tasks (timing-L2 experiments, Figures 4F and S16B). The control group (C57) exhibited the interval-timing predictive behavior and entrained activities in V1 (Figure 4G). Optogenetic suppression of visual responses in V1 of PV-ChR2 mice disrupted the licking accuracy at each stimulus (Figures S17A and S17B). In the prediction window, the distributions of the predictive licking behavior and LFPs were wider and more scattered than control mice (Figures 4G, 4H, and S17C). Predictive accuracy and fraction of entrained activities (Figures 4I and 4J), scalar property (Figures S17D and S17E) as well as the correlation between LFPs and predictive behaviors (Figures S17F–S17I) in V1, were disrupted. Hence, periodic suppression during the training window not only disrupted predictive lickings at the point of each training stimulus but also negatively influenced timing behaviors in the prediction window. Optogenetic suppression of auditory cortex (AU1) of PV-ChR2 mice did not disrupt the predictive licking behavior, compared with control mice, indicating that inactivation of auditory cortex did not cause the mice to be

(G) Examples of predictive lickings and entrained LFP activities on training day 7 of the timing-L2 task. Gray and blue areas represent drifting gratings and blue laser, respectively. Black triangles, colored traces and colored vertical ticks indicate expected time of stimuli, entrained LFPs and predictive lickings, respectively. Gray traces represent PSD.

(H) Left: mean lick times over 1-s time bins in the prediction window of the timing-L2 ctr and the timing-L2 tasks on day 7. Vertical dashed lines represent the expected onset of the 5- or 10-s stimulus. Shaded areas represent SEM. Right: normalized percentage of intervals of LFP events in the prediction window from the timing-L2 ctr and the timing-L2 tasks. Timing-L2 ctr: C57 mice conducted the timing-L2 task.

(I and J) Predictive accuracy of predictive lickings and entrained LFP activities in the prediction window from the timing-L2 task.

(K) Schematics of optogenetic activation training in the R task. Periodic blue laser stimuli were delivered onto contralateral V1 and synchronized with water rewards in CaMKII $\alpha$ -ChR2 mice, referred to as R-L.

(L) Left: mean lick times over 1-s time bins in the prediction window of the timing-L1 ctr and the timing-L1 tasks on day 7. Vertical dashed lines represent the expected onset of the 5- or 10-s stimulus. Shaded areas represent SEM. Right: normalized percentage of intervals of LFP events in the prediction window from the R ctr and the R-L tasks. R-L: CaMKII $\alpha$ -ChR2 mice conducted the R task. R ctr: C57 mice conducted the R task.

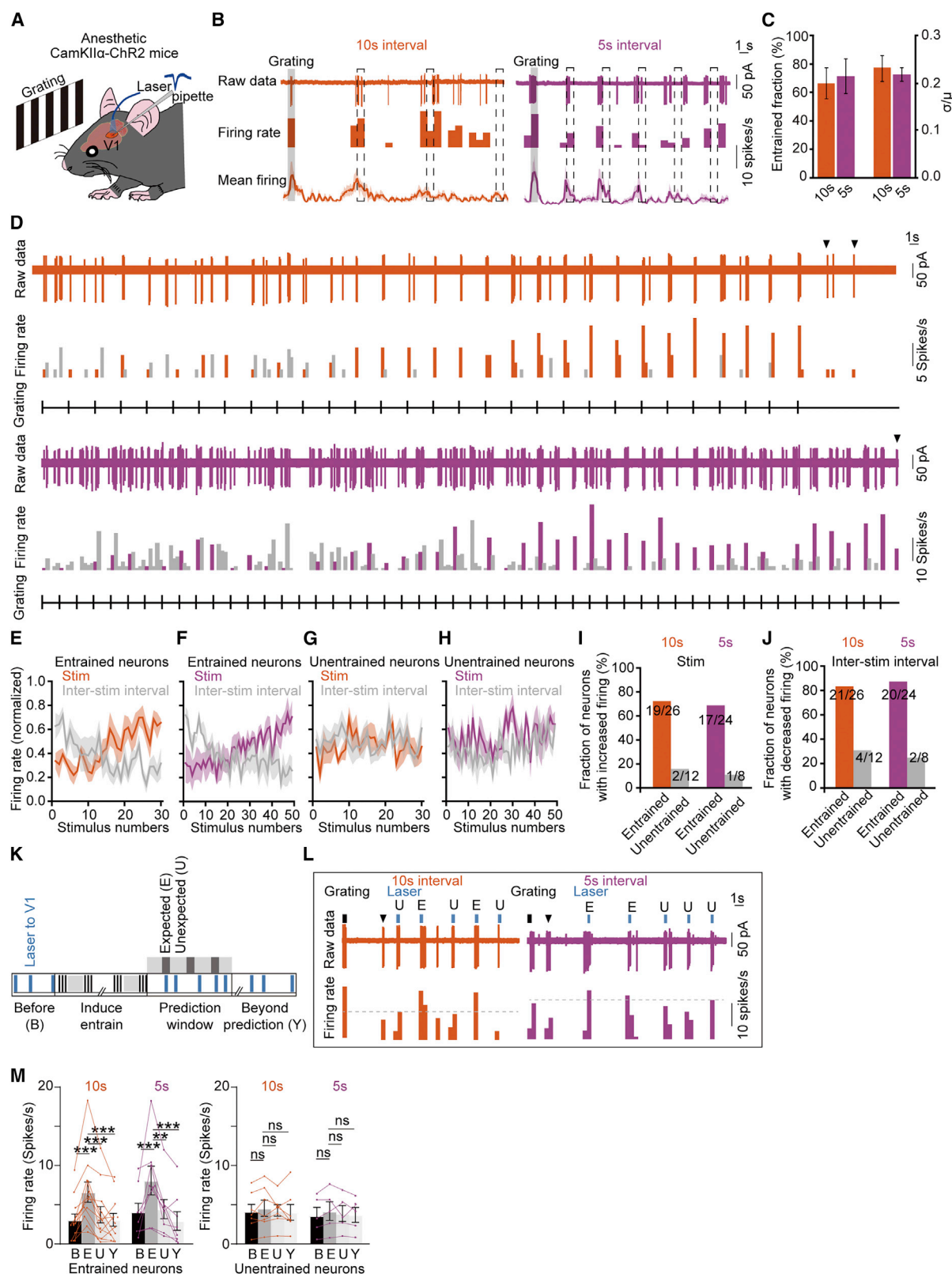
(M) Examples of predictive lickings and entrained LFP activities on training day 7 of the R-L task. Blue areas indicate blue laser. Black triangles, colored traces, gray traces, and colored vertical ticks indicate expected time of stimuli, entrained LFPs, PSD, and predictive lickings, respectively.

(N and O) Predictive accuracy of predictive lickings and entrained LFP activities in the prediction window of the R-L task.

5-s behavior: compared with timing-ctr,  $p = 0.0012$  and  $t = 4.32$  for timing-L1,  $p = 7.44 \times 10^{-4}$  and  $t = 4.63$  for timing-L2; compared with R-L ctr,  $p = 6.32 \times 10^{-4}$  and  $t = 4.72$  for R-L. 10-s behavior: compared with timing-ctr,  $p = 0.0036$  and  $t = 3.68$  for timing-L1,  $p = 6.19 \times 10^{-4}$  and  $t = 4.72$  for timing-L2; compared with R-L ctr,  $p = 4.28 \times 10^{-4}$  and  $t = 4.98$  for R-L. 5-s LFP: compared with timing-ctr,  $p = 1.27 \times 10^{-4}$  and  $t = 5.68$  for timing-L1,  $p = 2.39 \times 10^{-5}$  and  $t = 6.96$  for timing-L2; compared with R-L ctr,  $p = 8.61 \times 10^{-4}$  and  $t = 4.49$  for R-L. 10-s LFP: compared with timing-ctr,  $p = 2.40 \times 10^{-4}$  and  $t = 5.32$  for timing-L1,  $p = 8.65 \times 10^{-4}$  and  $t = 4.50$  for timing-L2; compared with R-L ctr,  $p = 7.66 \times 10^{-4}$  and  $t = 4.61$  for R-L. Paired  $t$  test.

$n = 9, 6, 8, 7, 7$ , and 5 mice for timing-L1 ctr, timing-L2 ctr, timing-L1, timing-L2, R-L, and R-L ctr in the 5-s task;  $n = 7, 5, 7, 8, 9$ , and 7 mice for timing-L1 ctr, timing-L2 ctr, timing-L1, timing-L2, R-L, and R-L ctr in the 10-s task. \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . Error bars represent SEM across mice in each task.

See also Figures S15–S21.



(legend on next page)

less apt to perform in behavioral tasks (Figure S18). These results indicated that entrained activities in V1 relied on periodic visual stimuli and contributed to the predictive behavior. We noticed no remaining entrained activities in timing-L1 experiments but some in timing-L2 experiments. These activities may be due to rhythmic retinal input to other visual nuclei. In summary, suppression of V1 responses to grating stimuli disrupted predictive licking behavior, further suggesting that V1 plays an important role in visually cued timing.

To confirm that periodic activities in V1 led to entrained activities and regulated the predictive behavior, we replaced periodic visual input with periodic optogenetic activation of V1 in R-L experiments. To activate V1 neurons, we injected adeno-associated viruses (AAV-EF1a-DIO-hChR2(H134R)-mCherry) encoding ChR2 into layer 2/3 of V1 of CaMKII $\alpha$ -Cre mice to express ChR2 in excitatory cortical neurons, namely “CaMKII $\alpha$ -ChR2 mice” (Figure S19A; Niell, 2015). Presentation of laser light in V1 induced robust spiking responses in CaMKII $\alpha$ <sup>+</sup> neurons in V1 (Figures S19B and S20). Periodic optogenetic stimuli-induced robust predictive licking behaviors and entrained LFPs in the prediction window in CaMKII $\alpha$ -ChR2 mice (Figures 4K–4O and S19C–S19F). The quantification of timing prediction behaviors, including successive numbers of predictive licking behaviors and entrained LFPs (Figures S19G and S19H), as well as cross-correlations between predictive licking behaviors and LFPs (Figures S19I–S19L), were similar to those in visually cued timing tasks in Figure 3. We also found that as few as 3 periodic stimuli could induce entrained activities 70 s after the last optogenetic stimulation (Figure S21). These results suggested that periodic V1 activation is sufficient to induce timing prediction behavior in mice.

### Spike activities in entrained neurons in V1 exhibit dynamic changes

We next explored the dynamics of V1 activities at the single-cell level using *in vivo* cell-attached recording (Figure 5A; see STAR Methods). When periodically stimulating C57 mice with visual cues (30 drifting gratings for 10-s intervals and 50 drifting gratings for 5-s intervals, 1-s duration), 26 out of 38 V1 neurons

exhibited entrained properties in 10-s interval and 24 out of 32 V1 neurons exhibited entrained properties in 5-s interval (Figures 5B and 5C). The scalar property ( $\sigma/\mu$ ) of entrained firing after two stimulus intervals (5 or 10 s) also showed similar features (Figure 5C). Interestingly, the firing rates of entrained neurons increased progressively during visual stimulation but decreased at inter-stimulus intervals for both 10- and 5-s intervals (Figures 5D–5F). Such dynamic changes did not occur in unentrained neurons (Figures 5G and 5H). Indeed, most entrained neurons in V1 showed robustly enhanced activities during stimulus presentation and depressed activities during inter-stimulus intervals (Figures 5I and 5J).

To further elucidate whether the differential changes in the activity level between stimuli and inter-stimulus intervals extends into the prediction window (if no behaviors were measured), we presented 5 random laser stimuli (0.5-s duration) to V1 in the prediction window in CaMKII $\alpha$ -ChR2 mice to measure the excitability of the neurons (Figure 5K). Firing rates of entrained neurons at the expected time (E) were significantly higher than those at the unexpected time (U) and those before (B) or beyond (Y) the prediction window (Figures 5L and 5M). In contrast, firing rates of unentrained neurons remained unchanged in B, U, E, and Y windows. These results suggested that V1 neurons became more excitable when stimuli were expected to arrive at the later stage of the stimulus window and in the prediction window.

