Main Manuscript for

Auditory stimuli extend the temporal window of visual integration by modulating alpha-band oscillations

Mengting Xu[a,b†], Biao Han[a,b†], Qi Chen [a,b*], Lu Shen[a,b*]

[a] Center for Studies of Psychological Application, South China Normal University, Guangzhou 510631, China

[b] School of Psychology, South China Normal University, Guangzhou 510631, China

[*]**Correspondence:**

lu.shen2013@gmail.com (LS)

qi.chen27@gmail.com (QC)

[†] These authors contribute equally to the present work.

Abstract

In multisensory environments, how inputs from different sensory modalities interact to shape perception are not fully understood. In this study, we investigated how auditory stimuli influence the temporal dynamics of visual processing using electroencephalography (EEG). Participants were presented with two consecutive visual flashes, either accompanied by an auditory beep or without, and were asked to report their perception of one or two flashes. Behaviorally, we found that the introduction of auditory input induced a longer temporal window for integration. Alpha frequency analysis further revealed that the presence of auditory stimuli led to poststimulus alpha frequency degradation, positively correlating with the prolonged temporal window, supporting the idea that alpha oscillations represent the temporal window of visual integration. Further exploration of prestimulus alpha oscillations revealed that auditory stimuli diminished the predictive role of prestimulus alpha frequency while enhancing the predictive role of prestimulus alpha phase in shaping perceptual outcomes. A follow-up transcranial alternating current stimulation (tACS) experiment confirmed that alpha oscillations have a causal role in modulating visual perception in the absence of auditory stimuli but not when auditory stimuli were present. To probe the underlying mechanisms, we developed a computational model based on the phase-resetting hypothesis and perceptual cycle theory, which successfully replicated the core findings. These results reveal that auditory input extends the temporal window of visual integration by resetting alpha oscillations in the visual cortex, leading to alpha frequency reduction and an altered perception of visual events. This study advances the understanding of cross-modal interactions and highlights the dynamic, adaptive processes underlying sensory integration.

Keywords

Multisensory perception, fusion illusion, alpha oscillations, phase reset

Introduction

In our daily lives, we are constantly bombarded with information from various senses. Our brains face the challenge of organizing and integrating these signals into coherent perceptual experiences, and this becomes more complex due to the interactions between different sensory modalities (Cooke et al., 2019; Shams et al., 2002; Watkins et al., 2006). Numerous studies consistently highlight the noteworthy influence of sound signals on visual perception, particularly in the temporal domain. For example, the introduction of sound has been observed to influence the perceived duration of visual stimuli or rate of visual events (Gebhard and Mowbray, 1959; Shipley, 1964; Walker and Scott, 1981; Welch et al., 1986). The temporal relationship between visual and auditory stimuli plays a crucial role, either enhancing or degrading visual temporal resolution (Fendrich and Corballis, 2001; Shimojo et al., 2001). These findings intriguingly suggest that auditory stimuli may exert a profound influence on the temporal processing of the visual system. However, the neural mechanisms that underlie this cross-modal influence remain elusive.

Temporal windows, a fundamental attribute of the visual system, refer to time intervals during which discrete stimuli interact to influence perception (Samaha and Romei, 2023). A simple example is the fusion of two brief stimuli into a unified percept when they occur within a certain time range. The rhythmic patterns of alpha oscillations, serve as a potential mechanism for defining this temporal window (Baumgarten et al., 2015; Cecere et al., 2015; Samaha and Postle, 2015). This notion gains support from research demonstrating that the phase and frequency of alpha oscillations shape the temporal integration process. Alpha oscillation phase influences visual signal detectability at the perceptual threshold (Busch and VanRullen, 2010; Busch et al., 2009; Mathewson et al., 2009), the perceived timing of stimuli (Chakravarthi and Vanrullen, 2012; Milton and Pleydell-Pearce, 2016), and the occurrence of perceptual illusions (Gulbinaite et al., 2017; Rohe et al., 2019; Ronconi et al., 2017). Moreover, emerging evidence highlights a correlation between the frequency of alpha oscillations and temporal

integration (Cecere et al., 2015; Cooke et al., 2019; Han et al., 2023; Samaha and Postle, 2015; Shen et al., 2019). Notably, this frequency not only correlates with but also causally influences observers' tendency to integrate signals within (Han et al., 2023; Samaha and Postle, 2015; Shen et al., 2019) and across different sensory modalities (Cecere et al., 2015; Cooke et al., 2019; Keil and Senkowski, 2017). However, it is important to note that some studies have failed to find a significant correlation between alpha oscillations and temporal parsing of visual signals (Buergers and Noppeney, 2022; Gray and Emmanouil, 2020; Ruzzoli et al., 2019). A recent study provide evidence that pre-stimulus alpha frequency does not substantially influence the temporal segmentation of visual signals in visual or audiovisual perception(Buergers and Noppeney, 2022). This highlights the need for further research to confirm the relationship between alpha oscillations and the temporal window within and across the senses.

In the context of audiovisual processing, the presentation of two visual flashes with an auditory stimulus leads to an increased likelihood of perceiving them as a single flash, inducing fusion illusions (Andersen et al., 2004). This observation suggests that cross-modal stimulation has the potential to impact the temporal window of visual processing. Neural studies have revealed significant activation changes in the primary visual cortex (V1) when visual stimuli are presented concurrently with incongruent auditory stimuli, which are closely linked to reports of visual illusions (Watkins et al., 2006). Furthermore, research on animals and humans has demonstrated that auditory stimuli can modulate early visual cortex processing through phase-reset mechanisms (Kayser et al., 2008; Keil et al., 2014; Lakatos et al., 2007; Mercier et al., 2013), particularly in the alpha frequency range. Building upon these findings, one potential hypothesis is that cross-modal sound stimuli may influence the temporal window of visual integration by exerting an influence on alpha oscillations.

To test this hypothesis, we conducted an experiment where observers performed a visual flash discrimination task while their EEG data was recorded. We analyzed the

frequency and phase of alpha oscillations for two experimental conditions: one with only two visual flashes (F2) and the other with two visual flashes accompanied by an auditory beep (F2B1) (Figures 1A and B). Upon comparing different conditions, our investigation revealed that the introduction of auditory stimuli effectively prolonged the temporal window of integration by reducing the instantaneous frequency of alpha oscillations. Notably, in the F2 condition, the instantaneous frequency of alpha oscillations plays a predictive role in perceptual outcomes, while in the F2B1 condition, this predictive role is absent. Conversely, the phase of alpha oscillations exhibits a predictive role in perceptual outcomes in the F2B1 condition, but not in the F2 condition. Further transcranial alternating current stimulation (tACS) experiment confirmed the casual role of alpha frequency in modulating perception specifically in F2 condition. To elucidate the underlying mechanisms, we introduced a sound-induced phase-resetting model. Our simulation results indicate that the introduction of auditory input resets the phase of alpha oscillations, subsequently extending the temporal window through a reduction in alpha frequency. This modification significantly influences the predictive roles of both phase and frequency.

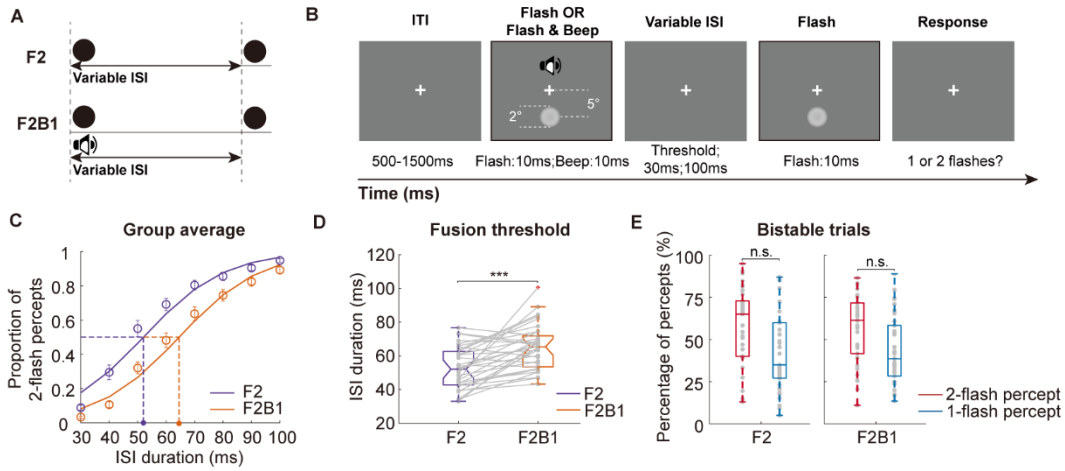

Figure 1. Paradigm and behavioral results. (A) Schematic illustration of the F2 and F2B1 conditions. (B) Illustration of one trial sequence. On each trial, the first visual stimulus presented below the fixation, either accompanied by an auditory beep or without it (the sound icon in the figure is not part of the actual experiment). After a variable ISI, the second visual stimulus was presented. Participants were required to report whether they perceived one or two flashes. Participants were more likely to report perceiving one flash with a short ISI (30 ms) and two flashes with a long ISI (100 ms). Perception became bistable during trials where the ISI matched fusion threshold, individually determined for each participant, resulting in approximately equal proportions of 1-flash and 2-flash reports. (C) Psychometric curves represent the best fit of the average probability of perceiving the two flashes plotted as a function of different ISIs in the F2 and F2B1 conditions. Error bars represent ±1 SEM. (D) Fusion threshold ISIs derived from the psychophysical procedure for each condition. (E) Percentage of 1-flash and 2-flash percepts for the bistable trials in the F2 and F2B1 conditions of the main EEG experiment. ***$p < 0.001$.

Results

Psychophysical Results

Prior to the main EEG experiment, participants underwent a psychophysical pretest to determine the fusion threshold for each condition. Psychometric curves were fitted to the proportion of perceiving two flash percepts at each of the eight ISIs for the F2 and F2B1 conditions, respectively (Figure 1C). The fusion threshold, representing the ISI at which equal proportions of 1-flash and 2-flash trials were reported, was derived for each condition and participant (Figure S1). This threshold provides an estimate of the temporal window for integration within the visual modality (F2) and across the audiovisual modality (F2B1). A two-tailed paired t-test revealed that the fusion threshold was significantly shorter ($t_{(33)} = 5.23$, $p = 9.00\times10^{-6}$; Figure 1D) in the F2 (52.68 ± 12.25 ms, mean ± SD) than in the F2B1 condition (65.18 ± 13.17 ms, mean ± SD), indicating a significant shift in the fusion threshold between the two conditions. Further analysis revealed that the auditory stimulus affected participants' choices differently across ISIs ($F_{(7, 231)} = 11.12$, $p = 4.11 \times 10^{-12}$; Figure S2), with the strongest effect at intermediate ISIs, where the stimulus was most ambiguous. This suggests that the auditory input actively shapes perception rather than simply response bias.