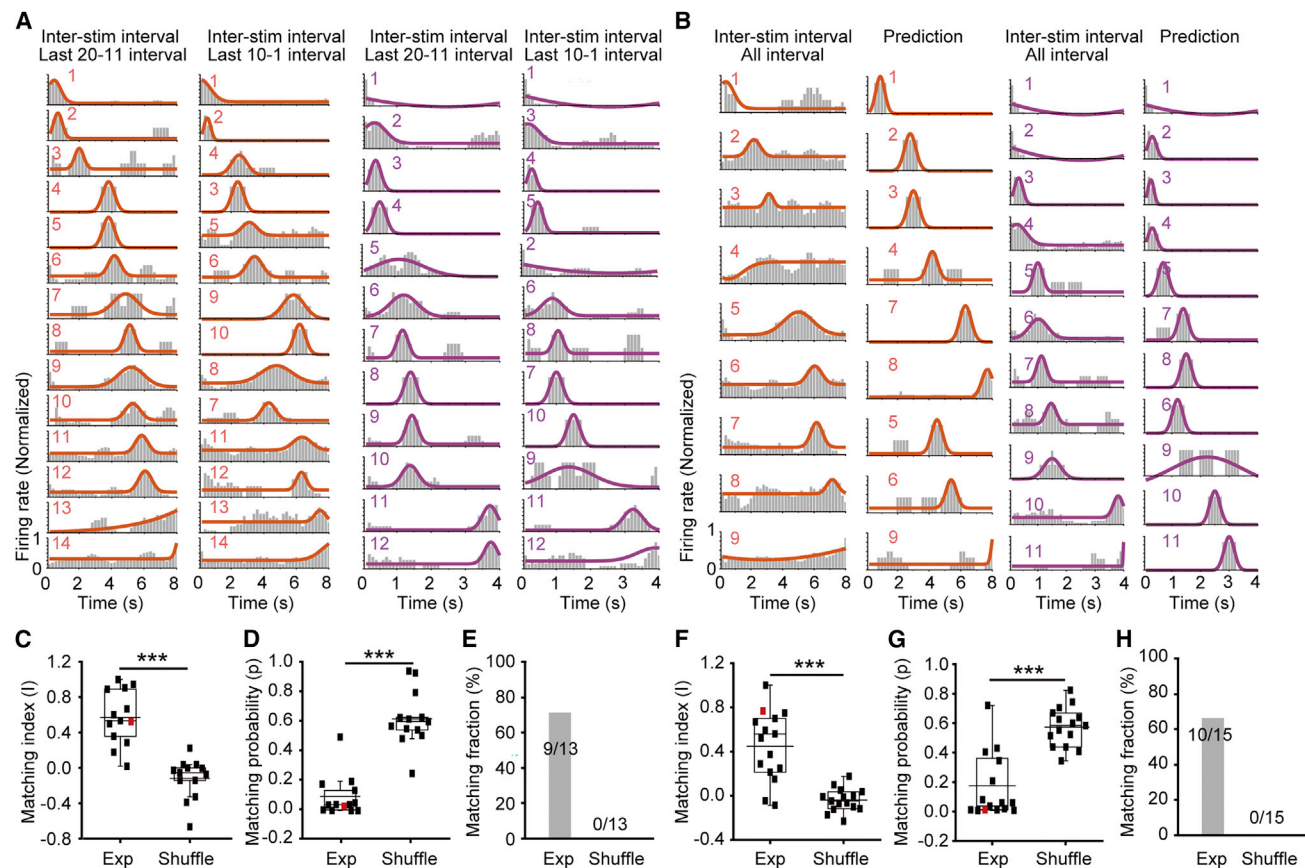
### Population firing in V1 replayed in the prediction window

Entrained neurons in V1 exhibited their capability to retain interval-timing information for a short period of time through dynamic changes in firing patterns and excitability. We next explored the possible source of timing information in V1 (Gu and Gong, 2016; Townsend et al., 2017). Eichenbaum suggested that timing information may be coded through neuronal sequential firing (Eichenbaum, 2014). This finding inspired us to investigate the replay of firing sequence in a neuronal population during the inter-stimulus intervals using a multi-electrode array while C57 mice conducted the timing task. The firing sequence was defined by the time of peak firing rate during inter-stimulus intervals or within the

### Figure 5. Excitability dynamics of entrained neurons in V1

(A) Schematics of the *in vivo* patch-clamp recording setup. Periodic drifting gratings were presented to one eye of anesthetic CaMKII $\alpha$ -ChR2 mice, and CaMKII $\alpha$ <sup>+</sup> neural activities of contralateral V1 were recorded.  
(B) Entrained spiking activities after 10- or 5-s periodic stimuli in example neurons. Gray areas: drifting grating stimuli. Dashed lines: expected stimuli.  
(C) Fraction of entrained neurons and scalar property ( $\sigma/\mu$ ) of spiking activities after 10- or 5-s periodic stimuli.  
(D) Examples of excitatory responses during stimulation. Purple and orange histograms represent the firing rate of visual-evoked responses with 5- and 10-s intervals, respectively; gray histograms represent the firing rate of inter-stimulus intervals; black triangles indicate the entrained spiking activities.  
(E and F) Average firing rate of entrained neurons during inter-stimulus intervals (gray curve) and 10- or 5-s periodic stimuli. Orange curves represent the 10 s interval case and purple curves represent the 5-s interval case.  
(G and H) Average firing rate of unentrained neurons during inter-stimulus intervals and 10- or 5-s periodic stimuli.  
(I and J) Fraction of entrained or unentrained neurons with increased firing rate during stimuli or inter-stimulus intervals.  
(K) Schematics of excitatory responses recording in anesthetic CaMKII $\alpha$ -ChR2 mice. Laser stimuli were presented before the periodic drifting grating (B), during the expected (E) and unexpected time (U) in the prediction window, and beyond prediction (Y).  
(L) Examples of excitatory responses during the expected and unexpected time. Black squares: drifting grating stimuli. Black triangles: entrained spiking activities. Blue squares: laser stimuli. Gray dashed lines: the highest firing rate during the unexpected time (U).  
(M) Quantification of excitatory responses during laser stimuli. Dots with connected lines indicate data from individual neurons. For 10-s intervals (left),  $p = 3.61 \times 10^{-6}$  for E versus B;  $p = 1.82 \times 10^{-4}$  for E versus U;  $p = 7.11 \times 10^{-3}$  for E versus Y. For 5-s intervals (right),  $p = 4.43 \times 10^{-6}$  for E versus B;  $p = 6.07 \times 10^{-3}$  for E versus U;  $p = 3.89 \times 10^{-7}$  for E versus Y. Student's t test.  
 $n = 38$  neurons from 24 mice for 10-s intervals and  $n = 32$  neurons from 18 mice for 5-s intervals. \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ . Error bars represent SEM across neurons in each task.





**Figure 6. Population firing sequences in V1 in the training and prediction windows**

(A) Example of sequential firing during the inter-stimulus interval. Gray histograms represent PSTHs of each neuron; orange and purple lines indicate peak function curves with Gaussian fitting for 5- and 10-s intervals, respectively. The matching index ( $I$ ) is 0.97 and the matching probability ( $p_{match}$ ) is  $4.3 \times 10^{-5}$  for the 10-s interval. The index is 0.84 and the probability is  $7.5 \times 10^{-3}$  for the 5-s interval.

(B) Example of sequential firing of all inter-stimulus intervals and the prediction window. Note that  $-8$  to  $-2$  s in prediction window is defined as “unexpected” window for 10-s intervals and  $-4$  to  $-1$  s in prediction window is defined as “unexpected” window for 5-s intervals. The matching indices are 0.92 and 0.94, and the probabilities are  $2.6 \times 10^{-4}$  and  $1.9 \times 10^{-4}$  for 10 and 5 s, respectively.

(C and D) Box plotting of matching index ( $I$ ) and probability ( $p_{match}$ ) of experimental and shuffled data from the comparison between last 20-11 and last 10-1 stimuli.  $p = 7.56 \times 10^{-5}$  and  $t = 6.03$  for matching index and  $p = 1.21 \times 10^{-5}$  and  $t = 7.42$  for probability between experimental and shuffled data. Paired  $t$  test. Each dot represents the data from a single mouse. Red dots indicate the data corresponding to the firing sequences illustrated in Figures 6A and 6B.

(E) Matching fraction of the firing sequence between last 20-11 stimuli and last 10-1 stimuli.

(F and G) Similar to (C) and (D). The firing sequence between all inter-stimulus intervals and the prediction window.  $p = 6.96 \times 10^{-4}$  and  $t = 4.66$  for matching index and  $p = 6.86 \times 10^{-4}$  and  $t = 4.67$  for probability between experimental and shuffled data. Paired  $t$  test. Red dots indicate the data corresponding to the firing sequences illustrated in Figures 6A and 6B.

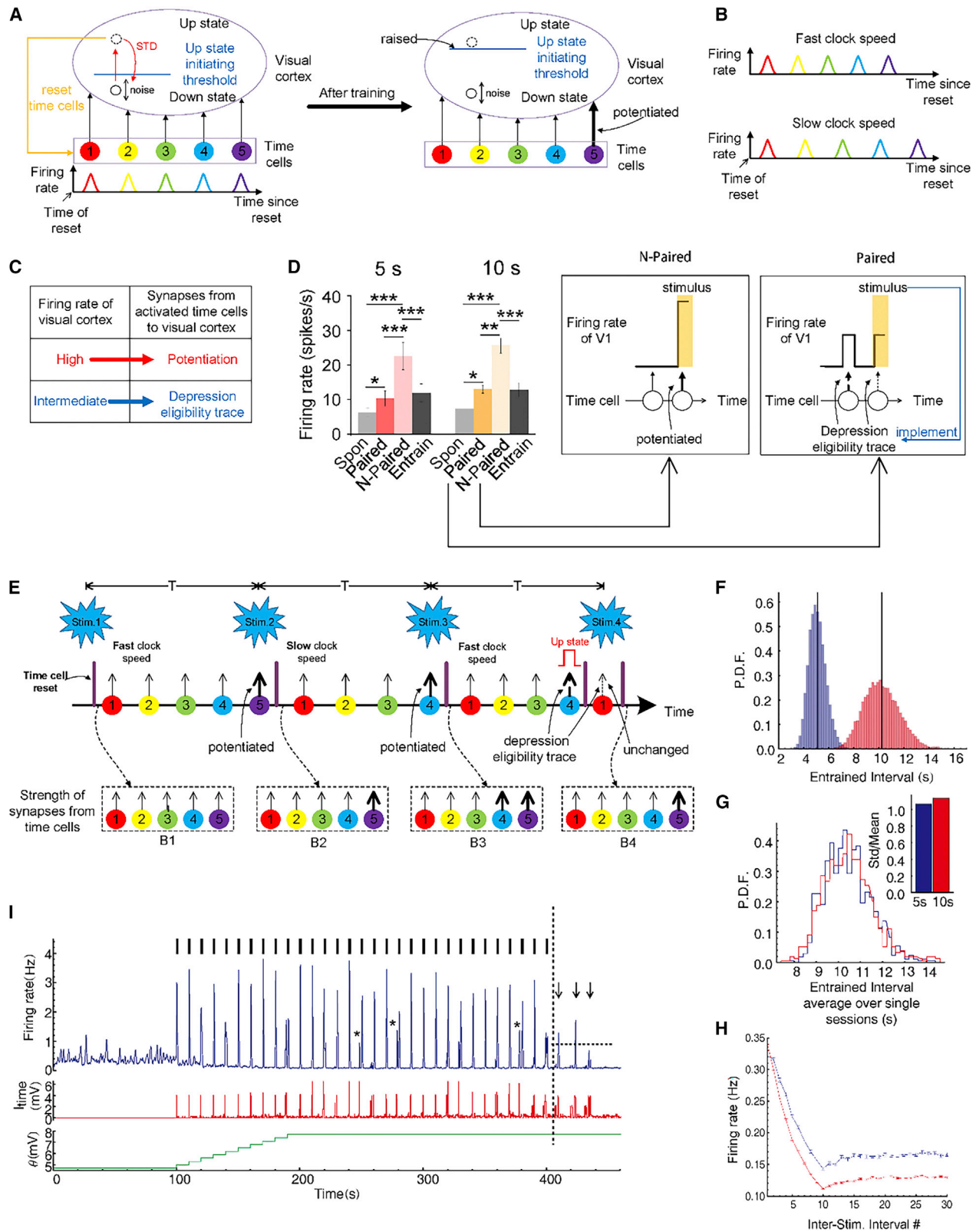
(H) Matching fraction of firing sequence for all inter-stimulus interval. \*\*\* $p < 0.001$ .

$n = 13$  mice for 5-s intervals and  $n = 15$  mice for 10-s intervals. Error bars represent SEM across mice in each task.

See also Figure S22.

prediction window. We used a combined method to quantify the sequence replay using matching index  $I$  and matching probability  $p$  (see STAR Methods) (Ji and Wilson, 2007; Lee and Wilson, 2004). Specifically, for any given two sequences, the matching index  $I_{match}$  measures the similarity between them, and the matching probability  $p_{match}$  calculates the fraction of matching indices equal to or larger than  $I_{match}$  among all possible shuffled sequences. A threshold of  $p_{match} < 0.05$  was used as a criterion to evaluate matching significance between the two sequences (see STAR Methods). We compared the sequence in the last 20-11 stimuli with that in the last 10-1 stimuli

and the sequence calculated from all inter-stimulus intervals with that in the prediction window. Most sequential neurons were active in the “unexpected” window, indicating that these neurons were “untrained” neurons (Figures 6A and 6B). The sequential neurons fire at similar rates during the inter-stimulus intervals in most trials (Figure S22), confirming that sequential neurons were “untrained” (Figures 5G and 5H). As expected, the sequential neurons fired around the same time in most trials during inter-stimulus intervals (Figure S22). In both cases, compared with shuffled data, experimental data had significantly larger  $I_{match}$  values and smaller  $p_{match}$  values, with a higher



(legend on next page)

matching fraction (the fraction of sequence pairs that met the criterion of  $p\_match < 0.05$ , [Figures 6C–6H](#)). These results demonstrated that the replay of firing sequences within inter-stimulus intervals extends into the prediction window.

### Computational model for entrained activities in V1

Experiments in humans and mice showed that VC plays a crucial role in visually cued timing experiments ([Figures 1 and 4](#)). More interestingly, sequential firing on the neuronal level and elevated excitability toward the end of the training window in visual cortical neurons are timing-related features ([Figures 5 and 6](#)). Inspired by these experimental observations, we hypothesized that the sequential firing pattern retains timing features during the periodic stimulation training, which leads to elevated excitability through plastic changes in the cortical network. Guided by this hypothesis, we proposed a computational model with a two-layer network (see [STAR Methods](#)) to elucidate potential mechanisms of the entrained activities ([Figure 7A](#)). The first layer models VC, whose recurrent dynamics has two attractors representing up and down states ([Holcman and Tsodyks, 2006](#)). The second layer models time cells that start to fire in a stereotyped sequence (the sequence is Cells  $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5$  in [Figure 7A](#)) after being stimulated by a short strong input ([Buonomano and Laje, 2010](#); [Cone and Shouval, 2021](#); [Mi et al., 2013](#); [Veliz-Cuba et al., 2015](#); [Voloh and Womelsdorf, 2016](#)), a process we call “time-cell resetting.” In the down states, neuronal noises sporadically initiate the attractor network to the up state ([Figure 7A](#), left subplot). In the up states, the high neuronal activity triggers short-term depression (STD) of the recurrent connections in the attractor network, which soon drags the network back to the down state ([Neske, 2015](#)). Training raises the up-state-initiating threshold of the attractor network and potentiates the synapses from some time cells (say, Cell 5 in the right subplot of [Figure 7A](#)) to the attractor network. As a result of the raised threshold in training, the noises can hardly initiate up state before

the firing of Cell 5, in the prediction window without stimulus, but when Cell 5 fires, the strong synaptic input drives the attractor network to up state. This up state resets the time cells, which starts to fire in the sequence  $1 \rightarrow 2 \rightarrow 3 \rightarrow \dots$  until Cell 5 fires again, causing another up state that resets the sequence again. This is the mechanism of the periodic entrained-up states.

Due to the fluctuation of the dopamine level ([Ruskin et al., 1999](#); [Soares et al., 2016](#)), time cells have clock-speed variability ([Matell and Meck, 2004](#)), which means that the sequential updating speed is variable after each resetting ([Figure 7B](#)). To explain the entrainment phenomenon under clock-speed variability, we propose a plasticity rule similar to Bienenstock-Cooper-Munro (BCM) rule for the synapses from the time cells to the attractor network ([Bienenstock et al., 1982](#); [Shouval et al., 2002](#); [Figure 7C](#)): if the attractor network has high firing rate, the synapses from the co-firing time cells are potentiated; if the attractor network has intermediate firing rate, then the synapses from the co-firing time cells are tagged by eligibility traces ([Gerstner et al., 2018](#)) and become subject to depression by an external stimulus. The firing rate of the mice VC neurons during a visual stimulus presentation was high (or relatively low) if there was no (or a) spontaneous up state happening immediately before the stimulus, while the firing rate during an entrained-up state was also relatively low ([Figure 7D](#), left panel). Inspired by these experimental observations, in our model, a stimulus potentiates the synapses from the time cells (or tagged depressive eligibility traces onto the synapses) if there is no (or a) time-cell-initiated up state immediately before the stimulus. Time-cell-initiated up states also tag depressive eligibility traces onto the synapses ([Figure 7D](#), middle and right panels). See [Figure 7E](#) for an illustration of the training process.

With clock-speed variability and this plasticity rule, the intervals between the entrained-up states are equal to the stimulus intervals on average and obey the scalar property ([Figures 7F and 7G](#)) in our simulation (see [Figure 7H](#) for an example). Due

### Figure 7. Computational model to explain entrained activities

- (A) The model contains an attractor network modeling visual cortex and a group of time cells. Left: the attractor network can be kicked to up state by noises and return to down state due to short-term depression (STD). The brief up state resets the time cells that start to fire in a stereotypical sequence (color bumps). Right: training raises the up-state-initiating threshold of the attractor network and potentiates the synapses from Cell 5 to the attractor network.
- (B) Clock-speed variability. The speed of sequential updating of the time cells may be fast (upper subplot) or slow (lower subplot) after resetting.
- (C) The proposed BCM-like plasticity rule of the synapses from the time cells to the attractor network.
- (D) Left panel: experimental data of mice. Neuronal firing rates in V1 immediately before the stimulus during spontaneous (Spon) up states, entrained (Entrain) up states, stimulus-evoked up states with (paired) or without (N-paired) pairing spontaneous up states, when  $T = 5$  (left bars) and 10 s (right bars). Data are expressed as mean  $\pm$  SEM over neurons. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ . Middle and right panels: schematics of the learning rule.
- (E) Schematics of the training process. Stimuli (blue sparks at the top) with period  $T$  train the network. Color circles: the activated time cells at specific times. Black arrows: weak (thin arrow) or strong (thick arrow) synapses. Dashed curved arrows: configurations of the synaptic strengths at the time immediately after each stimulus. Purple thick vertical lines: time-cell resetting. See further illustration of the training process in [Figure S24B](#).
- (F) The probability distribution functions (PDFs) of the intervals between entrained-up states when stimuli periods are 5 (blue) and 10 s (red), with vertical black lines indicating their means.
- (G) There may be several entrained intervals in each simulation session (see I for illustration). Here we first averaged over the entrained intervals in single sessions and plotted the distributions of the average entrained intervals when stimuli periods are 5 (blue) and 10 s (red). Inset: the ratios of the standard deviation over mean of these two distributions.
- (H) The change of firing rate of the attractor network within inter-stimulus intervals during training when stimuli periods are 5 (blue) and 10 s (red). Error bars represent SEM over different simulation sessions.
- (I) Firing rate, synaptic input from time cells  $I_{time}$  and neuronal firing threshold  $\theta$  of the attractor network in an example session. Black bars: stimuli. Black arrows: entrained activities. Periodic stimuli are removed after the dashed vertical line. The last entrained up state cannot reset the time cells, which in our model requires the firing rate higher than the dashed horizontal line. Asterisks indicate some spontaneous up states immediately before stimuli, which were also observed in mice experiment ([Figure S23](#)). See also [Figures S23–S25](#).

to a homeostatic mechanism (see [STAR Methods](#)) raising up the initialization threshold of up states ([Figure 7A](#), right panel), the firing rate of the attractor network in between stimuli decreases with the number of stimuli ([Figures 7H and 7I](#)), which is consistent with the experimental observation of the firing decrease of entrained neurons ([Figures 5G and 5H](#)). Details of the modeling can be found in [supplemental information](#) and [Figures S24 and S25](#).

## DISCUSSION

One challenge in understanding time perception is whether there is a core timing circuit, or whether timing properties exist in multiple neural circuits ([Merchant et al., 2013](#); [Paton and Buonomano, 2018](#)). To address this question, we first investigated how distributed or localized are the brain areas involved in timing perception on the scale of seconds. Previous studies identified the cortico-thalamic-basal ganglia circuit as one of the core timing structures ([Merchant et al., 2013](#)). Specifically, theta (near alpha frequency) oscillations have been shown to be correlated with visually cued interval timing ([Zold and Shuler, 2015](#)) and visually cued action timing ([Levy et al., 2017](#)) in the primary VC. With simultaneous SEEG recordings from 28 brain regions in humans, our study showed a significant increase in the correlation of alpha oscillation power in VC with time perception. There is also enhanced phase synchrony in alpha oscillations within the prediction window. These observations are consistent with previous studies using fMRI and EEG recordings that indicated oscillations in VC may be involved in temporal expectations ([Buetti et al., 2010](#); [Cravo et al., 2013](#); [Keitel et al., 2017](#); [Levy et al., 2017](#); [Mathewson et al., 2009](#); [Thut et al., 2006](#); [Zold and Shuler, 2015](#)). The absence of elevated oscillation power or phase synchronization in basal ganglia, thalamus, and other cortical areas suggests that time information is represented in a localized manner during the prediction of timing with the absence of sensory cues. We also found that VC did not exhibit significant ISPC with other brain regions, suggesting that low-frequency oscillations in VC with timing prediction features were not likely dependent on other brain areas, in line with other studies indicating that VC itself could be a site of time generation ([Chubykin et al., 2013](#); [Monk et al., 2020](#); [Namboodiri et al., 2015](#)). One limitation in SEEG recording, however, is the incomplete coverage of all brain areas, so we do not rule out the possibility that other unrecorded brain areas are important for interval timing.

Due to the low incidence of epilepsy in the occipital lobe and the challenge of direct manipulation of neural activities in well-defined brain areas in humans, it is difficult to further identify the causal role of VC in time perception in humans. By conducting optogenetic manipulations in the same time prediction behavioral paradigm in mice, the causal role of VC in visually cued timing perception was further confirmed, consistent with a previous study ([Liu et al., 2015](#)). Moreover, we found two key properties in VC that contribute to time perception: sequential replay and the existence of “entrained neurons,” respectively, represented by the time-cell layer and the attractor-network layer in our model. These two layers do not correspond to specific layers in V1 anatomically but are instead convenient for implementation. It is likely that the same cortical circuit has both

highly activated up state and a stereotypical temporal sequence during its down state. Our experiment found that the firing rates of entrained neurons in between stimuli decreased with the number of stimuli ([Figures 6E and 6F](#)), which was consistent with the attractor network in the model ([Figure 7H](#)). The sequential neurons, which were unentrained neurons, have similar firing rates in different inter-stimulus intervals ([Figure S22](#)): this is an exact property of the time cells in the model.

Formal models of learning cued interval timing in V1 were supported by experimental evidence in mice ([Cone and Shouval, 2021](#); [Huertas et al., 2015](#); [Veliz-Cuba et al., 2015](#)). Sequential replay of awake experience during sleep has been reported previously in V1 ([Ji and Wilson, 2007](#)). Computationally, sequential replay can be realized through a number of mechanisms, such as tapped delay lines, reservoirs of response time constants or oscillation frequencies, and recurrent networks containing high dimensional line attractors ([Buonomano and Laje, 2010](#); [Mauk and Buonomano, 2004](#)). Exploration of the source of sequential replay may help answer the long-standing question of how the brain codes time. Future studies on entrained neurons using single-cell resolution functional imaging may give more profound insight into how the animal predicts timing on the order of seconds ([Carandini et al., 2015](#); [Madisen et al., 2015](#)). For example, it is plausible to study whether the entrained neurons respond to similar visual stimuli and have overlapping receptive fields, or whether activation of these entrained neurons is necessary or sufficient for the time perception behavior of the animal.

Previous studies showed that neuronal activities in the auditory cortex in a synchronization-and-continuation task are required for anticipatory auditory-motor response in humans ([Pecenko et al., 2013](#)) and mice ([Li et al., 2017](#)), indicating that auditory- and visual-based time perception may share the same mechanism. Our results, together with these previous findings, suggest that second-scale interval timing, at least with respect to the anticipation of predictable sensory stimuli, may substantially manifest in the sensory processing regions that respond to the relevant stimuli. The “clock” for such second-scale timing may not have a fixed meter, but instead can be flexibly tuned by recent experience. Repeated occurrence of sensory stimuli with given intervals in the recent past may suggest the likelihood that similar stimuli will occur at the particular interval in the near future. Accurate anticipation of such stimuli may prepare individuals to avoid possible threats or obtain potential rewards.

Clock-speed variability, which is the variability of the updating speed of the time-keeping neuronal firing sequence, has been proposed to be the source of the variability of behavioral temporal estimation ([Matell and Meck, 2004](#)). Our computational model demonstrates a mechanism whereby the scalar-timing variability of up states can be learned under clock-speed variability, which establishes a generic theory of behavioral scalar-timing variability together with our experimental finding of the correlation between cortical up state and licking ([Figure 3](#)).

Previous studies found that V1 can learn the single interval between a stimulus and an expected reward by persistently firing at a high rate during that interval ([Shuler and Bear, 2006](#)). Our study implemented a beat-timing paradigm, which may have different neural substrates and mechanisms with the single-interval case



(Teki et al., 2011). However, our work and a previous work (He et al., 2015) suggest that V1 uses a similar learning principle for both beat and single-interval timing: a self-correcting process where synapses are potentiated and depressed respectively to correct delayed and premature timing expectation, except that the synapses from time cells to the attractor network are plastic in our model, whereas the recurrent synapses in the network are plastic in the previous work (He et al., 2015). Therefore, self-correcting process is a universal computational principle to acquire timing ability.

In summary, we showed that time information on the second scale is mostly represented in VC through local circuit dynamics. Our study also opens up the possibility of dissecting brain networks by combining human SEEG recordings with circuit-level studies in mice.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Participants
  - Animals
- METHOD DETAILS
  - Experimental design
  - Eye movements
  - SEEG recording and preprocessing
  - Electrode location
  - Frequency decomposition and power
  - Inter-trial phase clustering (ITPC)
  - Inter-site phase clustering (ISPC)
  - Virtual reality (VR) system
  - Visual stimuli presentation
  - Stereotaxic surgery and LED implantation
  - Prediction tasks in VR system
  - *In vivo* extracellular e-phs recording
  - Potentiation analysis
  - *In vivo* cell-attached patch-clamp recording
  - Immunohistochemistry
  - Automated detection of licking behavior
  - Detection of upstate from PSD of LFPs
  - Pairwise correlation analysis
  - Matching analysis of firing sequence
  - Computational model for timing activity in V1
  - Our learning rule vs classic Hebbian rule
  - Advantages of high plasticity rate
- QUANTIFICATION AND STATISTICAL ANALYSIS

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.neuron.2022.09.008>.

## ACKNOWLEDGMENTS

We thank Dr. M. Brainard for his comments on the manuscript. Funding: this work was supported by the Ministry of Science and Technology, People's Republic of China (2022ZD0208604 and 2022ZD0208605 to J.Z.; 2022ZD0210000 to B.Y.), National Natural Science Foundation of China (31771195, 81790640, and 820712002 to J.Z.; 32100803 to B.Y.; 32000694 to Z.B.; 11975194 to C.Z.), Shanghai Municipal Science and Technology Major Project (2018SHZDZX01 to J.Z.) and ZJLab, the Key Research and Development Program of Ningxia, China (2022BEG02046 to J.Z.), Sanming Project of Medicine in Shenzhen, China (SZSM202011015) to J.Z., Key Scientific Technological Innovation Research project by Ministry of Education to J.Z.; Talents Program of Shanghai Municipal Health Commission (to L. Chen).