In the main EEG experiment, each condition comprised three types of trials: the explicit 1-flash trials with a short ISI (30 ms), the explicit 2-flash trials with a long ISI (100 ms), and the bistable trials with the fusion threshold ISI obtained from the psychophysical procedure. Participants exhibited high accuracy rates in the explicit trials for both the F2 and F2B1 conditions (F2 with short ISI: 89.16 ± 11.81%; F2 with long ISI: 95.96 ± 4.31%; F2B1 with short ISI: 95.48 ± 5.86%; F2B1 with long ISI: 89.52 ± 14.91%, mean ± SD), indicating their ability to accurately discriminate between the explicit percepts in both conditions. Regarding the bistable trials with the threshold ISI, the proportions of 1-flash and 2-flash percepts were comparable in both the F2 ($t_{(33)} = 1.85$, $p = 0.07$, two-tailed) and F2B1 conditions ($t_{(33)} = 1.92$, $p = 0.06$, two-tailed; Figure 1E),

suggesting that participants experienced a similar level of perceptual ambiguity in the two conditions.

Auditory input extended the temporal window for visual integration by decreasing occipital alpha frequency

The EEG data showed a clear peak in the alpha-band power and a posterior scalp distribution of alpha power during the peri-stimulus period (-600 to 600 ms relative to the first stimulus onset) for all the participants in both F2 and F2B1 conditions (Figure 2A). These results guided further analysis by confirming the frequency of interest to 8-13 Hz and the region of interest to the posterior electrodes (Oz, O1, O2, POz, PO1, PO2, PO3, PO4, PO5, PO6).

Building upon our behavioral observations, we noted a significant extension in the temporal window for visual integration when auditory input was introduced. If we assume that alpha frequency serves as an indicator of the temporal window within the visual cortex, it is plausible to hypothesize that alpha frequency in the F2B1 condition might decrease compared to the F2 condition. To test this hypothesis, we performed a direct comparison of poststimulus instantaneous alpha frequency (IAF) between the F2 and F2B1 conditions by collapsing trials across 1-flash and 2-flash precepts while ensuring an equal number of trials in each condition (See Materials and Methods). The instantaneous frequency was calculated as the temporal derivative of the instantaneous phase, which extracted the alpha-band phase angle time series over the occipital electrodes via the Hilbert transform. Consistent with our hypothesis, in the bistable trials, we found that the frequency of F2B1 was significantly lower than that of F2 after the presentation of auditory stimulus (0-280 ms relative to the first stimulus onset, $p = 1.00\times10^{-3}$, cluster-based correction, two-tailed; Figure 2B).

Subsequently, we conducted a between-subject correlation analysis between the poststimulus IAF differences in the F2B1 and F2 conditions and the fusion threshold ISI differences between these conditions, as measured through psychophysics.

Interestingly, we observed a significant correlation between the decrease in alpha frequency and the increase in fusion threshold in the F2B1 condition relative to the F2 condition (Pearson correlation: $r = 0.45$, $p = 7.55 \times 10^{-3}$). Specifically, the greater the reduction in alpha frequency induced by the auditory stimulus, the longer the fusion threshold in the F2B1 condition compared to the F2 condition (Figure 2C). To ensure the robustness of this correlation, we performed a permutation analysis by randomly shuffling the pairing of post-stimulus IAF differences and fusion threshold ISI differences across participants. This procedure, repeated 1,000 times, generated a null distribution of correlation coefficients expected by chance. The observed correlation coefficient remained significant (Figure S3; $p = 2.00 \times 10^{-3}$), further validating the association. These findings suggest that the concurrent presentation of auditory input may lead to a decrease in alpha frequency, thereby extending the temporal window of visual integration.

Given the difference in fusion thresholds between the F2 and F2B1 conditions, which could potentially contribute to the observed poststimulus differences between the two conditions, we conducted an additional analysis to address this concern. Specifically, we examined the IAF difference between the two conditions in trials with explicit short ISI and long ISI, respectively. In the explicit trials, ISIs were consistent across participants and conditions, with the only distinguishing factor between F2B1 and F2 being the presence of auditory input in F2B1. Notably, our results revealed that the alpha frequency was consistently lower in F2B1 compared to F2 for both explicit conditions (explicit short ISI: 0-180 ms relative to the first stimulus onset, $p = 0.02$, cluster-based correction; explicit long ISI: 0-370 ms relative to the first stimulus onset, $p = 1.00\times10^{-3}$, cluster-based correction; Figures 2D and E). This control analysis robustly confirmed that the decrease in alpha frequency is indeed attributable to the auditory input, regardless of the variations in stimulus ISI.

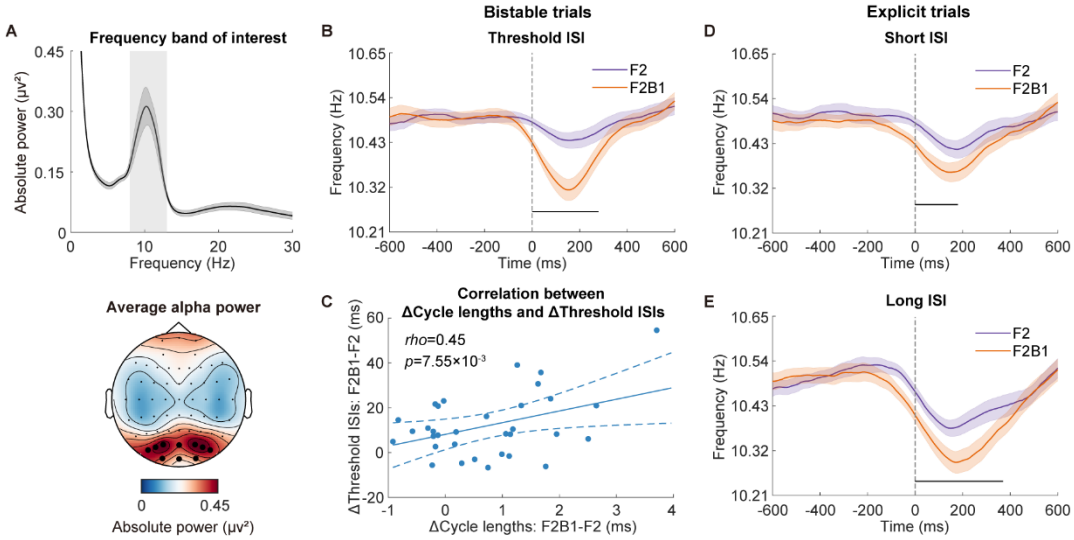

Figure 2. The relationship between decreased post-stimulus IAFs and perceptual precision. (A) *Upper panel*: the power spectrum obtained from all bistable trials (from -600 to 600 ms relative to the onset of the first flash) in F2 and F2B1 conditions, collapsed across all electrodes in all participants, clearly revealed a distinct peak in the alpha band. The gray shading represents ±1 within-subjects SEM. The light gray rectangle represents the chosen alpha frequency band (8-13Hz). *Lower panel*: Topography of absolute alpha-band power (8-13Hz) recorded from 64-channel EEG, averaged over F2 and F2B1 conditions, showed a clear occipital scalp distribution. Black dots indicated the chosen posterior channels. (B) IAFs for F2 and F2B1 bistable trials averaged over the occipital channels. The within-subject findings of the IAFs revealed that the post-stimulus alpha frequency of F2B1 decreased more than that of F2. (C) The between-subject correlation was conducted to examine the relationship between the post-stimulus IAFs difference of F2B1 and F2 in the main EEG experiment and the threshold ISIs difference of F2B1 and F2 in the psychophysical pretest. 95% confidence intervals around the linear fit line are shown by the dashed line. (D) The result is identical to (B), but for explicit short ISI trials. (E) The result is identical to (B), but for explicit long ISI trials.

Auditory input disrupted the predictive role of prestimulus alpha frequency

Extensive research has demonstrated the significant impact of prestimulus alpha frequency on the outcomes of visual integration. In our study, we sought to investigate whether prestimulus alpha frequency exerted a similar influence on the perceptual outcomes in both the F2 and F2B1 conditions. First, we conducted a within-subject comparison of the instantaneous alpha frequency (IAF) variations between different perceptual outcomes of bistable trials in the F2 and F2B1 conditions, respectively. Our results demonstrated that in the F2 condition, the IAF was significantly higher in the 2-flash trials compared to 1-flash trials (Figure 3A), both in the pre-stimulus period (from -330 to -30 ms relative to the first stimulus onset, $p = 0.02$, cluster-based correction, two-tailed) and post-stimulus period (from 130 to 540 ms, $p = 4.00\times10^{-3}$, cluster-based correction, two-tailed). These findings are in line with previous research, indicating that slower alpha frequencies are associated with a wider temporal window, thereby increasing the likelihood that two visual stimuli fall within the same window and are integrated as one percept (Samaha and Postle, 2015). However, in the F2B1 condition, no significant difference in IAF between different perceptual outcomes was observed (Figure 3B).

To further compare the different patterns in perception between the two conditions, we averaged the IAF over the posterior electrodes and the entire time of interest (-600 to 600 ms). These mean IAFs were then subjected to a 2 (percepts: 1-flash vs. 2-flash) × 2 (conditions: F2 vs. F2B1) repeated-measures ANOVA. The results revealed a significant main effect of percepts, $F_{(1,33)} = 4.29$, $p = 0.046$, indicating that the IAF was significantly higher in the 2-flash percepts compared to the 1-flash percepts. Additionally, there was a marginally significant main effect of conditions, $F_{(1,33)} = 3.40$, $p = 0.07$, suggesting a trend towards different IAFs between the F2 and F2B1 conditions. More importantly, the interaction effect was significant, $F_{(1,33)} = 9.86$, $p = 3.55\times10^{-3}$. Specifically, in the F2 condition, the frequency of 2-flash percepts was significantly higher than 1-flash percepts ($t_{(33)} = 5.23$, $p = 3.87\times10^{-3}$), while in the F2B1 condition,

no significant difference between the two percepts was observed ($t_{(33)} = -0.49$, $p = 0.63$; Figure 3C). These results indicate statistically that the influence of IAF on perceptual outcomes differs between the two conditions.