## AUTHOR CONTRIBUTIONS

Conceptualization, J.Z., L. Chen, and Y. Mao.; methodology, J.Z., L. Chen, Y. Mao, Z.B., and C.Z.; investigation, Q.Y., S.J., Z.W., X.L., X.T., Y.X., L.X., H.F., L. Cui, and S.W.; formal analysis and software, H.C., Y.W., Y. Miao, B.Y., Z.B., C.Z., G.W., and L.L.; writing – original draft, J.Z., B.Y., and Z.B.; writing – review & editing, J.Z., Y.W., K.L.; visualization, Q.Y., H.C., B.Y., and Y.W.; resources, X.W., Y.Y., and L.M.; funding acquisition, J.Z., L. Chen, B.Y., and C.Z.; supervision, J.Z.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 19, 2021

Revised: March 15, 2022

Accepted: September 7, 2022

Published: October 3, 2022

## REFERENCES

- Bartolo, R., and Merchant, H. (2015). beta oscillations are linked to the initiation of sensory-cued movement sequences and the internal guidance of regular tapping in the monkey. *J. Neurosci.* 35, 4635–4640.
- Besle, J., Schevon, C.A., Mehta, A.D., Lakatos, P., Goodman, R.R., McKhann, G.M., Emerson, R.G., and Schroeder, C.E. (2011). Tuning of the human neocortex to the temporal dynamics of attended events. *J. Neurosci.* 31, 3176–3185.
- Bienenstock, E.L., Cooper, L.N., and Munro, P.W. (1982). Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J. Neurosci.* 2, 32–48.
- Bueti, D., Bahrami, B., Walsh, V., and Rees, G. (2010). Encoding of temporal probabilities in the human brain. *J. Neurosci.* 30, 4343–4352.
- Buonomano, D.V., and Laje, R. (2010). Population clocks: motor timing with neural dynamics. *Trends Cogn. Sci.* 14, 520–527.
- Bushara, K.O., Weeks, R.A., Ishii, K., Catalan, M.J., Tian, B., Rauschecker, J.P., and Hallett, M. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nat. Neurosci.* 2, 759–766.
- Buzsáki, G., and Wang, X.J. (2012). Mechanisms of gamma oscillations. *Annu. Rev. Neurosci.* 35, 203–225.
- Calderone, D.J., Lakatos, P., Butler, P.D., and Castellanos, F.X. (2014). Entrainment of neural oscillations as a modifiable substrate of attention. *Trends Cogn. Sci.* 18, 300–309.
- Cang, J., and Feldheim, D.A. (2013). Developmental mechanisms of topographic map formation and alignment. *Annu. Rev. Neurosci.* 36, 51–77.
- Carandini, M., Shimaoka, D., Rossi, L.F., Sato, T.K., Benucci, A., and Knöpfel, T. (2015). Imaging the awake visual cortex with a genetically encoded voltage indicator. *J. Neurosci.* 35, 53–63.
- Carrillo-Reid, L., Yang, W., Bando, Y., Peterka, D.S., and Yuste, R. (2016). Imprinting and recalling cortical ensembles. *Science* 353, 691–694.

- Cheng, R.K., Scott, A.C., Penney, T.B., Williams, C.L., and Meck, W.H. (2008). Prenatal-choline supplementation differentially modulates timing of auditory and visual stimuli in aged rats. *Brain Res.* 1237, 167–175.
- Chubykin, A.A., Roach, E.B., Bear, M.F., and Shuler, M.G. (2013). A cholinergic mechanism for reward timing within primary visual cortex. *Neuron* 77, 723–735.
- Cone, I., and Shouval, H.Z. (2021). Learning precise spatiotemporal sequences via biophysically realistic learning rules in a modular, spiking network. *eLife* 10, e63751.
- Cravo, A.M., Rohenkohl, G., Wyart, V., and Nobre, A.C. (2013). Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *J. Neurosci.* 33, 4002–4010.
- Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. *Nat. Rev. Neurosci.* 15, 732–744.
- Ekstrom, A.D., Kahana, M.J., Caplan, J.B., Fields, T.A., Isham, E.A., Newman, E.L., and Fried, I. (2003). Cellular networks underlying human spatial navigation. *Nature* 425, 184–188.
- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* 9, 474–480.
- Gavornik, J.P., and Bear, M.F. (2014). Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nat. Neurosci.* 17, 732–737.
- Gavornik, J.P., Shuler, M.G., Loewenstein, Y., Bear, M.F., and Shouval, H.Z. (2009). Learning reward timing in cortex through reward dependent expression of synaptic plasticity. *Proc. Natl. Acad. Sci. USA* 106, 6826–6831.
- Gerstner, W., Lehmann, M., Liakoni, V., Corneli, D., and Brea, J. (2018). Eligibility traces and plasticity on behavioral time scales: experimental support of NeoHebbian three-factor learning rules. *Front. Neural Circuits* 12, 53.
- Gu, Y., and Gong, P. (2016). The dynamics of memory retrieval in hierarchical networks. *J. Comput. Neurosci.* 40, 247–268.
- Hastings, M.H., Maywood, E.S., and Brancaccio, M. (2018). Generation of circadian rhythms in the suprachiasmatic nucleus. *Nat. Rev. Neurosci.* 19, 453–469.
- He, K., Huertas, M., Hong, S.Z., Tie, X., Hell, J.W., Shouval, H., and Kirkwood, A. (2015). Distinct eligibility traces for LTP and LTD in cortical synapses. *Neuron* 88, 528–538.
- Helfrich, R.F., Breska, A., and Knight, R.T. (2019). Neural entrainment and network resonance in support of top-down guided attention. *Curr. Opin. Psychol.* 29, 82–89.
- Helfrich, R.F., Schneider, T.R., Rach, S., Trautmann-Lengsfeld, S.A., Engel, A.K., and Herrmann, C.S. (2014). Entrainment of brain oscillations by transcranial alternating current stimulation. *Curr. Biol.* 24, 333–339.
- Henry, M.J., Herrmann, B., and Obleser, J. (2014). Entrained neural oscillations in multiple frequency bands modulate behavior. *Proc. Natl. Acad. Sci. USA* 111, 14935–14940.
- Holcman, D., and Tsodyks, M. (2006). The emergence of UP and Down states in cortical networks. *PLoS Comput. Biol.* 2, e23.
- Huertas, M.A., Hussain Shuler, M.G., and Shouval, H.Z. (2015). A simple network architecture accounts for diverse reward time responses in primary visual cortex. *J. Neurosci.* 35, 12659–12672.
- Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., and Smith, S.M. (2012). *Fsl. Neuroimage* 62, 782–790.
- Ji, D., and Wilson, M.A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nat. Neurosci.* 10, 100–107.
- Keitel, C., Thut, G., and Gross, J. (2017). Visual cortex responses reflect temporal structure of continuous quasi-rhythmic sensory stimulation. *Neuroimage* 146, 58–70.
- Kirkby, L.A., Luongo, F.J., Lee, M.B., Nahum, M., Van Vleet, T.M., Rao, V.R., Dawes, H.E., Chang, E.F., and Sohal, V.S. (2018). An amygdala-hippocampus subnetwork that encodes variation in human mood. *Cell* 175, 1688–1700.e14.
- Kösem, A., Bosker, H.R., Takashima, A., Meyer, A., Jensen, O., and Hagoort, P. (2018). Neural entrainment determines the words we hear. *Curr. Biol.* 28, 2867–2875.e3.
- Kravitz, D.J., Saleem, K.S., Baker, C.I., and Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nat. Rev. Neurosci.* 12, 217–230.
- Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., Kochunov, P.V., Nickerson, D., Mikiten, S.A., and Fox, P.T. (2000). Automated Talairach atlas labels for functional brain mapping. *Hum. Brain Mapp.* 10, 120–131.
- Lee, A.K., and Wilson, M.A. (2004). A combinatorial method for analyzing sequential firing patterns involving an arbitrary number of neurons based on relative time order. *J. Neurophysiol.* 92, 2555–2573.
- Levy, J.M., Zold, C.L., Namboodiri, V.M.K., and Hussain Shuler, M.G. (2017). The timing of reward-seeking action tracks visually cued theta oscillations in primary visual cortex. *J. Neurosci.* 37, 10408–10420.
- Li, J., Liao, X., Zhang, J., Wang, M., Yang, N., Zhang, J., Lv, G., Li, H., Lu, J., Ding, R., et al. (2017). Primary auditory cortex is required for anticipatory motor response. *Cereb. Cortex* 27, 3254–3271.
- Liu, C.H., Coleman, J.E., Davoudi, H., Zhang, K., and Hussain Shuler, M.G. (2015). Selective activation of a putative reinforcement signal conditions cued interval timing in primary visual cortex. *Curr. Biol.* 25, 1551–1561.
- Madisen, L., Garner, A.R., Shimaoka, D., Chuong, A.S., Klapoetke, N.C., Li, L., van der Bourg, A., Niino, Y., Egolf, L., Monetti, C., et al. (2015). Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. *Neuron* 85, 942–958.
- Madisen, L., Mao, T., Koch, H., Zhuo, J.M., Berenyi, A., Fujisawa, S., Hsu, Y.W., Garcia, A.J., 3rd, Gu, X., Zanella, S., et al. (2012). A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. *Nat. Neurosci.* 15, 793–802.
- Makino, H., and Komiyama, T. (2015). Learning enhances the relative impact of top-down processing in the visual cortex. *Nat. Neurosci.* 18, 1116–1122.
- Matell, M.S., and Meck, W.H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Brain Res. Cogn. Brain Res.* 21, 139–170.
- Mathewson, K.E., Gratton, G., Fabiani, M., Beck, D.M., and Ro, T. (2009). To see or not to see: prestimulus alpha phase predicts visual awareness. *J. Neurosci.* 29, 2725–2732.
- Mauk, M.D., and Buonomano, D.V. (2004). The neural basis of temporal processing. *Annu. Rev. Neurosci.* 27, 307–340.
- Meck, W.H., Penney, T.B., and Pouthas, V. (2008). Cortico-striatal representation of time in animals and humans. *Curr. Opin. Neurobiol.* 18, 145–152.
- Merchant, H., Harrington, D.L., and Meck, W.H. (2013). Neural basis of the perception and estimation of time. *Annu. Rev. Neurosci.* 36, 313–336.
- Mi, Y., Liao, X., Huang, X., Zhang, L., Gu, W., Hu, G., and Wu, S. (2013). Long-period rhythmic synchronous firing in a scale-free network. *Proc. Natl. Acad. Sci. USA* 110, E4931–E4936.
- Monk, K.J., Allard, S., and Hussain Shuler, M.G. (2020). Reward timing and its expression by inhibitory interneurons in the mouse primary visual cortex. *Cereb. Cortex* 30, 4662–4676.
- Namboodiri, V.M., Huertas, M.A., Monk, K.J., Shouval, H.Z., and Hussain Shuler, M.G. (2015). Visually cued action timing in the primary visual cortex. *Neuron* 86, 319–330.
- Neske, G.T. (2015). The slow oscillation in cortical and thalamic networks: mechanisms and functions. *Front. Neural Circuits* 9, 88.
- Niell, C.M. (2015). Cell types, circuits, and receptive fields in the mouse visual cortex. *Annu. Rev. Neurosci.* 38, 413–431.
- Nobre, A.C., and van Ede, F. (2018). Anticipated moments: temporal structure in attention. *Nat. Rev. Neurosci.* 19, 34–48.
- O'Connor, D.H., Wittenberg, G.M., and Wang, S.S. (2005). Graded bidirectional synaptic plasticity is composed of switch-like unitary events. *Proc. Natl. Acad. Sci. USA* 102, 9679–9684. <https://doi.org/10.1073/pnas.0502332102>.