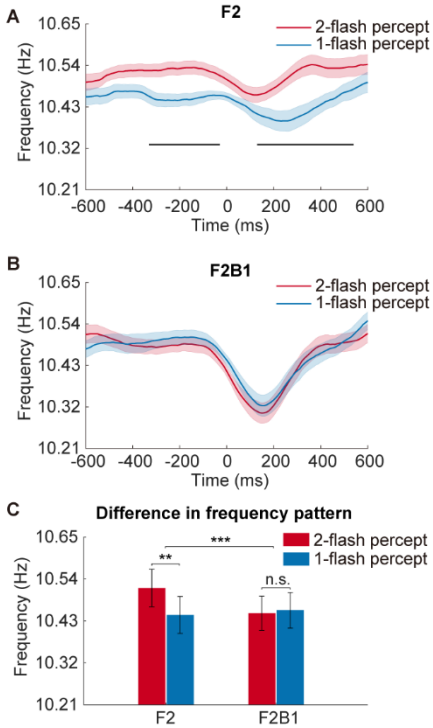

Figure 3. Relationship between the instantaneous alpha frequency (IAF) and perception. (A) IAFs over time were averaged over posterior channels (shown in Fig 2a) for different perceptual outcomes in the F2 condition. Shaded areas show ±1 within-subjects SEM. Significant time epochs are indicated with black lines ($p < 0.05$; permutation test; cluster corrected). (B) Similar to a, but for F2B1 condition. (C) IAFs during the time of interest (-600 to 600 ms) were averaged separately for F2 and F2B1 conditions with different perceptual outcomes. The error bars represent ±1 SEM. $^{**}p < 0.01$. $^{***}p < 0.001$.

Moreover, to address any potential bias in the calculation of instantaneous alpha frequency (IAF) resulting from the 1/f slope effect, we applied a demodulation process with attenuated 1/f characteristics on the time-domain signal (Samaha and Cohen, 2022) and repeated the IAF analysis. The observed pattern of results remained consistent (Figure S4). In addition, to rule out the possibility that the instantaneous alpha frequency modulation is attributed to the oscillatory power, we performed a statistical test on the instantaneous alpha power within the time and channels of interest between the two percepts for F2 and F2B1 conditions, respectively. No significant differences between the two perceptual outcomes were found in either the F2 or F2B1 conditions (Figure S5).

Auditory input enhanced the predictive role of prestimulus alpha phase

The spontaneous alpha-band phase has previously been shown to mirror the cyclic alterations in cortical excitability, thereby supporting the concept of perceptual cycles. In light of this, we extended our investigation to explore the influence of prestimulus alpha phase on perceptual outcomes by computing the phase opposition sum (POS), a measure that quantifies the difference in phase distributions between two sets of trials (VanRullen, 2016a). If the phase of spontaneous alpha oscillations prior to the first flash is predictive of the subsequent perceptual outcome, we should observe a strong phase clustering around a certain phase angle for 1-flash trials and distinct phase clustering around another phase angle for 2-flash trials, resulting in a significant value of POS.

In the F2B1 condition, we indeed observed a significant phase opposition between 1-flash and 2-flash trials, spanning from -460 to 0 ms relative to the first flash onset, within the frequency range of 8 to 13 Hz ($p = 3.03\times10^{-3}$, cluster-based correction; Figure 4A, *upper panel*). However, no significant phase effect manifested in the F2 condition (Figure 4A, *lower panel*). For demonstration purpose, to further visualize the mean phase angle differences across participants, phases at the time-frequency point with maximal POS effect (at 10.28 Hz; -300 ms) at electrode Oz were pooled into 16 bins. In line with the POS analysis, the directions of the preferred phase for 1-flash and

2-flash percepts exhibited significant differences across participants in the F2B1 condition ($F_{(1,66)} = 15.31$, $p = 2.18\times10^{-4}$, Figure 4B, *upper panel*). Conversely, no significant phase difference between the two percepts was detected in the F2 condition ($F_{(1,66)} = 3.30$, $p = 0.07$, Figure 4B, *lower panel*).

Furthermore, we evaluated whether the observed difference in prestimulus POS patterns between F2B1 and F2 was statistically significant using a permutation procedure. Specifically, we computed the mean POS difference between F2B1 and F2 in the frequency range of 8 to 13 Hz and time range from -600 to 0 ms relative to the first flash onset. Subsequently, we compared this mean POS difference value against a distribution of POS difference values expected by chance (see Materials and Methods; Figure 4C). This analysis confirmed a statistically significant distinction in POS patterns between F2B1 and F2 conditions, $p = 4.60\times10^{-3}$, indicating a prevailing phase modulation effect on perceptual outcomes in the F2B1 condition.

The relationship between the perceptual performance and prestimulus phase was further assessed by categorizing single trials according to the phase at the optimal time-frequency point (-300 ms, 10.28 Hz, Oz) and calculating the proportion of 2-flash percepts in each of 9 phase bins (Figure 4D). For each participant, the phase angles were shifted so that the phase corresponding to the highest proportion of 2-flash percepts was aligned to a phase angle of zero. Consequently, due to this artificial alignment, the 2-flash proportion is necessarily maximal at a phase angle of zero; therefore, the zero-phase bin was discarded from further analyses. Most importantly, if the prestimulus phase indeed modulated the perceptual outcomes, we expected that the proportions of 2-flash percepts would peak at the optimal phase angle, gradually decreasing as the phases deviate from this optimal angle. Therefore, we tested for the quadratic main effect for phase bins (8 levels) and their interaction with different conditions (F2B1 vs. F2). There was no significant quadratic effect of phase bins ($F_{(1,33)} = 2.47$, $p = 0.13$). Consistent with the observation that phases exerted a more pronounced influence on behavior when auditory input was introduced, a significant

quadratic interaction emerged between conditions and phase bins ($F_{(1,33)} = 6.45$, $p = 0.02$). Post hoc tests further revealed a significant phase modulation effect only for F2B1 ($F_{(1,33)} = 7.02$, $p = 0.01$), but not for F2 condition ($F < 1$).

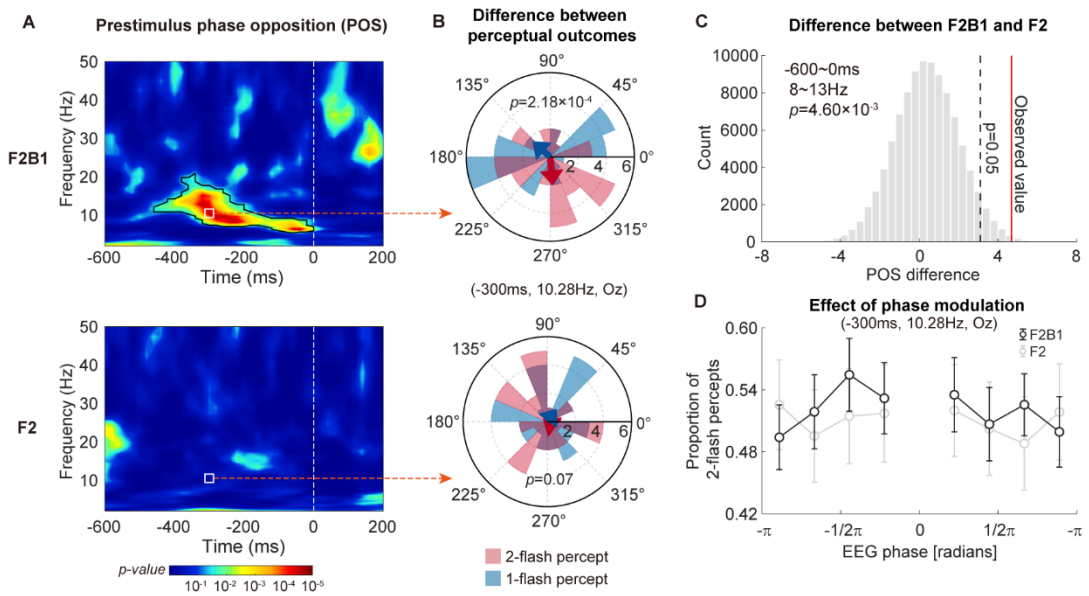

Figure 4. Prestimulus alpha phase difference between 1-flash and 2-flash trials. (A) Time-frequency representation of p values (for occipital regions), computed as a proportion of surrogate phase opposition values (distribution of phase opposition values expected under null hypothesis) that exceeded empirically observed phase opposition values for the F2B1 (*upper panel*) and F2 conditions (*lower panel*), respectively. The outlined area indicates significant effects corrected for multiple comparisons using cluster-based correction. The white square indicates the time-frequency point with strongest POS effect in the F2B1 condition. (B) The circular histograms of mean phase angles (at 10.28 Hz; -300 ms) at the occipital channel Oz during 1-flash and 2-flash trials across participants for the F2B1 (*upper panel*) and F2 conditions (*lower panel*), respectively. The horizontal black line indicates the number of participants with the mean phase angle in each bin. The direction of the arrows corresponds to the mean phase angle across participants, and the length of the arrows indicates the extent of phases clustering around the mean. (C) Statistical analysis of the differential phase effects between F2B1 and F2 condition within the entire prestimulus period (-600~0ms) and alpha-band range (8~13 Hz). The red line indicates the observed mean POS difference between F2B1 and F2 within this cluster. The gray bars represent the surrogate distribution of mean POS difference expected by chance for this cluster, estimated from shuffled data. Dashed black line indicates the 95th percentile value of the surrogate distribution. (D) The relationship between the proportion of two-flash percepts and pre-stimulus phase (at 10.28 Hz; -300 ms) at the occipital channel Oz. Single trials were categorized into 9 phase bins, centered on the preferred phase bin with the highest proportion of 2-flash percepts for each participant (the central bin is removed due to artificial shifting). The error bars represent ±1 SEM.

tACS-modulated alpha frequency causally influences the temporal integration window specifically in the F2 condition

To address the limitations of EEG in providing only correlational evidence, we conducted a follow-up experiment with a new group of participants using transcranial alternating current stimulation (tACS). This approach allowed us to actively modulate the alpha frequency and investigate its causal role in perceptual processes.

During the experiment, continuous tACS was applied to the posterior brain region (Figure S6) while participants performed the same task as in the psychophysical pretest, which required reporting their perception of two flashes across eight ISIs under the F2 and F2B1 conditions. Each participant completed three tACS sessions in a randomized order: low-alpha (8 Hz), high-alpha (13 Hz), and a sham condition with no active stimulation. Psychometric curves were fitted for each tACS session using the proportion of two-flash perceptions across eight ISIs, separately for F2 and F2B1 conditions. From these fitted curves, we calculated the fusion threshold ISI—the ISI at which participants perceived two flashes in 50% of trials—as an estimate of the temporal integration window (Figure 5A and B).

A 3 (tACS sessions: 8 Hz vs. sham vs. 13 Hz) × 2 (stimuli conditions: F2 vs. F2B1) repeated-measures ANOVA on fusion thresholds revealed a significant main effect of conditions, $F_{(1,29)} = 39.58$, $p = 7.19\times10^{-7}$, consistent with the psychophysical pretest, showing that auditory input prolonged the temporal window for visual integration. No significant main effect of tACS sessions was observed, $F_{(1,29)} = 1.25$, $p = 0.27$. Crucially, a significant interaction between tACS session and stimulus condition was observed, $F_{(1,29)} = 5.41$, $p = 0.03$. Specifically, in the F2 condition, the fusion threshold at 13 Hz was significantly lower than at 8 Hz ($t_{(29)} = 2.09$, $p = 0.045$), indicating that higher alpha frequencies shortened the temporal window for perceptual integration. In contrast, no significant differences were observed between different tACS conditions in the F2B1 condition ($F < 1$) (Figure 5C). These findings suggest that alpha oscillations

causally modulate the temporal window for visual perception in the F2 condition but not in the F2B1 condition.

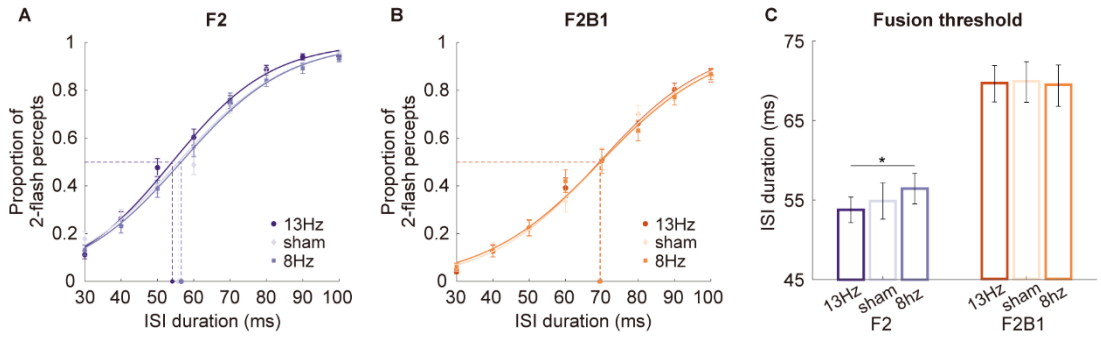

Figure 5. Results of tACS experiment. (A) Psychometric curves show the average probability of perceiving two flashes as a function of different ISIs, under three tACS conditions: 13 Hz, sham, and 8 Hz, in the F2 condition. (B) Similar to (A), but for the F2B1 condition. (C) Fusion threshold ISIs from the three tACS sessions in both F2 and F2B1 conditions. Error bars represent ±1 SEM. *$p < 0.05$.

Auditory input-induced phase resetting model explains the decrease in alpha frequency and the altered predictive roles of alpha frequency and phase

What could be underlying cause of the decrease in poststimulus alpha frequency observed in F2B1 condition? Could this cause explain the divergent effects of the altered predictive roles of alpha frequency and phase? Previous research has shown that transient auditory stimuli can reset the phase of ongoing oscillations in the visual cortex (Lakatos et al., 2009, 2007). This, in turn, has a notable impact on multisensory integration processes(Mercier et al., 2015, 2013). Importantly, this phase resetting has the potential to disrupt the timing and synchronization of periodic oscillations, which could subsequently lead to a shift in oscillation speed.

To test this hypothesis, we introduced a phase-resetting model for examining perceptual processing (Figure 6). Developed based on perceptual cycles theoretical framework (VanRullen, 2016b), this model utilized simulations to replicate our key observations. In the context of perceptual cycle theory, both frequency and phase play fundamental roles in shaping perceptual outcomes. Frequency governs the temporal resolution of the visual system—higher frequencies favor the perception of two distinct flashes, while lower frequencies tend to amalgamate them into a single event. The oscillatory phase represents brain states fluctuating between high and low cortical excitability (Engel et al., 2001; Lakatos et al., 2007), thereby influencing perceptual processing intensity. Our model differentiates percepts by assessing whether two flash stimuli fall within a single cycle. Good phases near the peak enhance processing efficiency, resulting in clearer perceptions of 1-flash and 2-flash percepts. Conversely, bad phases near the trough lead to less efficient processing and increased perceptual ambiguity. Figure 6 visually illustrates our model. When the second flash aligns with a good phase, efficient processing occurs, leading to a clear perception of 1-flash or 2-flash based on whether they fall within the same or different cycles. In contrast, when the second flash aligns with a bad phase, less efficient processing increases perceptual ambiguity, making both interpretations plausible. For the F2B1 condition, the introduction of auditory input,

known to induce delayed responses in the visual cortex relative to visual stimuli (Land et al., 2012; Oeur et al., 2023), induces phase resetting of the ongoing oscillations. The specific parameters governing phase resetting are derived from the model proposed by Canavier in 2013(Canavier et al., 2013), as detailed in the Methods section.

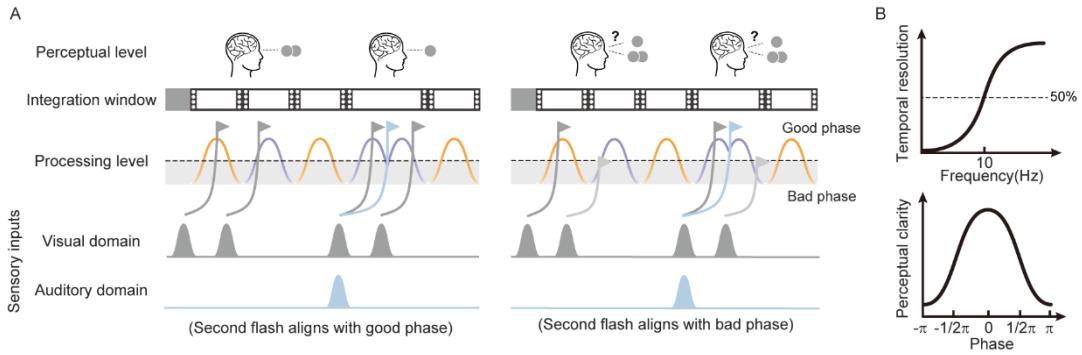

Figure 6. Hypothesized phase-resetting model of perceptual processing. (A) The perceptual outcomes depend on the temporal alignment of two stimuli and the phase at which the second flash occurs. When the second flash coincides with good phases, around the peak of the oscillation, effective processing occurs, and the perceptual outcomes are determined based on whether they fall within the same temporal window defined by the alpha cycle. In cases where the two visual stimuli fall in different temporal windows, the 2-flash percept is reported. The introduction of auditory stimuli concurrently with visual stimuli induces phase resetting, elongating the temporal window. Consequently, the second flash, originally falling in different temporal windows, aligns with the first flash, resulting in a 1-flash percept. Conversely, when the second flash aligns with bad phases, around the trough of the oscillation, effective processing is hindered, leading to ambiguity in 1-flash and 2-flash percepts, irrespective of their temporal alignment. (B) *Upper panel*: The frequency of the alpha rhythm is anticipated to modulate perceptual outcomes, with higher alpha frequencies correlating with an increased proportion of 2-flash percepts. *Lower panel*: The phase of the alpha rhythm is expected to dictate perceptual clarity. In instances where the second flash falls in bad phases (e.g., [-π, -1/2π] and [1/2π, π]), perceptual clarity is low, making it challenging to distinguish between 1-flash and 2-flash percepts. Conversely, when the second flash aligns with good phases (e.g., [-1/2π, 1/2π]), perceptual clarity is high, and the perception of 1-flash or 2-flash depends on whether the two stimuli fall within the same alpha cycle or not.

Building upon the established model, we simulated 1000 trials under two conditions (F2 vs. F2B1). Initial analysis of the model-generated behavioral data yielded results consistent with real participants, notably indicating a prolonged integration window under the F2B1 condition compared to F2 (Figure 7A). Subsequent analysis of alpha oscillations in the simulated data, using the same methodology as for real data, consistently unveiled a lower alpha frequency under F2B1 compared to F2 (Figure 7B). For the alpha frequency effects in perceptual outcomes, mirroring our real data, in our simulations, we averaged the IAF over the entire time of interest (artificially defined from -600 to 600 ms). The simulation results highlighted a more pronounced modulation of alpha frequency in perceptual outcomes in the F2 compared to the F2B1 condition (Figure 7C).

Furthermore, in line with phase results in our real data, simulated data demonstrated more pronounced phase effects in the F2B1 condition compared to the F2 condition at the onset of flash 1 (Figure 7D). Since the phase at which the second flash occurs determines perceptual clarity between the two percepts in our model, we further presented the phase distribution of the second flash for the F2 and F2B1 conditions, respectively (Figure 7E). Notably, auditory input-induced phase resetting made the second flash more likely to fall into the range of a favorable phase (Figure 7E), leading to higher subjective sensitivity (Figure 7F), compared to the F2 condition.

Our own findings lend credence to this hypothesis. We observed a higher alpha inter-trial coherence (ITC) in the F2B1 condition compared to the F2 condition (Figures S7A and B). Because an increase in ITC following a stimulus may reflect a stimulus-evoked response, we also computed changes in alpha power after stimulus onset. A stimulus-evoked response, characterized by an additional signal superimposed on ongoing oscillations, would manifest as both an increase in ITC and power following the stimulus, whereas a pure phase-resetting of ongoing oscillations will manifest as an increase in ITC with no accompanying power change or a power decrease (Makeig et al., 2004; Shah et al., 2004). Our findings showed no significant difference in alpha

power between the two conditions (Figure S7C). Furthermore, we observed no substantial correlation between the magnitude of alpha power and alpha ITC increase after stimulus onset (Figure S7D). Therefore, although we cannot rule out a stimulus-evoked explanation for alpha ITC, we believe the increase in alpha ITC at least partially reflects phase-resetting of spontaneous oscillations in occipital regions.