- Paton, J.J., and Buonomano, D.V. (2018). The neural basis of timing: distributed mechanisms for diverse functions. *Neuron* 98, 687–705.
- Pecenka, N., Engel, A., and Keller, P.E. (2013). Neural correlates of auditory temporal predictions during sensorimotor synchronization. *Front. Hum. Neurosci.* 7, 380.
- Penney, T.B., Gibbon, J., and Meck, W.H. (2008). Categorical scaling of duration bisection in pigeons (*Columba livia*), mice (*Mus musculus*), and humans (*Homo sapiens*). *Psychol. Sci.* 19, 1103–1109.
- Pitkow, X., and Meister, M. (2012). Decorrelation and efficient coding by retinal ganglion cells. *Nat. Neurosci.* 15, 628–635.
- Redgrave, P., and Gurney, K. (2006). The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.* 7, 967–975.
- Ruskin, D.N., Bergstrom, D.A., Kaneoke, Y., Patel, B.N., Twery, M.J., and Walters, J.R. (1999). Multisecond oscillations in firing rate in the basal ganglia: robust modulation by dopamine receptor activation and anesthesia. *J. Neurophysiol.* 81, 2046–2055.
- Salvioni, P., Murray, M.M., Kalmbach, L., and Buetti, D. (2013). How the visual brain encodes and keeps track of time. *J. Neurosci.* 33, 12423–12429.
- Shouval, H.Z., Bear, M.F., and Cooper, L.N. (2002). A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. *Proc. Natl. Acad. Sci. USA* 99, 10831–10836.
- Shuler, M.G. (2016). Timing in the visual cortex and its investigation. *Curr. Opin. Behav. Sci.* 8, 73–77.
- Shuler, M.G., and Bear, M.F. (2006). Reward timing in the primary visual cortex. *Science* 311, 1606–1609.
- Soares, S., Atallah, B.V., and Paton, J.J. (2016). Midbrain dopamine neurons control judgment of time. *Science* 354, 1273–1277.
- Teki, S., Grube, M., Kumar, S., and Griffiths, T.D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *J. Neurosci.* 31, 3805–3812.
- Thaut, M.H. (2003). Neural Basis of Rhythmic Timing Networks in the Human brain. *Ann. N. Y. Acad. Sci.* 999, 364–373.
- Thut, G., Nietzel, A., Brandt, S.A., and Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J. Neurosci.* 26, 9494–9502.
- Toda, K., Lusk, N.A., Watson, G.D.R., Kim, N., Lu, D., Li, H.E., Meck, W.H., and Yin, H.H. (2017). Nigroretectal stimulation stops interval timing in mice. *Curr. Biol.* 27, 3763–3770.e3.
- Townsend, R.G., Solomon, S.S., Martin, P.R., Solomon, S.G., and Gong, P. (2017). Visual motion discrimination by propagating patterns in primate cerebral cortex. *J. Neurosci.* 37, 10074–10084.
- Turrigiano, G. (2011). Too many cooks? Intrinsic and synaptic homeostatic mechanisms in cortical circuit refinement. *Annu. Rev. Neurosci.* 34, 89–103.
- Veliz-Cuba, A., Shouval, H.Z., Josić, K., and Kilpatrick, Z.P. (2015). Networks that learn the precise timing of event sequences. *J. Comput. Neurosci.* 39, 235–254.
- Voloh, B., and Womelsdorf, T. (2016). A role of phase-resetting in coordinating large scale neural networks During attention and goal-directed behavior. *Front. Syst. Neurosci.* 10, 18.
- Xu, M., Zhang, S.Y., Dan, Y., and Poo, M.M. (2014). Representation of interval timing by temporally scalable firing patterns in rat prefrontal cortex. *Proc. Natl. Acad. Sci. USA* 111, 480–485.
- Xu, S., Jiang, W., Poo, M.M., and Dan, Y. (2012). Activity recall in a visual cortical ensemble. *Nat. Neurosci.* 15, 449–455.
- Yao, H., Shi, L., Han, F., Gao, H., and Dan, Y. (2007). Rapid learning in cortical coding of visual scenes. *Nat. Neurosci.* 10, 772–778.
- Zold, C.L., and Hussain Shuler, M.G. (2015). Theta oscillations in visual cortex emerge with experience to convey expected reward time and experienced reward rate. *J. Neurosci.* 35, 9603–9614.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE                           | SOURCE                                 | IDENTIFIER                                                                                                                                    |
|-----------------------------------------------|----------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Antibodies</b>                             |                                        |                                                                                                                                               |
| Rabbit anti-DsRed                             | Takara Clontech, Japan                 | RRID: AB_10013483                                                                                                                             |
| Rhodamine (TRITC)-donkey anti-rabbit          | Jackson ImmunoResearch                 | RRID: AB_2340588                                                                                                                              |
| <b>Deposited data</b>                         |                                        |                                                                                                                                               |
| Original data for main figures                | This paper                             | <a href="https://doi.org/10.17632/smvt9jtx53.1">https://doi.org/10.17632/smvt9jtx53.1</a>                                                     |
| <a href="#">Supplemental information</a>      | This paper                             | <a href="https://doi.org/10.17632/z629248f2x.1">https://doi.org/10.17632/z629248f2x.1</a>                                                     |
| <b>Experimental models: Organisms/strains</b> |                                        |                                                                                                                                               |
| Mouse: C57BL/6                                | Slac Laboratory Animal Co.             | N/A                                                                                                                                           |
| Mouse: CaMKII $\alpha$ -Cre                   | Jackson Laboratory                     | RRID: IMSR_JAX:005359                                                                                                                         |
| Mouse: PV-Cre                                 | <a href="#">Madisen et al., 2012</a>   | N/A                                                                                                                                           |
| Mouse: loxP-flanked-ChR2-tdTomato             | Jackson Laboratory                     | RRID: IMSR_JAX:012567                                                                                                                         |
| Human: epileptic patients                     | Huashan Hospital                       | N/A                                                                                                                                           |
| <b>Software and algorithms</b>                |                                        |                                                                                                                                               |
| Original codes                                | This paper                             | <a href="https://doi.org/10.17632/cjhm2n625z.1">https://doi.org/10.17632/cjhm2n625z.1</a>                                                     |
| Adobe Illustrator CC                          | Adobe                                  | <a href="http://www.adobe.com/products/illustrator.html">http://www.adobe.com/products/illustrator.html</a> ; RRID:SCR_010279                 |
| EEG-1200C                                     | Nihon Kohden                           | <a href="https://www.nihonkohden.com/">https://www.nihonkohden.com/</a>                                                                       |
| FMRIB Software Library                        | <a href="#">Jenkinson et al., 2012</a> | RRID:SCR_002823                                                                                                                               |
| MATLAB                                        | MathWorks                              | <a href="https://matlab.mathworks.com">https://matlab.mathworks.com</a> ; RRID:SCR_001622                                                     |
| Neuralynx system                              | Neuralynx Inc.                         | <a href="https://neuralynx.com/">https://neuralynx.com/</a>                                                                                   |
| Neurofax EEG-1200 acquisition system          | Nihon Kohden                           | <a href="https://www.nihonkohden.com/">https://www.nihonkohden.com/</a>                                                                       |
| Origin software                               | OriginLab Corporation                  | <a href="http://www.originlab.com/index.aspx?go=PRODUCTS/Origin">http://www.originlab.com/index.aspx?go=PRODUCTS/Origin</a> ; RRID:SCR_014212 |
| Psychopy                                      | Open-resource                          | <a href="https://www.psychopy.org/">https://www.psychopy.org/</a> ; RRID:SCR_006571                                                           |
| Python                                        | Python                                 | <a href="http://www.python.org/">http://www.python.org/</a> ; RRID:SCR_008394                                                                 |
| Spike2 system                                 | Cambridge Electronic Design            | <a href="https://ced.co.uk/">https://ced.co.uk/</a> ; RRID:SCR_000903                                                                         |
| Talairach Daemon software                     | <a href="#">Lancaster et al., 2000</a> | RRID:SCR_000448                                                                                                                               |
| TRACKPixx3                                    | VPixx Technologies Inc.                | <a href="https://vpixx.com/">https://vpixx.com/</a>                                                                                           |
| Virtual reality system                        | PHENOSYS                               | <a href="https://www.phenosys.com/">https://www.phenosys.com/</a> ; RRID:SCR_004072                                                           |

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jiayi Zhang ([jiayizhang@fudan.edu.cn](mailto:jiayizhang@fudan.edu.cn)).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

Related data have been deposited at Mendeley Data and are publicly available as of the date of publication. DOIs are listed in the [key resources table](#).

Original code for data analysis and computational model has been deposited at Mendeley Data and is publicly available as of the date of publication. DOIs are listed in the [key resources table](#).

Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.



## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Participants

Participants included 11 patients (7 males and 4 females; age:  $25.6 \pm 5.3$  years old) with drug-controllable epilepsy. They were implanted with chronically intracranial depth electrodes as a preoperative procedure to evaluate epilepsy surgery or to localize seizure focus for potential surgical resection. All the parameters of implanted electrodes were determined by clinical surgery need rather than the needs of our research. All patients were treated in Huashan Hospital (Shanghai, China), Fudan University (Shanghai, China) and signed informed written consent to participate in the study. All procedures were approved by Huashan Hospital's institutional review board.

### Animals

All animal care and experiments were performed in accordance with Fudan University Shanghai Medical College Institutional Animal Care and Use Committee guidelines. Wild-type (C57BL/6J) mice were obtained from the Slac Laboratory Animal Co. (Shanghai, China). CaMKII $\alpha$ -Cre mice (Jackson lab, US. Stock #005359, RRID: IMSR\_JAX:005359) and PV-Cre mice (Madisen et al., 2012) were crossed with loxP-flanked-ChR2-tdTomato mice (Jackson lab, US. Stock #012567, RRID: IMSR\_JAX:012567). Mice pups were bred with a housing group of no more than 3 adult mice. 8-12 weeks male mice were used in the experiments, separately reared in a housing group of no more than 5. All mice were maintained in a 12-hour light-dark cycle at 18-22°C. Mice were numbered blindly for behavioral training or electrophysiological recording experiments. All experiments were conducted during the light cycle.

## METHOD DETAILS

### Experimental design

We designed 2 timing perception tasks for each participant, named tapping and grating. In each task, participants needed to focus on the center of a monitor (13-inch MacBook pro, Macintosh) that was placed about 50 cm in front of the patient's eyes and displayed visual stimuli. Visual stimuli were generated by custom-written codes in the Psychopy software (a Python-based psychophysics package). These periodic drifting gratings were square waves with 100% contrast, temporal frequency of 2 Hz and spatial frequency of 0.05 cycles per degree. Different time intervals (5-s or 10-s intervals; 1s duration, followed by 4-s or 9-s grey background;  $n = 20$ ) were used in the study. The participants were unaware of the time intervals throughout the experiments. In the tapping task, participants were instructed to tap the spacebar of the keyboard at the exact time point when they predicted that gratings would appear. The accurate tapping and grating time were imported in real time into a stereo-electroencephalography (SEEG) recording system by custom-written codes in Arduino platform. In the grating task, participants were instructed to stay relaxed and look at the monitor screen to avoid overthinking as much as possible.

### Eye movements

TRACKPixx3 (VPixx Technologies Inc., Montreal, Canada) was used, according to product instructions, to track the eye movements of participant Y when performing tasks. The pupil center and iris of two eyes were recognized by the software during recording. The vertical eye movements were calculated in further analysis using custom-written MATLAB codes.

### SEEG recording and preprocessing

SEEG signals were acquired using a clinical recording system (EEG-1200C, Nihon Kohden, Irvine, CA) with a sampling rate of 2000 Hz and were band-pass filtered between 0.5 Hz and 600 Hz using Neurofax EEG-1200 acquisition system (NIHON KOHDEN Corporation, Japan). After signal recording, signals were notch-filtered to eliminate 50 Hz noise and harmonics. All SEEG data were collected when the participants did not show any epileptic symptoms. Data of 11 participants were collected for 7 consecutive sessions. Two tasks were performed in the sequence of tapping and grating groups for each session.

### Electrode location

The locations of each SEEG electrode contact were extracted from pre-implant MRI and postoperative CT scans before further analysis. We used FMRIB Software Library (FSL) (Jenkinson et al., 2012) for co-registration to MNI space and marked each electrode contact manually. Then we remapped the MNI coordinates into the Talairach space with Bioelectromagnetism toolbox (<http://eeg.sourceforge.net/bioelectromagnetism.html>) in MATLAB and extracted the Talairach label with the Talairach Daemon software (Lancaster et al., 2000). All electrodes were classified into discrete brain areas based on the Brodmann area and their functions.

### Frequency decomposition and power

Signals were analyzed mostly by custom-written scripts in MATLAB and some functions in EEGLAB. First, signals were cut into event-related epochs, including 8 s before event marker and 7 s after event marker. The segmented data were zero-padded to minimize filter-induced edge effects. A wavelet transformation ranged from 2 Hz to 150 Hz through 30 logarithmically spaced frequencies with the complex-valued Morlet wavelet was performed on raw signals of each segment. The event-induced power was calculated by squaring the analytic amplitude envelope and was normalized to the pre-trial baseline (-2 s to -1 s) using the relative change in decibel

conversion (dB). We then averaged through the same brain area across all the participants. To examine whether the event-induced power differed between the timing prediction task and the visual task, we quantified the significance of an index between the two tasks. The index is defined as:

$$\frac{\text{power}_{-2\sim 0s} - \text{power}_{0\sim 2s}}{|\text{power}_{-2\sim 0s}| + |\text{power}_{0\sim 2s}|}$$

Power<sub>-2~0s</sub> is the average power from -2 s to 0 s.

In addition, to further explore the percentage of power-increased channels, we calculated Z-scores for theta band (3-8 Hz), alpha band (8-12 Hz), beta band (12-30 Hz), low gamma band (30-80 Hz) and high gamma band (80-150 Hz), respectively. The channels with absolute Z-scores no less than 1.96 before time 0 were regarded as being significant.

### Inter-trial phase clustering (ITPC)

The strength of inter-trial neural synchrony was quantified by inter-trial phase clustering (ITPC):

$$\text{ITPC}_{\text{tf}} = \left| n^{-1} \sum_{\text{trial}=1}^n e^{ik_{\text{tf}}} \right|$$

where  $n$  is the number of trials and  $e^{ik_{\text{tf}}}$  is from Euler's formula and provides the complex polar representation of a phase angle  $k$  on a certain trial at time-frequency point  $\text{tfr}$ . ITPC measures the extent to which the phase angles of different trials are clustered at each time-frequency point in polar space, with values approaching 1 indicating stronger phase synchrony. For each brain region, the ITPC was then averaged by channels across all the participants. To test the statistical significance of ITPC, the observed ITPC was compared with a null distribution, which was created by repeating the procedure of shuffling the labels of time series and computing ITPC of the permuted signal for 500 times. The observed ITPC was considered significant when it was above the threshold, which was the 95<sup>th</sup> percentile of the surrogate ITPC. The non-significant ITPC was set to 0. Then we counted the number of significant points and defined it as the area during periods of -2~0 s and 0~2 s in both timing and visual tasks, in 4 oscillation bands (delta band, theta band, alpha band and beta band) respectively. A similar analysis was also performed when computing ISPC.

### Inter-site phase clustering (ISPC)

The phase-based connectivity between two brain regions was measured by inter-site phase clustering (ISPC):

$$\text{ISPC}_{\text{tf}} = \left| n^{-1} \sum_{\text{trial}=1}^n e^{i(\varphi_{\text{atf}} - \varphi_{\text{btf}})} \right|$$

It calculates the phase angle difference  $\varphi_{\text{atf}} - \varphi_{\text{btf}}$  between the channel pair (a, b) at time point  $t$  and frequency  $f$  averaged across  $n$  trials. For each brain region, the ISPC was then averaged by channel pairs across all the participants. Then the procedure of shuffling the labels of time series and computing ISPC of the permuted data was repeated for 500 times, and a null distribution was created. The observed ISPC was compared with the null distribution and the part that exceeded 95<sup>th</sup> percentile of the surrogate ISPC was considered significant. The non-significant ISPC was set to 0. And we counted the number of significant points and defined it as the area during periods of -2~0 s and 0~2 s in both timing and visual tasks, in 4 oscillation bands (delta band, theta band, alpha band and beta band) respectively.

### Virtual reality (VR) system

The virtual reality system applied to our behavioral experiments was custom-built based on a commercial virtual reality system (PHENOSYS, Germany). Our virtual reality system included five TFT monitors, 200 mm jet balls, a ball holder with a X/Y motion sensor, a controllable device with air puff and water rewards. The jet ball was located on an air-cushioned spherical treadmill allowing head-restrained mice to navigate in Y direction. Five TFT surround monitors (19-inch, resolution 1280 × 1024, refresh rate 75 Hz, 33 cm away from the eyes) displayed custom-written images that covered 216 degrees of the visual field. *PhenoSoft VR* software controlled the virtual reality that was displayed on the surrounding TFT monitors by a series of parameters defined in an XML configuration file, such as the size, shapes, colors or textures of the walls. In our behavioral experiment, we designed a 400-meter-long corridor with low-contrast (50%) grating images. Another software *Phenosoft control* delivered programmed water rewards or air puffs and recorded the accurate position of the mice in the virtual corridor. The water reward was delivered through a tube driven by a peristaltic pump at a rate of 50  $\mu\text{L}$  per second.