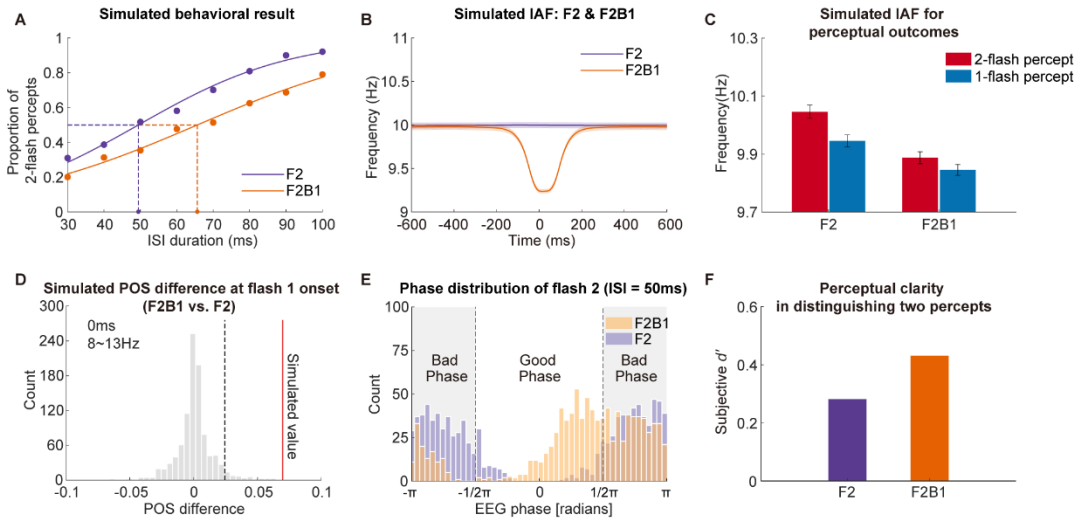

Figure 7. Simulation results. (A) Psychophysical results. Psychometric curves depict the best-fit average probability of perceiving two flashes plotted against different ISIs in simulated F2 and F2B1 conditions. (B) IAFs for simulated F2 and F2B1 trials. (C) IAFs, averaged during the time of interest (-600 to 600 ms), separately for simulated F2 and F2B1 trials with different perceptual outcomes. Error bars represent ±1 within-trials SEM. (D) Statistical analysis of the differential phase effects between F2B1 and F2 condition at stimulus onset (0 ms), prior to the auditory input-induced phase-resetting. The red line indicates the mean POS difference between simulated F2B1 and F2 trials. Dashed black line indicates the 95th percentile value of the surrogate distribution. (E) Phase distribution at the onset of flash 2 (ISI = 50 ms; following the phase-resetting effect). (F) Subjective perceptual sensitivity in distinguishing between 1-flash and 2-flash percepts for simulated F2 and F2B1 trials.

Discussion

In this study, we explored how auditory stimuli influence visual processing, specifically focusing on the temporal window of integration. Behaviorally, we uncovered a prolonged fusion threshold in the presence of auditory stimuli, indicative of an extended temporal window for visual perception induced by cross-modal interactions (Figures 1C and D). The analysis of alpha oscillations provided valuable insights, demonstrating that the introduction of auditory stimuli resulted in poststimulus alpha frequency degradation, which correlated with the extended window, affirming the role of alpha oscillations in shaping the temporal dynamics of visual processing (Figure 2C). Notably, our findings revealed distinct contributions from prestimulus alpha frequency and phase in predicting perceptual outcomes in unimodal and cross-modal conditions (Figures 3 and 4). To elucidate the underlying mechanisms, we employed a computational model based on the phase resetting hypothesis and perceptual cycle theory, successfully replicating key behavioral and neural findings (Figure 7). The convergence of behavioral, neurophysiological, and computational evidence underscores the interplay between auditory and visual modalities in shaping the temporal dynamics of perception.

Our findings reveal that concurrent auditory input significantly extends the temporal window of visual integration compared to the no-sound condition (Figure 1C). This is in line with consistent findings from previous studies demonstrating the impact of sound signals on the perceived duration or rate of visual stimuli (Gebhard and Mowbray, 1959; Shipley, 1964; Walker and Scott, 1981; Welch et al., 1986). Intriguingly, this prolonged temporal window is positively correlated with the decrease in poststimulus alpha frequency induced by auditory stimuli. These findings contribute to the enduring hypothesis associating the human alpha rhythm with temporal windows, positing that higher frequencies correspond to shorter windows and heightened temporal resolution (Cecere et al., 2015; Cooke et al., 2019; Keil and Senkowski, 2017; Samaha and Postle, 2015). A lengthened integration window, as indicated by a longer alpha cycle, implies that two successive visual flashes are more prone to falling within a single cycle.

Consequently, the visual system requires an extended duration to effectively differentiate between the two flashes. This established connection between altered alpha frequency and the resultant temporal window expansion provides a nuanced understanding of how auditory stimulation influences the temporal dynamics of visual perception. Essentially, our findings suggest that auditory stimuli can influence the temporal resolution of visual processing, potentially providing a longer temporal window for the integration of cross-modal information.

Building upon these findings, a critical inquiry arises: how does auditory stimulation contribute to a decrease in alpha frequency? Previous fMRI studies on multisensory perception have shown that auditory stimuli could modulate activity in primary visual cortex(Watkins et al., 2007, 2006). Neurophysiological evidence further emphasizes that transient auditory stimuli can reset the phase of the ongoing oscillations, particularly in alpha-band, within the visual cortex (Lakatos et al., 2009; Mercier et al., 2013; Romei et al., 2012). These insights suggest that cross-modal auditory stimuli might exert influence on alpha oscillations by resetting the phase in the visual cortex. Our empirical evidence, as indicated by the increased phase coherence in the F2B1 condition according to the ITC analysis, lends support to this hypothesis (Figures S7A and B). To unravel the underlying mechanisms, we developed a computational model that integrates the perceptual cycle theory and auditory-induced phase resetting. As expected, the introduction of auditory-induced phase resetting, causing a delay in the originally achieved phase, led to a prolonged temporal integration window and a decrease in post-stimulus alpha frequency. This implies that cross-modal auditory stimuli may disrupt the original trajectory of periodic signals, impacting the integration of incoming inputs in the visual cortex (Luo et al., 2010; Romei et al., 2012; Wutz et al., 2014). Together, our combined empirical and computational evidence provides a comprehensive understanding of how auditory stimuli influence the temporal dynamics of visual perception by altering alpha frequency and temporal integration windows in the visual cortex.

Moreover, our findings highlight a clear distinction in the predictive role of prestimulus alpha frequency between F2 and F2B1 conditions. Consistent with previous studies (Han et al., 2023; Samaha and Postle, 2015; Shen et al., 2019; Wutz et al., 2018), we observed that higher alpha frequencies in the F2 condition correlate with a shorter integration window and more rapid sampling, enabling accurate discrimination of temporally close stimuli. In contrast, lower frequencies correspond to a longer integration window and slower sampling, leading to fusion illusions (Figure 3A). These results align with prior research using similar individual alpha frequency (IAF) analyses (Cohen, 2014a; Samaha and Postle, 2015; Shen et al., 2019; Wutz et al., 2018), which also reported relatively small changes in alpha frequency across conditions (Figure 3A). A possible explanation for this small effect lies in the fact that EEG alpha frequency signals represent an average of all neural populations, including both task-relevant and irrelevant groups. This averaging process likely attenuates the effect size due to noise introduced by task-irrelevant neural activity.

In contrast, this frequency-perception link was disrupted in the F2B1 condition (Figure 3B), where the auditory stimulus appeared to negate the predictive power of alpha frequency in visual integration. This observation aligns with recent null effects reported by Buergers and Noppeney (Buergers and Noppeney, 2022). Further supporting this interpretation, the tACS results demonstrated that alpha frequency modulation influenced perceptual outcomes only in the F2 condition, underscoring the causal role of alpha frequency in visual integration without auditory interference. Moreover, these findings ruled out the possibility that the auditory beep evoked an alpha response contaminating occipital alpha recordings and masking the alpha frequency-perception relationship in the F2B1 condition. Notably, even when alpha frequency was externally manipulated via tACS, no behavioral differences emerged between high- and low-frequency conditions in the F2B1 condition (Figure 5B). This suggests that the auditory-induced disruption of the alpha frequency-perception link was not due to neural noise but rather reflected the fundamental influence of auditory input on alpha oscillatory dynamics. We propose that the auditory stimulus induces phase resetting,

aligning alpha oscillation phases in a consistent direction and slightly reducing poststimulus alpha frequency (Figure 2B). Although subtle, this reduction is meaningful: it increases the probability that stimuli will fall within a longer temporal integration window, thereby favoring 1-flash percepts in the F2B1 condition (Figure 3B). Thus, even if a higher prestimulus alpha frequency might typically predict more 2-flash percepts, the auditory-induced frequency reduction overrides this pattern, diminishing alpha frequency's predictive role. These findings suggest that concurrent auditory input extends the temporal window by modulating the dynamics of alpha oscillations, ultimately changing perceptual integration in a cross-modal context.