### Visual stimuli presentation

Visual stimuli were generated by custom-written codes in Psychopy software (a Python-based psychophysics package), containing periodic drifting gratings (1-s duration, followed by 4-s or 9-s grey background, 5-s or 10-s interval). These stimuli were square waves with 100% contrast, temporal frequency of 2 Hz and spatial frequency of 0.05 cycles per degree. In VR experiments, drifting gratings were delivered via five surrounding TFT monitors. In non-VR experiments, drifting gratings were delivered via an LCD monitor (150 × 90 mm<sup>2</sup>, 800 × 480 pixels, 60 Hz. FEELWORLD, China) placed 8 cm from the eye.

473 nm blue laser (DPSSL DRIVER, China) was used to optogenetically manipulate the neural activities of V1 in PV-ChR2 and CaMKII $\alpha$ -ChR2 mice. In PV-ChR2 experiments, we enucleated one eyeball and implanted a LED onto the surface of V1 ipsilateral to the enucleated eye to minimize the ipsilateral influence when inhibiting contralateral V1 activities. We provided two types of optogenetic inhibition patterns: one was 20 periodic blue laser stimuli with 2.5-s duration for 5-s intervals and 5-s duration for 10-s intervals to inhibit the spontaneous and visual-evoked activities of V1 by periodic gratings; the other was 30s continuous blue laser stimuli to V1 after drifting gratings. In CaMKII $\alpha$ -ChR2 experiments, two optogenetic activation patterns were implemented: one was periodic blue laser stimuli with 1-s duration for both 5-s and 10-s intervals to excite V1 activities; the other was 5 random-interval blue laser stimuli with 0.5-s duration to V1 30 s and 3 min after drifting gratings, respectively.

### Stereotaxic surgery and LED implantation

All mice were anesthetized with isoflurane and placed into a stereotaxic apparatus (RWD, China). 1% lidocaine (10 mg/mL lidocaine in saline, MP Biomedicals, US) was subcutaneously injected under the scalp. After removal of the scalp, the skull was dried and roughened. A titanium alloy head bar (about 5 × 30 mm) was implanted on left hemisphere with two skull screws and dental cement (C&B Metabond, Parkell, US). A craniotomy window in the skull of the right hemisphere was created stereotactically in V1 (3.1–4.2 mm posterior to bregma and 2.0–3.2 mm lateral) and the dura was removed carefully. Fresh warm (37 °C) low-gelling temperature agarose (1.5% in sterile buffered saline, A9414, Sigma, US) covered the craniotomy. A custom-built LED (1.6 × 0.8 mm,  $\lambda$  = 473 nm, King-bright, China) was placed onto the surface of exposed V1 and sealed to the skull by dental cement. Mice were given at least 3 days to recover from surgery before behavioral experiments.

### Prediction tasks in VR system

Mice implanted with head bars were put on a water-restriction schedule, in which they only received 1–1.5 mL water per day in three consecutive days. Their body weight was checked daily to maintain no less than 80% before water restriction. Head-restrained mice were first trained to run on a jet ball in 100 meters virtual corridor. 200  $\mu$ L water rewards were delivered after mice arrived at the destination. Each mouse was trained for 3–5 days until it reached the destination in 5 minutes for three consecutive sessions. Mice were then trained to seek rewards with timing tasks in 4 training phases: 1) running 60s in the corridor to accustom to the behavioral environment; 2) 20 high-contrast drifting gratings (details described previously) coupled with a half drop of water reward (about 0.02 mL) were presented on two opposite monitors; 3) the prediction window with grey background were presented for 60s; 4) another 60s running in the corridor. Mice were trained for at least 6 sessions per day during 7 training days. In the reward-only task (R), water rewards were periodically presented 20 times with 5-s or 10-s intervals in the virtual corridor. In the optogenetic manipulation task, head-fixed PV-ChR2 and CaMKII $\alpha$ -ChR2 mice were trained with blue laser stimuli presented to V1 (similar to Timing or R task, see [visual stimuli presentation](#) for detailed stimulation pattern). During all behavioral experiments, animal behaviors were monitored with infrared cameras (JAI, Denmark) or a digital video recorder (Sony, Japan), with a sampling rate at 25 frames per second. The 5-s and 10-s tasks were all conducted in different cohorts of mice.

### In vivo extracellular e-phys recording

In awake-behaving recordings, we performed two different recording paradigms: chronic recording and acute recording. For acute recording, we trained at least 8 head-fixed mice as a group, and randomly chose one of them to acutely record its V1 activities each day. By day 7, V1 activities of all remaining mice (in general, 2 or 3 mice were the rest) were acutely recorded V1 activities. In awake and anesthetic recordings, mice were deeply anesthetized with 2.5% isoflurane in oxygen for 20 min before surgery. The nose and mouth were placed in a respiratory mask and mice were kept on a heating pad (FHC Inc., US), followed by subcutaneous injection of 1% lidocaine under the scalp. The mice were fixed in a holder and placed on a braked jet ball. After removal of the scalp, a craniotomy window in the skull was created stereotactically in V1, and the dura was removed carefully. Microelectrode arrays were stereotactically inserted into L2/3 of V1 to collect neural activities. Mineral oil covered the whole exposed V1 to keep the brain moist, while a warm melted mixture of wax and mineral oil covered the recording sites to eliminate the movement artifact of mice. The electrophysiological recording began after mice were completely awake after surgery. None of the mice cohorts were trained with cues and rewards at 5 s or 10 s interval before the anesthetic experiments. Mice with water restriction for at least one day were trained to drink water with a tube on a VR setup for about two days. We randomly presented water drops to mice with intervals ranging from 2 s to 20 s such that no mice experienced rewards with fixed intervals at 5 s or 10 s prior to the recording. None of the anesthetized mice received any visual cues before acute recording.

For chronic recording, we implanted a microelectrode array (16 channels, 125- $\mu$ m intercontact spacing, Microprobes, US) into the primary visual cortex in the right hemisphere of head-fixed mice. Firstly, the scalp was removed, and the skull was dried and roughened. Then two stainless screws, which were soldered with 2 cm silver wire, were inserted into the front and tail of the skull as ground and reference electrodes. A craniotomy window in the skull was drilled in V1, ranging about 2 × 2 mm, and then the dura was removed carefully. The microelectrode array was micro-managed into the craniotomy of layer 2/3 of V1, safeguarded with low-gelling temperature agarose and Kwik-Sil and Kwik-Cast silicones. When the agarose and silicones became gel phase, we used dental cement to fasten all the devices on the skull, including the head bar, screws and the microelectrode array. A cylinder copper mesh flapping around the probes was clinging to the surface of the skull. The ground and reference electrodes from the two screws

were soldered onto the ground lines of the microelectrode array, and then soldered on the cylinder copper mesh to shield all environmental noise. After recovery of at least 5 days, the behavioral training schedule and neural activity collection were conducted.

All electrophysiological data were amplified and digitized using the Spike2 system (Cambridge Electronic Design, UK) or the Neuralynx system (Neuralynx Inc., US). Neural signals in V1 were recorded continuously with a range of 0.3–5,000 Hz and sampled at 30 kHz; spiking signals were band-pass filtered at 300–5,000 Hz and local field potentials were band-pass filtered at 0.3–300 Hz.

### Potentiation analysis

To calculate the duration of entrained LFP as well as predictive licking behaviors, we plotted the average licking rate or number of entrained LFP with a 1 s bin size in the prediction window. An exponential decay fitting function was performed to fit the average firing curve and to obtain the time constant of exponential decay. The exponential decay fitting function was:  $y = y_0 + A_1 \cdot \exp(-x/t_1)$ , where  $y_0$  = offset;  $A$  = amplitude;  $t_1$  = time constant. Here,  $t_1$  is the decay time constant.

### In vivo cell-attached patch-clamp recording

We performed *in vivo* cell-attached patch-clamp recording in CaMKII $\alpha$ -Chr2 mice. Firstly, we prepared the recording pipettes with optimum resistance 5–7 M $\Omega$  by a filament micropipette puller (Sutter, US) a few hours before the experiment. The V1 area in anesthesia CaMKII $\alpha$ -Chr2 mice was exposed and covered with warm (37°C) sterile buffered saline. Recording pipettes filled with artificial cerebrospinal fluid (ACSF, 124 mM NaCl, 3 mM KCl, 2 mM CaCl<sub>2</sub>·2H<sub>2</sub>O, 1 mM MgCl<sub>2</sub>, 1.25 mM NaH<sub>2</sub>PO<sub>4</sub>, 26 mM NaHCO<sub>3</sub>, 10 mM Glucose) were exerted about 300 mbar positive pressure and moved forward into V1 by a micro-manipulator (Sutter, US). The positive pressure in the recording pipette was tuned down to about 30 mbar once in the brain and moved at 10  $\mu$ m per second. The resistance of the recording pipette fluctuated dramatically when it touched the soma or axon of a neuron. An increase in pipette resistance (stable at about 15–20 M $\Omega$ ) indicated that the pipette was approaching a nearby neuron. Positive pressure was quickly removed and retained for a few seconds. Negative pressure at about -30 mbar was applied if the resistance stopped increasing. Experiments started at about 50–300 M $\Omega$ , indicating close attachment between the recording pipette and the cell membrane.

### Immunohistochemistry

PV- or CaMKII $\alpha$  mice were deeply anesthetized by overdose isoflurane and transcardially perfused with 4% paraformaldehyde afterwards. The brain was dissected and fixed in 4% paraformaldehyde overnight at 4°C. The brain was dehydrated with 30% Sucrose in PB (324 mM Na<sub>2</sub>HPO<sub>4</sub>·12H<sub>2</sub>O, 76 mM NaH<sub>2</sub>PO<sub>4</sub>·2H<sub>2</sub>O) overnight at 4°C, mounted in O.C.T at -80°C for 30 min, and sectioned at 30  $\mu$ m using a cryostat (LEICA CM1950; Germany). Brain slices were washed with TBS (12.4 mM Tris, 37.7 mM Tris-HCl, 154.0 mM NaCl, PH = 7.4) for 4 times, permeabilized with 0.5% Triton-X100 in TBS for 20 min at room temperature, and then blocked with 10% donkey serum and 2 mg ml<sup>-1</sup> BSA in TBS containing 0.05% Triton-X100 for 1–2 hour at room temperature. After incubating in the primary antibody (Rabbit anti-DsRed; Takara Clontech, 632496, Japan, RRID: AB\_10013483; 1:1000 for 24 hours, respectively) at 4°C, the slices were washed in TBS for 4 times every 15 min. Slices were then incubated for 4 hours at room temperature in the secondary antibody (Rhodamine (TRITC)-donkey anti-rabbit, 711-025-152; 1:200; Jackson ImmunoResearch, US, RRID: AB\_2340588). These fluorescently stained samples were washed in TBS for 15 min, counterstained with DAPI (Sigma, 200 ng ml<sup>-1</sup>) for 3–5 min followed by 4 washes in TBS every 10 min and mounted with AQUA-MOUNT (Thermo Scientific, US).

### Automated detection of licking behavior

We developed an automatic licking-behavior algorithm based on the core package OpenCV in Python. The behavioral movies were imported into Python and separated into single frames. The position of Region of Interest for licking (ROI<sub>licking</sub>, 2 × 2 pixels) was determined around the jaws of the mice. The position of ROI for control (ROI<sub>control</sub>, 2 × 2 pixels) was 15 pixels below ROI<sub>licking</sub> with the bright background area. In general, the luminance of the jaw or the tongue was darker than that of control areas. Customized codes in Python automatically compared the luminescence of ROI<sub>licking</sub> and ROI<sub>control</sub> for all frames and sorted the frames into licking behavior when luminescence of ROI<sub>licking</sub> was less than 90% of that of ROI<sub>control</sub>. The rest frames were sorted as non-licking behaviors. The intervals between two successive licking behaviors were calculated. If the intervals of two licking frames were smaller than 1 s, they were considered to be from the same licking. Customized codes in Python printed the licking intervals in CSV format.

### Detection of upstate from PSD of LFPs

Upstate-like activities were identified from the LFP signals. Mice with no observable upstate-like events were excluded from the analysis, typically smaller than 1-fold of the signal to noise ratio. Time-dependent power spectra density of LFP signals (band-passed range was 0.3–300 Hz) were calculated in Spike2 software, with the frequency resolution 0.5 Hz and the sample interval 0.25 s. An automatic thresholding MATLAB program (custom-written, MathWorks, US) calculated the event-detecting thresholds for the power spectra density. Briefly, the data were initially segmented into two parts using a randomly-selected threshold. A new threshold was then computed as the average of the means of these two parts. This process was repeated until the threshold stopped changing. The starting and ending time points when the power was larger than the threshold value were recorded. The event intervals were calculated as the time difference between two neighboring starting time points. In order to identify a valid event, the following criteria were used: 1) to exclude stimulation and movement artifacts, the duration of an event was larger than 0.15 s; 2) the time difference between two neighboring events was larger than 0.4 s. To visualize the distribution of LFP events, we counted the number of LFP



events in every 1 s time bin each training day and calculated the percentage of LFP events of each bin over the total number of LFP events that day. We then normalized the LFP percentage by dividing the percentage of LFP events of each time bin by the largest percentage of LFP events over 7 training days.

### Pairwise correlation analysis

Pairwise correlation analysis was calculated between the licking behavior and PSD in the Origin software (OriginLab Corporation, US). PSD was calculated from LFPs by Fast Fourier Transformation in Spike2 software and the onset of the licking behavior was changed to continuous curves by a kernel of Gaussian fitting in Spike2 software. Based on our previous successive predictions data, we chose 30-s PSD and licking data to calculate the pairwise correlation. We chose 30-s spontaneous data before training as pre, the last 3 stimuli data as during and first 30-s after stimuli as post training to compare the dynamics between PSD and the licking behavior. Both PSD and licking data were imported into Origin, and normalized to [0, 1] according to each data set. The cross-correlation curves were computed between licking curves and PSD in pre, during and post conditions in Origin. Peaks in correlation coefficient curves were detected using a Peak Analyzer toolbox in Origin, which offers an interactive interface to walk us through the steps of advanced peak analysis. All detected peaks were plotted using probability density distribution to visualize the peaks distribution around stimulation interval, such as 0,  $\pm 10$ ,  $\pm 20$ ,  $\pm 30$  for 10-s intervals and 0,  $\pm 5$ ,  $\pm 10$ ,  $\pm 15$ ,  $\pm 20$ ,  $\pm 25$ ,  $\pm 30$  for 5-s intervals.