Prior research has shown that phase encodes rapidly changing stimuli effectively (Schyns et al., 2011), though temporal precision is often compromised by stimulus-evoked neural responses that dominate phase estimates(Harris, 2023; VanRullen, 2016a). As a result, prestimulus phase has become a valuable proxy for predicting perception, although its effectiveness has been debated. While some studies consistently support the predictive role of prestimulus phase (Busch et al., 2009; Gruber et al., 2014; Mathewson et al., 2009; Milton and Pleydell-Pearce, 2016), others report null effects (Benwell et al., 2022, 2017; Keitel et al., 2022; Ruzzoli et al., 2019; van Diepen et al., 2015). In our study, in the F2B1 condition, the pre-stimulus alpha rhythm exhibited a significant phase opposition effect, whereas this modulation effect was much weaker in the unimodal F2 condition. According to our model, both the frequency and phase of alpha oscillations play distinct roles: frequency determines whether two flashes are perceived as one or two, while phase alignment influences the clarity of that perception. In the F2 condition, flashes falling within the same alpha cycle tend to be perceived as a single flash, while flashes falling in separate cycles favor a 2-flash percept. However, even when both flashes are detected as separate, perception clarity depends on favorable phase alignment: favorable alignment sharpens the distinction, whereas less favorable alignment reduces clarity. In the F2B1 condition, the auditory stimulus induces phase resetting, aligning the second flash within a more favorable phase window, improving perceptual clarity. This mechanism allows the

second flash to reliably fall within optimal phase windows, making it easier to distinguish between one- and two-flash outcomes. Therefore, in the cross-modal F2B1 condition, the predictive role of prestimulus phase is enhanced, as the auditory-driven phase resetting sharpens perceptual distinctions by aligning the stimulus to optimal phases.

Previous studies on the influence of oscillatory phase on perception primarily analyzed absolute phase angles within a cycle, categorizing specific phases as favorable or unfavorable for perception(Busch et al., 2009; Mathewson et al., 2010; Neuling et al., 2012; Ng et al., 2012). This framework, however, creates a potential contradiction: one stimulus could be presented at a favorable phase and another at an unfavorable phase, leading to the erroneous perception of only one stimulus, irrespective of whether they occur in the same or different cycles. Such an explanation conflicts with our behavioral data, which suggest a more nuanced interplay between frequency and phase. Notably, the above-mentioned studies, however, used near-threshold stimuli, raising concerns about their applicability to suprathreshold stimuli, which are presumably perceived with less dependence on specific phase variations (Baumgarten et al., 2015). In contrast, our model integrates the roles of both frequency and phase. Frequency governs the temporal resolution of the visual system, while phase modulates the clarity with which two perceptual states are distinguished (Ten Oever et al., 2020). This expanded role of phase as a modulator of perceptual clarity helps reconcile discrepancies in the literature. For example, Buergers and Noppeney (2022) reported null effects of alpha frequency on sensitivity for discriminating between one-flash and two-flash percepts in both unimodal and cross-modal contexts (Buergers and Noppeney, 2022). Our model suggests that while frequency may not directly influence sensitivity, alpha phase plays a critical role in modulating perceptual sensitivity. By accounting for both frequency and phase, our framework offers a cohesive explanation for both consistent and null results in prior research, advancing our understanding of oscillatory dynamics in perception.

Additionally, the observed phase and frequency effects appear to be unrelated to fluctuations in oscillatory amplitude, as our analyses have revealed. This indicates that, at least in the alpha band, the phase and frequency of ongoing neural oscillations may exert an influence on stimulus processing that is distinct from amplitude variations. This finding aligns with prior evidence (Han et al., 2023; Milton and Pleydell-Pearce, 2016). Moreover, the discernible distinctions in patterns between the phase and frequency of alpha oscillations in cross-modal integration imply that they independently contribute, to some degree, to the modulation of temporal integration processes.

Our findings offer a novel perspective on phase resetting, emphasizing its influence on cross-modal interactions beyond its traditional role in facilitating multisensory integration. For example, prior research has shown that cross-modal phase resetting, such as visual stimuli modulating auditory alpha-band activity, is linked to faster reaction times and enhanced efficiency(Mégevand et al., 2020; Thorne et al., 2011). Similarly, auditory-driven phase resetting in the visual cortex has been demonstrated to facilitate audiovisual integration (Mercier et al., 2013). These studies underscore the critical role of phase resetting in optimizing sensory processing, proposing that its underlying mechanism involves the reorganization of neural oscillations through input from another sensory domain (Senkowski and Engel, 2024). Supporting this, a study in non-human primates by Lakatos et al. (2007) found that somatosensory inputs reset oscillatory phases in the auditory cortex by modulating subthreshold membrane potentials, rather than increasing neural firing rates, thereby enhancing responsiveness (Lakatos et al., 2007). In this study, we reveal a distinct and underexplored consequence of phase resetting. Specifically, we show that auditory inputs to the visual cortex extend the temporal window for visual processing. While this might initially suggest a reduction in the sensory resolution of the visual system, it actually facilitates the seamless integration of auditory and visual inputs. This broader temporal window reflects an adaptive trade-off that enhances cross-modal coherence. Thus, phase resetting not only improves integration efficiency but also dynamically reshapes the

temporal properties of sensory modalities to meet task demands, enabling flexible adaptation to complex environments. This aligns with a recent study in animals, which showed that responses in a primary sensory area's preferred modality (e.g., somatosensory cortex for tactile stimuli) are highly plastic and can adapt to task demands and environmental changes (Kato and Bruno, 2024).

Please note that our model, based on perceptual cycles and phase reset, offers a speculative hypothesis that necessitates additional empirical evidence for validation. Incorporating studies with variable perceptual clarity or sensitivity would enhance our understanding of the specific role played by phase. Moreover, acknowledging the physiological limitations in our EEG study, we recognize the complexity of phase resetting induced by transient events in the brain. Our model simplifies this intricacy by focusing on sound-induced phase resetting, but further empirical evidence is crucial for validation and refinement. Distinguishing evoked phase locking from pure phase resetting poses challenges, and intracranial recordings with higher resolution could provide a potential solution(Bauer et al., 2020). In our upcoming studies, we plan to utilize intracranial recordings in future studies to address and overcome this challenge, providing further clarity on the role of phase in perceptual processing.

In summary, our study unveils the impact of simultaneous auditory input on the temporal perception of two visual stimuli. The findings indicate that the introduction of auditory stimuli extends the temporal window of integration, with a notable influence on alpha oscillations. We propose that this effect is driven by auditory input-induced phase resetting in visual areas, which extends the duration of the alpha cycle. Consequently, this elongation strengthens the dominant predictive role of the alpha phase while simultaneously reducing the influence of alpha frequency in multisensory processing. In combination with previous studies investigating similar paradigms, our current findings lend further support to the perceptual cycle theory and offer a novel interpretation for the divergent outcomes observed in previous research on frequency and phase.

Materials and methods

Participants

Thirty-four healthy right-handed participants (23 females; age 20.8 ± 1.9, Mean ± SD) participated in the EEG experiment. Another group of 30 participants participated in the tACS experiment (18 females; age 22.4 ± 2.2). The sample size was pre-specified to ensure at least 80% power to detect experimental effect sizes ranging from moderate to large (Cohen's d > 0.5). All participants have normal or corrected-to-normal visual acuity and have no history of neurological or psychiatric disease.

Stimuli

The visual stimuli consisted of two identical white circles, each with a radius of 2 degrees, presented against a gray background. These circles were displayed rapidly one after the other, positioned peripherally below the fixation point at an eccentricity of 5 degrees of visual angle. Each visual stimulus had a duration of 10 milliseconds. The auditory stimulus was a 10-millisecond pure tone with a frequency of 3000 Hz and an approximate sound level of 50 decibels. To provide a spatially congruent experience for the participants, the auditory stimuli were delivered through a pair of loudspeakers situated on either side of the monitor. This arrangement created the perception that the auditory cues originated from the center of the screen, closely aligned with the visual stimuli. Visual stimuli were presented using Presentation software (Neurobehavioral Systems; http://www.neurobs.com) on a 24-inch high-refresh-rate LCD monitor (resolution: 1920 × 1080, refresh rate: 100 Hz, viewing distance 57 cm) and viewed through a chin-and-forehead rest. The monitor was carefully calibrated using an oscilloscope before the experiment. The study consisted of two conditions: the F2 condition, involving the presentation of two visual stimuli, and the F2B1 condition, where the first visual flash was consistently accompanied by an auditory beep.

Psychophysical procedures

To determine each subject's fusion threshold for the bistable condition, each participant completed a psychophysical pretest before the main experiment. Prior to the psychophysical pretest, participants completed a practice block consisting solely of explicit short (30 ms) and long trials (100 ms) until accuracy reached at least 90%. During the formal psychophysical pretest, the first flash was presented for a duration of 10 ms, after a variable ISI (eight levels: 30, 40, 50, 60, 70, 80, 90, or 100 ms), the second flash was presented for 10 ms as well. Each ISI comprised 35 trials for both F2 and F2B1 conditions. Participants were then asked to perform a two-alternative forced-choice task in which they had to choose between perceiving one flash or two flashes. A logistic function was used to fit the eight data points (one for each ISI) to a psychometric curve. The fusion threshold ISI, that is, the point at which 1-flash and 2-flash were equally likely to be reported, was calculated by estimating the 50% performance point on the fitted logistic function for each participant. The individual fusion threshold ISI derived from the psychophysical pretest was then utilized as ISI in the bistable trials of the subsequent EEG experiment.

EEG experimental procedures

Participants were instructed to fixate at a central fixation throughout the experiment without moving their eyes. The experimental task was to discriminate one- or two-flash by pressing one of two prespecified buttons on the keyboard using the index and middle fingers of their right hand. The mapping between the two response buttons and the two types of percepts was counterbalanced between participants. Each participant completed 560 experimental trials, including 15% explicit short trials (ISI 30 ms), 15% explicit long trials (ISI 100 ms), and 70% bistable trials (fusion threshold ISI), organized into ten blocks with short breaks in between. Each trial began with a central fixation lasting 500 to 1500 ms, followed by a 10 ms presentation of the first visual stimulus, accompanied by an auditory beep or not. After a variable ISI (30 ms, 100 ms, or the participant's individual fusion threshold), the second frame was displayed for 10 ms, and participants were required to respond within 2000 ms.