### Matching analysis of firing sequence

To compare the firing sequence of neurons in two different periods, we used a combinatorial method. Firstly, we calculated the spikes number of each bin from the PSTHs of neuron spiking in two different periods. The bin size of a PSTH was 0.1 s and width time was 4 s for 5 s interval or 9 s for 10 s interval stimuli. Secondly, we plotted the histogram distribution of spikes numbers, and calculated the firing peak time of different periods using Gaussian peak-fitting function in Origin. Thirdly, we set the firing sequence of one period as a template by sorting the firing peak time of each neuron, and then ranked the firing sequence of another period across the same neuron in the template. For example, we set the firing sequence of last 20-11 as the template, which was 123456789(10)(11)(12)(13)(14) for 10-s intervals and 123456789(10)(11)(12) for 5-s intervals, and then ranked the firing sequence of last 10-1 stimuli as 123569(10)87(11)(12)(13)(14) for 10-s intervals and 1245387109(11)(12) for 5-s intervals. We then defined a matching index  $I$  to measure the similarity between the template sequence and the ranked sequence. For example, we gave an example of dataset with a number of  $M$  neurons in a frame sequence and combined a total of  $M(M-1)/2$  neuron pairs. If  $m$  was the number of pairs with the same order in the frame and  $n$  was the number of pairs with a different order in the frame, we defined the matching index  $I$  as  $(m - n) / (m + n)$ , ranging from -1 to 1. When  $I$  was equal to 1, it indicated that the firing sequence of the neurons in the frame in two periods were exactly the same; when  $I$  was equal to -1, it indicated that they were completely opposite. We also defined a matching probability  $p_{match}$  to measure how significant the matching was, which is the proportion of a matching index equal or larger than the defined index  $I$  among all potential sequences. For small  $M$  ( $M \leq 11$ ), we actually generated all combinations of  $M$  neurons and calculated all matching index  $I$  and matching probabilities  $p_{match}$ . For large  $M$  ( $M > 11$ ), we had to randomly generate 10000 possible combinations as approximate data set because it was too large to generate all combinations. Unless otherwise specified, we used a threshold  $p_{match} < 0.05$  as a criterion to determine the significance between two sequences. We also randomly generated 3 sequences for each template sequence as ranked sequence, named shuffle. We computed the average matching index  $I$  and matching probability  $p_{match}$  from 3 shuffled ranked-sequence as control data.

### Computational model for timing activity in V1

In Figure 7, we simulated a two-layer network. The first layer of the network modeled V1, in which neurons were connected through synapses with short-term depression (STD). The mean activity of this network was approximated by the following dynamic system (Holcman and Tsodyks, 2006):

$$\tau \frac{dV}{dt} = -V + \mu U w_r r(V) + I_{time} + I_{stim} + I_b + \zeta \quad (\text{Equation 1})$$

$$\frac{d\mu}{dt} = \frac{1 - \mu}{\tau_{STD}} - \mu U r(V) \quad (\text{Equation 2})$$

where  $V$  is the average synaptic input (measured as the change of membrane potential in response to it, in unit of mV with a baseline at 0 mV) with the time constant  $\tau = 0.08$  s;  $w_r = 31.9$  mV/Hz is the average strength of the recurrent synapses in V1;  $U = 0.8$  and  $\tau_{std} = 0.8$  s are utilization parameter and recovery time constant of STD;  $\mu$  is the depression parameter of STD;  $I_b = 1.4$  mV is background input; and  $\zeta$  summarizes all sources of noise.  $r(V)$  is the average firing rate with supra-linear dependence on  $V$  when  $V$  is small and linear when  $V$  becomes large (Pitkow and Meister, 2012):

$$r(V) = \frac{\log(1 + \exp(\beta_2(V - \theta)))}{\beta_1} \quad (\text{Equation 3})$$

where  $\theta$  (4.7 mV before training) is the firing threshold, and  $\beta_1 = 3.72 \text{ Hz}^{-1}$  and  $\beta_2 = 0.25 \text{ mV}^{-1}$  are two shape parameters.  $I_{time}$  is the input current from the second layer of time cells, defined as

$$I_{time} = \sum_a J_a S_a \quad (\text{Equation 4})$$

with  $S_a = 0, 1$  being the state of the  $a$ th time cell (indicating whether the time cell is inactive or active), and  $J_a$  (0 mV before training) being the synaptic strength from the  $a$ th time cell to the attractor network. The noise  $\zeta$  is modeled as a Gaussian process with autocorrelation  $\langle \zeta(t), \zeta(t') \rangle = \sigma_n^2 \exp(-|t - t'| / \tau_n)$ , with  $\sigma_n = 0.56 \text{ mV}$  and  $\tau_n = 0.58 \text{ s}$ . During training, external stimuli came every  $T = 5 \text{ s}$  or  $10 \text{ s}$  time, lasting for  $1 \text{ s}$ , during which  $I_{stim} = 5 \text{ mV}$ ;  $I_{stim} = 0$  in the absence of external stimuli. The parameter values in our model were chosen so that some parameters (such as  $\tau$ ,  $U$ ,  $\tau_{STD}$  and  $\theta$ ) do not deviate too much from the values suggested in the literature (Holcman and Tsodyks, 2006), and the dynamic behavior of firing rate  $r$  looks similar as observed in our experiment.

As explained in reference (Holcman and Tsodyks, 2006), the dynamics of this model may contain two attractors (up and down states). While the transition between these two states is induced by noise, the duration that the system spends on the two states depends on model parameters. Numerically, we defined *spontaneous up state* in the absence of external stimuli as the following: the up state started when  $r$  rose up across  $r_{up} = 0.8 \text{ Hz}$  and ended when  $r$  fell down across  $r_{down} = 0.2 \text{ Hz}$ . During this period,  $r$  might fluctuate and went across  $r_{up}$  many times. We defined the *time of the up state* as the middle time between the period when  $r$  rose up across  $r_{up}$  for the first time and when its last time fell down across  $r_{up}$ . The duration between the times of two up states after training was what we used to calculate entrained intervals in Figure 7.

The second layer had 2000 time cells, each of which had two states  $S_a = 1, 0$ , representing whether or not it was active. Every  $\Delta t_{time}$  time, we randomly selected 20 time cells, letting their states be 1. Time-cell resetting happened 0.5s after the beginning of an external stimulus or a spontaneous up state. At time-cell resetting,  $\Delta t_{time}$  was randomly chosen from a Gaussian distribution centered at  $\overline{\Delta t_{time}} = 0.33 \text{ s}$  with a standard deviation of 0.15. This randomness of  $\Delta t_{time}$  changed the updating frequency of the activity pattern of the time cells after each resetting, modeling clock-speed variability. At the same time as time-cell resetting, the random number generator for the time-cell selection was reset to its initial state so that the sequence of activity pattern of the time cells started to repeat.

There are three ingredients in the plasticity rule of the synapses from the time cells to the attractor network (Figures 7C–7E). Firstly, the synapses from active time cells are potentiated at the time of a stimulus. Secondly, this potentiation does not happen if there is a spontaneous up state shortly before the external stimulus. Thirdly, a spontaneous up state induces depression eligibility trace in the synapses from the time cells that initiate the spontaneous up state, so that an external stimulus then depresses the synapses according to the eligibility trace.

The first and second terms above were modeled such that when an external stimulus comes, the synapses from the active time cells are potentiated by  $\Delta w_P = 0.15 \text{ mV}$ , until reaching the upper bound  $0.325 \text{ mV}$ . However, this potentiation does not happen if  $r > r_{up}$  at a certain time within 2.5s before the stimulus.

To model the third term above, we supposed that spontaneous up states of the attractor network in the absence of external stimuli induced depression eligibility trace in the synapses from the time cells to the attractor network (Gerstner et al., 2018). This eligibility trace does not change synaptic strength by itself. However, when a stimulus comes, neuromodulators can be released due to the saliency of the stimulus, which depresses the synapses according to the eligibility trace (Gerstner et al., 2018; Redgrave and Gurney, 2006). Specifically, we supposed that in the absence of external stimuli, the eligibility trace  $e_a$  in the synapse from the  $a$ th time cell to the attractor network was updated as

$$\frac{de_a}{dt} = -\eta_d S_a(t - \varepsilon_{pre}) \Theta(r - \theta_{dep}) \quad (\text{Equation 5})$$

where  $s_a(\cdot)$  is the state of the  $a$ th time cell as a function of time and  $\Theta$  is Heaviside step function. Therefore, when the firing rate  $r$  was larger than  $\theta_{dep} = 0.4 \text{ Hz}$ , then negative eligibility trace started to accumulate with a rate  $\eta_d = 0.5 \text{ mV} \cdot \text{Hz}^{-1} \cdot \text{s}^{-1}$  in the synapses from the time cells that were active  $\varepsilon_{pre} = 0.2 \text{ s}$  before the current time. Here we added a small  $\varepsilon_{pre}$  value to represent the causality between the pre- and post-synaptic neuronal activities. When an external stimulus came, the synapses from the time cells to the attractor network were depressed according to the eligibility trace  $J_a \rightarrow J_a + e_a$ , until reaching the lower bound  $0 \text{ mV}$ , then  $e_a$  was reset to 0.

External stimuli also induced intrinsic or synaptic homeostasis in the attractor network (Turrigiano, 2011). In Figures 7F–7I, we only considered intrinsic homeostasis, by supposing that when an external stimulus came, the firing threshold  $\theta$  (see Equation 3) of the attractor network was increased by  $0.3 \text{ mV}$ , until reaching the upper bound  $7.7 \text{ mV}$ .

In Figures 7F–7H, data were collected from 1024 training trials. In each trial, we trained the network using 31 periodic stimuli, and observed the entrained intervals before 400s after the last stimuli.

### Our learning rule vs classic Hebbian rule

Here we further explain the training process illustrated in Figure 7E in the main text. We compare the training process when the synapses from the time cells to the attractor network are changed according to our learning rule with the case when these synapses are changed according to the classic Hebbian rule (Figure S25B). We hope that readers can better understand the merit and the design motivation of our learning rule through this comparison.

We trained the network using the classic Hebbian rule, so that when the attractor network was in an up state, the synapse from the co-firing time cells were potentiated. We found that the entrained intervals were significantly smaller than the period  $T$  of the stimuli under the classic Hebbian rule (Figure S24A). Notably, the distribution of the entrained intervals tended to peak at two places: one at about 2s, and the other at a value slightly smaller than  $T$  (Figure S24A). The training process under the Hebbian rule is illustrated in the left column of Figure S24B. The two peaks in the entrained-interval distribution correspond to the entrained up states initiated by Cells 2 and 4 respectively.

Our learning rule was designed with the motivation to remove these two peaks (Figure S24B, right column). Firstly, the up state initiated by Cell 4 in the fourth row of the right column of Figure S24B tags a depression eligibility trace on the synapse from Cell 4, which is then implemented by the stimulus soon afterward, depressing the synapse from Cell 4. Secondly, the up state initiated by Cell 4 also reduces the firing rate of the attractor network during the stimulus to the intermediate level (see Figure 7D in the main text for experimental evidence) so that the synapse from Cell 2 that is co-active with this stimulus is also tagged by a depression eligibility trace immediately implemented by the stimulus itself. Because the synaptic strength from Cell 2 is already at the baseline level, this synapse effectively remains unchanged. Therefore, after training under our learning rule, the network can generate entrained up states with entrained intervals unbiased to  $T$  (Figure S24B, 5th row, right column).

Why is the depression eligibility trace an important ingredient of our learning rule? Why are the synapses not depressed immediately in response to an intermediate firing rate of the attractor network? The reason is that in that case the entrained-up states will be able to depress the synapses after the removal of the periodic stimuli, which makes the entrained intervals get successively longer (Figure S24C). This contradicts with the observation in our mice experiment: the entrained intervals got successively longer when the period of external stimuli  $T = 5$ s but got successively shorter when  $T = 10$ s, so that the entrained intervals gradually approached to the spontaneous inter-up-state intervals before training in both cases (Figure S10I). Besides, in anesthetic mice, the number of successive periodic entrained up states was more than that in awake mice (Figures 3F and 3I), which could be up to 6 for  $T = 5$ s and 3 for  $T = 10$ s; so it is unlikely that synapses could be immediately depressed by entrained up states.

### Advantages of high plasticity rate

In this section, we will discuss the advantages of a high plasticity rate during learning: fast entrainment, effectiveness and robustness.

- (1) By high plasticity rate, we mean that the change of synaptic weight during plasticity is fast enough: a stimulus-evoked up state potentiates a synapse by a large value  $\Delta w_p$  and a time-cell-initiated or stimulus-evoked up state depresses a synapse by a large value  $\Delta w_d$ . In our model,  $\Delta w_p$  and  $\Delta w_d$  are so large that a synapse gets close to its maximal strength once potentiated and gets close to its minimal strength once depressed. Physiologically, synaptic strength takes binary values with abrupt transition in between (O'Connor et al., 2005), which may be the biological realization of such a high plasticity rate.
- (2) By fast entrainment, we mean that only a few period stimuli are necessary to generate entrained up states. Taking the default parameters (see STAR Methods), our model has a high probability of generating entrained up states after 20 stimuli. The same number of stimuli was also used in our mice experiments.
- (3) By effectiveness, we mean unbiasedness and scalar property. Specifically, we denote the  $i$ th entrained interval of the  $a$ th session to be  $l_{ia}$ . We estimate the interval timing acquired by the network in the  $a$ th session by averaging all the entrained intervals (i.e., calculating  $E_i(l_{ia})$ ). By unbiasedness, we mean that the expectation of  $E_i(l_{ia})$  over sessions (i.e.,  $E_a(E_i(l_{ia}))$ ) is equal to  $T$ ; by scalar property, we mean that the standard deviation of  $E_i(l_{ia})$  over sessions (i.e.,  $\text{Stda}(E_i(l_{ia}))$ ) is proportional to  $T$ .
- (4) By robustness, we mean that the effectiveness of our plasticity rule is insensitive to the parameter values in the model.