Recording and preprocessing of EEG data

The electroencephalogram (EEG) was recorded at a sampling rate of 1000 Hz using a Neuroscan SynAmps2 system equipped with 64-channel Quick-Caps and Ag/AgCl electrodes. To capture horizontal and vertical electrooculograms, four extra electrodes were placed around the participants' eyes. The impedances of all electrodes were diligently maintained below 5 kΩ to ensure high signal quality. The signals were online-referenced to an electrode positioned between Cz and CPz. Subsequently, the data underwent offline processing and analysis using FieldTrip (Oostenveld et al. 2011) (www.fieldtriptoolbox.org) and custom MATLAB scripting. Data were filtered to eliminate interference from 50 Hz power lines, down-sampled to 500 Hz, re-referenced to the average reference, epoch from -1300 to 1000 ms. To correct for any baseline drifts, the mean amplitude was subtracted over a time window of -500 to 0 ms relative to the presentation of the first flash. It should be noted that the subtraction of the mean amplitude has no effect on the pre-stimulus alpha oscillations, as it mainly affects only the zero frequency component (Cohen, 2014a). Before segregating them into different categories, signals containing blinks, muscular movements, and other EEG artifacts were detected and removed through independent component analysis.

tACS experimental procedures

The alternating-current stimulation was administered using a 1 × 1 transcranial electrical stimulator (Soterix Medical, New York, NY) with rubber electrodes embedded in saline-soaked sponges. Electrodes were secured to the scalp with elastic bands. The reference electrode and stimulation electrode were positioned at the vertex (Cz) and posterior occipital area (Oz), respectively, following the international 10-20 EEG system. Each electrode measured 5 × 7 cm^2. A sinusoidal current was used with a DC offset of 0, and impedance was maintained below 5 kΩ. The current intensity was set at 1.5 mA. Participants completed three 16-minute tACS sessions in randomized order: (1) 8 Hz stimulation (low-alpha), (2) 13 Hz stimulation (high-alpha), and (3) sham stimulation. Sessions were spaced 30 minutes apart. During sham stimulation, the

current was ramped up over 30 seconds, then ramped down to zero over the subsequent 30 seconds, after which no further stimulation was applied. The Auto-Sham setting on the tACS controller was used to deliver sham stimulation. This ramping protocol mimicked the somatosensory sensation typically experienced during real transcranial electrical stimulation, particularly during current ramp-up or ramp-down phases. In each session, participants performed the same task as in the psychophysical pre-test, with 20 trials for each ISI condition.

Alpha oscillation analysis of EEG data

In our EEG data analysis, our primary focus was on the bistable trials within both the F2 and F2B1 conditions, unless stated otherwise. Fast Fourier Transform (FFT) was applied to generate a power spectrum for all electrodes and participants, ranging from 1 to 30 Hz. In each trial, the power spectrum was computed over the entire period of interest (-600 to 600 ms, relative to the presentation of the first flash). Subsequently, the grand-level power spectrum was obtained by averaging across all trials, electrodes, and participants, thereby identifying the frequency band where the peak frequency, characterized by the highest power, was located. The power topography map was constructed to depict the most prominent frequency band observed in the power spectrum.

Following established methods (Cohen, 2014a), we computed instantaneous alpha frequencies (IAF) over time. IAF was determined by selecting occipital electrodes (O1, O2, Oz, PO7, PO5, PO3, POz, PO4, PO6, and PO8) exhibiting maximum alpha power in the posterior region of interest for each participant. Specifically, in the time window of -600 to 600 ms, a plateau-shaped band-pass filter (8 to 13 Hz) with a 15% transition zone was applied to the data. Phase angle time series were extracted using the Hilbert transform, and the temporal derivative of the phase angle time series, indicative of instantaneous frequency (scaled by the sampling rate and 2π), was calculated as the phase varied over time. Previous studies have suggested that noise in the phase angle time series may result in sharp, unphysiological responses in the derivative (Cohen,

2014a). Therefore, the following disposition was made on the instantaneous frequencies with a median filter of order 10 and a maximum window size of 400 ms. To be more specific, the data were averaged across trials after being median filtered ten times using 10-time windows ranging from 10 to 400 ms. This analysis is theoretically independent of the oscillatory amplitude, except for cases when the amplitude is zero and the phase is indeterminate, as it only considers changes in instantaneous phase. To account for the possibility that differences in trial numbers may lead to discrepancies across conditions, we counterbalanced the number of trials between perceptual outcomes and different conditions. The average instantaneous alpha frequency (IAF) was then calculated for bistable 1-flash and 2-flash trials in F2 and F2B1 conditions.

For the control analysis of instantaneous alpha power, we employed the same filtering and Hilbert transform as utilized in the IAF analysis to extract instantaneous alpha amplitude. The power was then calculated by squaring the amplitude.

To account for the potential effects of the 1/f slope, we reconstructed the time-domain signals that had been demodulated by implementing 1/f attenuation, following the methods outlined in a previous study (Samaha and Cohen, 2022). Specifically, we utilized the Fast Fourier Transform (FFT) to extract the amplitude and phase spectra from each trial of the original signal. Polynomial fits were then applied to the amplitude spectra to obtain the best-fit values, the inverse of which was multiplied by the amplitude spectrum point-by-point to generate the demodulated amplitude spectrum. To reconstruct the demodulated time-domain signal, we employed the Inverse FFT to combine the phase and amplitude spectra of the original signal into complex values. Next, the demodulated signal underwent a filter with a fixed frequency limit of 8-13 Hz to perform the instantaneous frequency analysis.

Our primary goal was to investigate whether single-trial perception is influenced by the phase of ongoing oscillatory activity. To achieve this, we utilized the phase opposition sum (POS) method (VanRullen, 2016a) to evaluate the phase difference between the

two distinct perceptual outcomes in the F2 and F2B1 conditions. Given that the statistical power of POS is influenced by the absolute trial count within each set of trial, and that reliability of POS can be compromised when dealing with unequal 1-flash and 2-flash trial counts (an inherent feature of all phase-based time-frequency analyses methods) (Vinck et al., 2010), we balanced the trial counts for both perceptual outcomes in both F2 and F2B1 conditions.

Time-frequency transformations were first generated over occipital channels using fieldtrip (Oostenveld et al., 2011). Specifically, we used wavelets with logarithmically spaced cycles (ranging from 2 at the lowest frequency to 7 at the highest frequency) to obtain time-frequency representations (complex numbers) at 60 log-spaced frequency points spanning from 2 to 50 Hz. This yields a complex representation of the amplitude, A, and the phase, φ, for trial j at time t and frequency f:

$$tf_j = A_j(t,f)e^{i\varphi_j(t,f)} \tag{1}$$

The phase of this representation can be extracted by normalizing the complex vector to the unit length:

$$\Phi_j(t,f) = \frac{tf_j(t,f)}{|tf_j(t,f)|} \tag{2}$$

Inter-Trial Phase Coherence (ITC) measures the phase consistency across trials. We calculated the ITC using the method described previously (Cohen, 2014b):

$$ITC_{tf} = \left|\frac{1}{N}\sum_{n}^{N}\Phi_j(t,f)\right| \tag{3}$$

Where N is the number of trials in one group of trials. Note that the ITC is a scalar rather than a vector and it is always non-negative.

Subsequently, to assess whether trials associated with a specific perceptual outcome

(i.e., 1-flash or 2-flash) exhibited differences in phase distribution when compared to trials related to the other perceptual outcome, we computed the POS value using the following formula(VanRullen, 2016a):

$$POS = ITC_A + ITC_B + ITC_{ALL} \tag{4}$$

Where ITC_A and ITC_B are the ITC calculated separately for the two subgroups to be compared (i.e., 1flash or 2flash), and ITC_{ALL} is the ITC calculated across all trials regardless of condition.

Notably, when two conditions exhibit phase clustering at specific, distinct phase angles at a given time and frequency, this signifies that these conditions are phase-clustered with higher ITCs than the total ITC, leading to a higher POS value. Conversely, the POS value will be lower if the different conditions do not display stronger phase clustering or if the phases cluster at approximately the same phase angle.

To rigorously evaluate the significance of phase opposition without assuming any specific probability distribution of the POS values, a non-parametric permutation test was conducted. Initially, the POS values were computed for each point in the time-frequency plane from -600 to 200 ms, ranging from 2 to 50 Hz (in logarithmically spaced cycles) for each occipital electrode and subject, and then averaged over all occipital electrodes and subjects. The surrogate POS values were obtained by randomly assigning trials to one or the other condition for each subject while maintaining a constant number of trials in each condition, and then recalculating the grand mean POS values. The p-value of each time-frequency point was calculated as the proportion of grand mean surrogate POS values that exceeded the empirically observed grand mean POS. To ensure statistical rigor, 100,000 surrogates were employed, and a p-value of 1×10^{-5} was assigned to the points without more extreme POS values in the surrogates.

Statistical analyses

Cluster-based correction was applied when multiple time-frequency points (Figure 4A)

or time points (Figure 2B, 2D-E, 3A and S4A) were tested. For multiple comparison correction of time-frequency points, elements that passed a threshold corresponding to a p-value of 0.05 were marked in each surrogate. Neighboring marked elements were identified as clusters. The count of suprathreshold samples within a cluster was used to define the cluster size, and the largest cluster size was entered into a distribution of cluster size, which was expected under the null hypothesis (Maris and Oostenveld, 2007). The p-value of the cluster was calculated as the proportion of the largest cluster size that exceeded the empirically observed largest cluster size. Regarding multiple comparison correction for time points, paired t-statistics were initially computed between different conditions across participants. Elements that passed a threshold of $p < 0.05$ (two-tailed) were marked to identify temporal clusters. The t-values within a connected cluster were summed up to yield a cluster-level statistic. The p-value of the cluster was determined using a similar shuffling process within each participant.

Simulations of EEG data

Simulations of EEG data were conducted utilizing MATLAB (MathWorks Inc) (for visualization of the simulation procedure, please see Figure S8). The fundamental representation of an oscillator with constant parameters is expressed as:

$$g(t) = \psi\, cos[2\pi f t + \theta] \tag{5}$$

For each trial, signal (y_t) changes over time according to:

$$y_t = y_{t-1} + \Delta g(t-1) + \varepsilon_{t-1} \tag{6}$$

Where ε follows a normal distribution. Aligning with the phase reset hypothesis, which posits that a salient stimulus shifts the phase of ongoing neural oscillation to its optimal excitability phase(Lakatos et al., 2005; Lakatos et al., 2007; Lakatos et al., 2008; Thorne et al., 2011; Harris, 2023), we strategically adjust the stimulus occurrence to synchronize with the good phase range for both experimental conditions. Specifically, we randomly set the initial phase within the good phase range of [-1/2π, 1/2π].

For the F2B1 condition, specifically designed to incorporate sound-induced phase resetting, a phase delay of 25 ms was introduced to the signal after the presentation of a sound stimulus at 0 ms. This delay was randomly varied within the range of 0 to 50 ms.