In the rest of this section, we will present more simulation results of our biologically plausible model introduced in the main text and explain the effectiveness of the biologically plausible model by analytically studying a simplified model. Our simulation also demonstrated the robustness in high plasticity rate during learning.

The synaptic potentiation and depression during learning can be quantified by  $\Delta w_p$  and  $\overline{\Delta t_{time}}\eta_d$  respectively.  $\Delta w_p$  is the synaptic potentiation caused by a stimulus;  $\eta_d$  is the accumulation rate of the depressive eligibility trace; and  $\overline{\Delta t_{time}}$  is the average duration of the activity of a time cell. We simulated our model using the following two sets of parameters:

- (1) Keep  $\Delta w_p - \overline{\Delta t_{time}}\eta_d = 0$  and change  $\Delta w_p$  (so  $\eta_d$  is also changed accordingly). We studied different learning rates (quantified by  $\Delta w_p$  and  $\eta_d$ ) in this way and found that fast learning speed facilitated the generation of entrained up states using fewer stimuli (Figure S25A).
- (2) Keep  $\overline{\Delta t_{time}}\eta_d = A$  and increase  $\Delta w_p$  with the constraint that  $\Delta w_p - \overline{\Delta t_{time}}\eta_d \geq 0$ ; or keep  $\Delta w_p = A$  and increase  $\eta_d$  with the constraint that  $\Delta w_p - \overline{\Delta t_{time}}\eta_d < 0$ . We studied the potentiation-depression imbalance in this way. We denote  $l_{ia}$  to be the duration of the  $i$ th entrained interval in the  $a$ th training session. We quantify the mean of entrained intervals using

$E_a(E_i(I_{ia}))$ , and quantify the scalar property using  $z_{scalar}$  defined below:

$$z_{scalar} = \frac{|\mu_{ratio, 5s} / \sigma_{ratio, 5s} - \mu_{ratio, 10s} / \sigma_{ratio, 10s}|}{\sqrt{(\mu_{ratio, 5s} \mu_{ratio, 10s}) / (\sigma_{ratio, 5s} \sigma_{ratio, 10s})}} \quad (\text{Equation 6})$$

where  $\mu_{ratio, 5s} = E_a(E_i(I_{ia}))$  and  $\sigma_{ratio, 5s} = \text{Std}_a(E_i(I_{ia}))$  when  $T = 5s$  and  $\mu_{ratio, 10s}$  and  $\sigma_{ratio, 10s}$  are the corresponding values when  $T = 10s$ .  $z_{scalar}$  is the relative error between  $\mu_{ratio, 5s} / \sigma_{ratio, 5s}$  and  $\mu_{ratio, 10s} / \sigma_{ratio, 10s}$ . Good scalar property is indicated by small  $z_{scalar}$ . We found that with a high plasticity rate (i.e., large  $A$  value),  $E_a(E_i(I_{ia}))$  stayed close to  $T$  and  $z_{scalar}$  kept below 0.12 in a large range of  $\Delta wp - \Delta t_{time} \eta_d$  values, suggesting good unbiasedness and scalar property (Figures S25B and S25C). When the plasticity rate was small (i.e., small  $A$  value), both the mean and  $z_{scalar}$  became sensitive to the change of  $\Delta wp - \Delta t_{time} \eta_d$ : in other words, the parameter range that generated good unbiasedness and scalar property shrunk (Figures S25B and S25C).

In conclusion, a high plasticity rate facilitates the generation of entrainment using fewer stimuli and improves the robustness of the unbiasedness and scalar property of the entrained intervals with the change of potentiation-depression imbalance.

To understand the effectiveness of our plasticity rule under a high plasticity rate, we developed a simple analytically tractable model to better manifest the learning process. For the simplicity of the following discussion, we introduce the concept of *psychological time*. Because of clock-speed variability, the time when a time cell becomes active is variable after time-cell resetting (Figure 7B in the main text). Suppose on average, a specific time cell is active at time  $t$  after resetting, then we call it the psychological time of the animal when this time cell is active at time  $t$ , and denote this time cell as  $S_t$ . In other words, psychological time represents the timing preference of time cells, and measures the subjective timing sense of the animal.

Suppose  $w^t$  is the synaptic weight from  $S_t$  to the attractor network, and we consider an array of them  $\{w^t\}_t$ , with the psychological time  $t = \lceil \frac{(N+1)\Delta t_{time}}{2} + T \rceil + \Delta t_{time}$ ,  $\lceil \frac{(N+1)\Delta t_{time}}{2} + T \rceil + 2\Delta t_{time}$ ,  $\lceil \frac{(N+1)\Delta t_{time}}{2} + T \rceil + 3\Delta t_{time}$ ,  $\dots$ ,  $\lceil \frac{(N+1)\Delta t_{time}}{2} + T \rceil + N\Delta t_{time}$ , so the mean psychological time of  $\{w^t\}_t$  is  $T$  and the difference between the psychological time of adjacent weights is  $\Delta t_{time}$ . For simplicity, we denote  $w_i$  ( $i = 1, 2, \dots, N$ ) to be the weight with psychological time  $\text{Psy}(i) = \lceil \frac{(N+1)\Delta t_{time}}{2} + T \rceil + i\Delta t_{time}$ .

Training of this simple model is performed in discrete time. Each time step represents the synaptic change induced by a stimulus. At each time step, we perform three operations, listed below:

- (1) We randomly choose one weight  $w_\tau$  from the set of weights  $\{w_i\}_i$ . Because of clock-speed variability, not the same time cell fires at the times of different stimuli. If  $w_\tau$  is chosen, it means that the time cell  $S_{\text{Psy}(\tau)}$  fires at the time of the stimulus studied in this time step.
- (2) If all  $w_i$ s with  $i < \tau$  cannot induce up state of the attractor network,  $w_\tau$  will be potentiated by  $\Delta w_p$ . This means that the synapse from the time cell co-active with the stimulus will be potentiated if there is no spontaneous up state before the stimulus (see dashed boxes  $B_2$  and  $B_3$  in Figure 7E of the main text or the 1st-3rd rows of the right column of Figure S24B for this case).
- (3) If some  $w_i$  with  $i < \tau$  induces an up state,  $w_i$  will be depressed by  $\Delta w_d$  and no synapse is potentiated by this stimulus. This means that a spontaneous up state before the stimulus induces synaptic depression, see dashed box  $B_4$  in Figure 7E of the main text or the 4th row of the right column of Figure S24B for this case.

In this simplified model,  $w_i \in [w_{min}, w_{max}]$ , and  $w_{max}$  is large enough so that an up state will surely be induced by an active time cell whose synaptic weight is  $w_{max}$ , whereas  $w_{min}$  is small enough so that no up state will be induced by an active time cell whose synaptic weight is  $w_{min}$ . In the fast plasticity limit,  $\Delta w_p = w_{max} - w_{min}$  and  $\Delta w_d = w_{max} - w_{min}$ , which means that the change of synaptic strength at each time step is so large that  $w_i$  gets its maximal and minimal possible value once potentiated or depressed.

We denote the distribution of the weight index chosen in Step 1 to be  $p_{time}(\tau)$ . Here we make the following claim that is the foundation to understand the effectiveness of our plasticity rule:

*Claim 1: In the fast plasticity limit, if  $p_{time}(\tau)$  is a uniform distribution, then after a large number of training steps, the psychological times of the entrained intervals have expectation  $T$  and a standard deviation proportional to  $T$ .*

Note that the psychological time of an entrained up state depends on the first weight that takes  $w_{max}$ : if  $w_\tau = w_{max}$  and  $w_i = w_{min}$   $\forall i < \tau$ , then the entrained up state happen at psychological time  $\text{Psy}(\tau)$ ; if  $w_i = w_{min}$  for all, however, then no entrained up state will happen after training. After a large number of training steps, the last weight in the array  $\{w_i\}$  (i.e.,  $w_N$ ) has a probability 1 to be  $w_{max}$  (because it cannot be depressed back to  $w_{min}$  once potentiated). So the case of no entrained up state will not happen after long-term training nor be considered in the following discussion. From now on, we denote the *first-large weight* to be the first weight that takes  $w_{max}$ : in other words, if  $w_\tau = w_{max}$  and  $w_i = w_{min}$   $\forall i < \tau$ , then  $w_\tau$  is the first-large weight.

To prove Claim 1, we first prove a stronger claim:

*Claim 2: If  $p_{time}(\tau)$  is a uniform distribution, then after a large number of training steps, the probability that  $w_\tau$  to be the first-large weight is equal to  $\frac{1}{N}$  for any  $\tau$ .*



To prove Claim 2, we denote  $A_\tau = \{w_i | i \geq \tau\}$  and  $\bar{A}_\tau = \{w_i | i < \tau\}$ . First, let's think about the following training process and consider the first-large weight in  $A_\tau$ . Suppose that  $w_a$  is chosen in a time step. If  $a < \tau$ , no weights in  $A_\tau$  will be updated. If  $a \geq \tau$ , the weights in  $A_\tau$  can be updated only if  $w_i = w_{min}$  for all  $w_i \in \bar{A}_\tau$ . When the latter happens, each weight in  $A_\tau$  has equal probability to be chosen, just like the case when the training is operated on  $A_\tau$ . Therefore, although the weights in  $\bar{A}_\tau$  gate the updating of  $A_\tau$ , they do not influence this updating once it happens. After a sufficient number of training steps, the statistics of  $A_\tau$  should be equal as the case when  $\bar{A}_\tau$  is removed so that the training is directly operated on  $A_\tau$ . Now let us consider  $w_\tau$ , i.e., the first weight of  $A_\tau$ . When  $A_\tau$  is to be updated, if  $w_\tau$  is chosen, it will surely be  $w_{max}$  at the end of this step, so it will be the first-large weight of  $A_\tau$ ; but if  $w_\tau$  is not chosen, then it will surely be  $w_{min}$  at the end of this step, so it will not be the first-large weight of  $A_\tau$ . If  $p_{time}(\tau)$  is a uniform distribution, the probability that  $w_\tau$  is chosen is  $1/|A_\tau|$ , so the probability that  $w_\tau$  is the first-large weight of  $A_\tau$  is also  $\frac{1}{|A_\tau|}$ , with  $|A_\tau|$  being the number of weights in  $A_\tau$ . The probability that  $w_\tau$  is the first-large weight of the whole set  $A_{whole} = \{w_i | i = 1, 2, \dots, N\}$  is  $P_{FirstLarge}(w_\tau | A_{whole}) = [1 - P_{FirstLarge}(w_0 | A_0)] \cdot [1 - P_{FirstLarge}(w_1 | A_1)] \cdots [1 - P_{FirstLarge}(w_{\tau-1} | A_{\tau-1})] \cdot P_{FirstLarge}(w_\tau | A_\tau)$ , with  $P_{FirstLarge}(w_i | A_i)$  being the probability that  $w_i$  is the first-large weight in set  $A_i$ . Therefore,

$$P_{FirstLarge}(w_\tau | A_{whole}) = \left(1 - \frac{1}{N}\right) \left(1 - \frac{1}{N-1}\right) \cdots \left(1 - \frac{1}{|A_\tau - 1|}\right) \frac{1}{|A_\tau|} = \frac{N-1}{N} \frac{N-2}{N-1} \cdots \frac{|A_\tau - 1| - 1}{|A_\tau - 1|} \frac{1}{|A_\tau|} = \frac{1}{N}$$

This proves Claim 2.

Claim 2 means that the psychological times of the entrained up states have equal probability within the set  $\{Psy(i)\} (i = 1, 2, \dots, N)$  after training. The expectation of this uniform distribution is at its middle point  $T$  and its standard deviation is proportional to  $N$ , which is proportional to  $T$  (the reason why  $N \propto T$  is that because of clock-speed variability, the broadness of the distribution of psychological time around a given physical time  $T$  scales with  $T$ ). So Claim 1 immediately follows Claim 2. The unbiasedness and scalar property of our plasticity rule are then proved, after noting that  $E_i(I_{ia})$  (with  $I_{ia}$  being the  $i$ th entrained interval of the  $a$ th session) is the numeric estimation of psychological time of the entrained up states after training.

We numerically tested Claim 2. Consistently, we found that the distribution of first-large weight tended to be uniform after a sufficient number of time steps (Figure S25D). The change of synaptic weights  $\{w_i\}_i$  during training is illustrated in Figure S25E.

The above derivation assumes that the  $p_{time}(\tau)$  in Claims 1 and 2 is a uniform distribution. Physiologically, it is more plausible to assume that  $p_{time}(\tau)$  is not a uniform distribution, but instead a Gaussian distribution centering around  $T$ . We numerically tested Claim 1 when  $p_{time}(\tau)$  was Gaussian distributed around  $T$  (Figures S25F–S25I). We found that the expectation of the psychological time of the entrained intervals stayed very close to  $T$ , with almost constant ratio between the standard deviation and mean when  $N$  changes (Figures S25H and S25I). This suggests that the effectiveness of the plasticity rule is well preserved with biologically more plausible  $p_{time}(\tau)$ .

## QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis was performed in OriginLab and MATLAB. All data sets were tested for normality using Kolmogorov-Smirnov test and homogeneity of variance using F-test. Data with normal distribution were calculated with student's  $t$  test or Two-Way ANOVA with post-hoc Bonferroni's test as indicated. The paired  $t$  test was used for all paired comparisons. The Wilcoxon sign-rank test, was performed to compare the significance of differences in data that were not normally distributed. Statistical parameters including exact value of 'n' and what 'n' represents are reported in the figured legends. In all cases, the criterion for statistical significance was  $p < 0.05$  and error bars represented SEM.