To simulate behavioral and EEG results, we generated 1000 trials for F2 and F2B1 conditions, respectively, and then implemented eight-level ISIs, ranging evenly from 30 ms to 100 ms. The perceptual outcome depended on the temporal alignment of the two flashes and the phases at which the second flash occurred. Specifically, subjective perceptual outcome was determined based on whether the two flashes occurred within the same alpha cycle or different. The phase at which the second flash occurred played a crucial role in the clarity of differentiating between 1-flash and 2-flash percepts, thereby directly affecting the reported percepts. We assigned a clarity value of 1 for favorable phases [$-1/2\pi$, $1/2\pi$] and a clarity value of 0.5 for unfavorable phases [$-\pi$, $-1/2\pi$] and [$1/2\pi$, π]. Then we calculated the proportion of reported 2-flash percepts for each ISI. Applying a logistic function, we fitted the proportions of 2-flash percepts to determine the fusion threshold for each condition. Subsequently, the thresholds were used as ISI in F2 and F2B1 conditions for further frequency and phase analysis, followed the methodology outlined above for the analysis of real EEG data.

Moreover, we calculated the subjective sensitivity (d') for the F2 and F2B1 conditions, following the signal detection theory:

$$d' = z(HR) - z(FAR) \tag{7}$$

where the hit rate (HR) represents the proportion of cases where the subjective perceptual outcome is 2-flash (i.e. the two flashes fall in the different alpha cycles) and the reported outcome is also 2-flash. The false alarm rate (FAR) is the proportion of cases where the subjective perceptual outcome should be 1-flash (i.e., two flashes fall in the same alpha cycle), but the reported outcome is 2-flash percept.

Additional information

Competing interests

The authors declare no competing interests.

Fundings

This work was supported by grants from the National Natural Science Foundation of China (32000785, 32000741, 31871138, 32071052), the Guangdong Natural Science Foundation (2021A1515011185, 2021A1515011100, 2020A1515110223)

Author contributions

Mengting Xu, Conceptualization, Data curation, Formal analysis, Visualization, Methodology, Writing—original draft, Project administration, Writing—review and editing; Biao Han, Conceptualization, Data curation, Formal analysis, Supervision, Visualization, Methodology, Funding acquisition, Writing—review and editing; Qi Chen, Conceptualization, Supervision, Funding acquisition; Lu Shen, Conceptualization, Data curation, Formal analysis, Supervision, Visualization, Methodology, Funding acquisition, Writing—original draft, Project administration, Writing—review and editing

Author ORCIDs

Mengting Xu https://orcid.org/0009-0001-2865-0748
Biao Han https://orcid.org/0000-0002-6754-2926
Shen Lu https://orcid.org/0000-0002-7948-9341

Data and code availability

The data and code utilized in the study are openly accessible on the website (https://osf.io/2hsqw/). They have been made available for reuse or distribution in the public domain, as per the criteria set forth by the institute and in compliance with the ethical committee's endorsement.

Ethics

Human subjects: The study was authorized by the Ethics Committee of the School of Psychology, South China Normal University, in line with the Declaration of Helsinki, and each participant provided their informed permission.

Reference

Andersen TS, Tiippana K, Sams M. 2004. Factors influencing audiovisual fission and fusion illusions. *Brain Res Cogn Brain Res* **21**:301–308. doi:10.1016/j.cogbrainres.2004.06.004

Bauer A-KR, Debener S, Nobre AC. 2020. Synchronisation of Neural Oscillations and Cross-modal Influences. *Trends Cogn Sci (Regul Ed)* **24**:481–495. doi:10.1016/j.tics.2020.03.003

Baumgarten TJ, Schnitzler A, Lange J. 2015. Beta oscillations define discrete perceptual cycles in the somatosensory domain. *Proc Natl Acad Sci USA* **112**:12187–12192. doi:10.1073/pnas.1501438112

Benwell CSY, Coldea A, Harvey M, Thut G. 2022. Low pre-stimulus EEG alpha power amplifies visual awareness but not visual sensitivity. *Eur J Neurosci* **55**:3125–3140. doi:10.1111/ejn.15166

Benwell CSY, Tagliabue CF, Veniero D, Cecere R, Savazzi S, Thut G. 2017. Prestimulus EEG power predicts conscious awareness but not objective visual performance. *eNeuro* **4**. doi:10.1523/ENEURO.0182-17.2017

Buergers S, Noppeney U. 2022. The role of alpha oscillations in temporal binding within and across the senses. *Nat Hum Behav* **6**:732–742. doi:10.1038/s41562-022-01294-x

Busch NA, Dubois J, VanRullen R. 2009. The phase of ongoing EEG oscillations predicts visual perception. *J Neurosci* **29**:7869–7876. doi:10.1523/JNEUROSCI.0113-09.2009

Busch NA, VanRullen R. 2010. Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proc Natl Acad Sci USA* **107**:16048–16053. doi:10.1073/pnas.1004801107

Canavier CC, Wang S, Chandrasekaran L. 2013. Effect of phase response curve skew on synchronization with and without conduction delays. *Front Neural Circuits* **7**:194. doi:10.3389/fncir.2013.00194

Cecere R, Rees G, Romei V. 2015. Individual differences in alpha frequency drive crossmodal illusory perception. *Curr Biol* **25**:231–235. doi:10.1016/j.cub.2014.11.034

Chakravarthi R, Vanrullen R. 2012. Conscious updating is a rhythmic process. *Proc Natl Acad Sci USA* **109**:10599–10604. doi:10.1073/pnas.1121622109

Cohen MX. 2014a. Fluctuations in oscillation frequency control spike timing and coordinate neural networks. *Journal of Neuroscience* **34**:8988–8998. doi:10.1523/JNEUROSCI.0261-14.2014

Cohen MX. 2014b. Analyzing Neural Time Series Data: Theory and Practice, 1st ed. The MIT Press.

Cooke J, Poch C, Gillmeister H, Costantini M, Romei V. 2019. Oscillatory Properties of Functional Connections Between Sensory Areas Mediate Cross-Modal Illusory Perception. *J Neurosci* **39**:5711–5718. doi:10.1523/JNEUROSCI.3184-18.2019

Engel AK, Fries P, Singer W. 2001. Dynamic predictions: oscillations and synchrony in top-down processing. *Nat Rev Neurosci* **2**:704–716. doi:10.1038/35094565

Fendrich R, Corballis PM. 2001. The temporal cross-capture of audition and vision. *Percept Psychophys* **63**:719–725. doi:10.3758/bf03194432

Gebhard JW, Mowbray GH. 1959. On discriminating the rate of visual flicker and auditory flutter. *Am J Psychol* **72**:521. doi:10.2307/1419493

Gray MJ, Emmanouil TA. 2020. Individual alpha frequency increases during a task but is unchanged by alpha-band flicker. *Psychophysiology* **57**:e13480. doi:10.1111/psyp.13480

Gruber WR, Zauner A, Lechinger J, Schabus M, Kutil R, Klimesch W. 2014. Alpha phase, temporal attention, and the generation of early event related potentials. *Neuroimage* **103**:119–129. doi:10.1016/j.neuroimage.2014.08.055

Gulbinaite R, van Viegen T, Wieling M, Cohen MX, VanRullen R. 2017. Individual alpha peak frequency predicts 10 hz flicker effects on selective attention. *J Neurosci* **37**:10173–10184. doi:10.1523/JNEUROSCI.1163-17.2017

Han B, Zhang Y, Shen L, Mo L, Chen Q. 2023. Task demands modulate pre-stimulus alpha frequency and sensory template during bistable apparent motion perception. *Cereb Cortex* **33**:1679–1692. doi:10.1093/cercor/bhac165

Harris AM. 2023. Phase resets undermine measures of phase-dependent perception. *Trends Cogn Sci (Regul Ed)* **27**:224–226. doi:10.1016/j.tics.2022.12.008

Kato DD, Bruno RM. 2024. Stability of cross-sensory input to primary somatosensory cortex across experience. *Neuron*. doi:10.1016/j.neuron.2024.10.020

Kayser C, Petkov CI, Logothetis NK. 2008. Visual modulation of neurons in auditory cortex. *Cereb Cortex* **18**:1560–1574. doi:10.1093/cercor/bhm187

Keil J, Müller N, Hartmann T, Weisz N. 2014. Prestimulus beta power and phase synchrony influence the sound-induced flash illusion. *Cereb Cortex* **24**:1278–1288. doi:10.1093/cercor/bhs409

Keil J, Senkowski D. 2017. Individual Alpha Frequency Relates to the Sound-Induced Flash Illusion. *Multisens Res* **30**:565–578. doi:10.1163/22134808-00002572

Keitel C, Ruzzoli M, Dugué L, Busch NA, Benwell CSY. 2022. Rhythms in cognition: The evidence revisited. *Eur J Neurosci* **55**:2991–3009. doi:10.1111/ejn.15740

Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE. 2007. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* **53**:279–292. doi:10.1016/j.neuron.2006.12.011

Lakatos P, O'Connell MN, Barczak A, Mills A, Javitt DC, Schroeder CE. 2009. The leading sense: supramodal control of neurophysiological context by attention. *Neuron* **64**:419–430. doi:10.1016/j.neuron.2009.10.014

Land R, Engler G, Kral A, Engel AK. 2012. Auditory evoked bursts in mouse visual cortex during isoflurane anesthesia. *PLoS ONE* **7**:e49855. doi:10.1371/journal.pone.0049855

Luo H, Liu Z, Poeppel D. 2010. Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. *PLoS Biol* **8**:e1000445. doi:10.1371/journal.pbio.1000445

Makeig S, Debener S, Onton J, Delorme A. 2004. Mining event-related brain dynamics. *Trends Cogn Sci (Regul Ed)* **8**:204–210. doi:10.1016/j.tics.2004.03.008

Maris E, Oostenveld R. 2007. Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* **164**:177–190. doi:10.1016/j.jneumeth.2007.03.024

Mathewson KE, Fabiani M, Gratton G, Beck DM, Lleras A. 2010. Rescuing stimuli from invisibility: Inducing a momentary release from visual masking with pre-target entrainment. *Cognition* **115**:186–191. doi:10.1016/j.cognition.2009.11.010

Mathewson KE, Gratton G, Fabiani M, Beck DM, Ro T. 2009. To see or not to see: prestimulus alpha phase predicts visual awareness. *J Neurosci* **29**:2725–2732. doi:10.1523/JNEUROSCI.3963-08.2009

Mégevand P, Mercier MR, Groppe DM, Zion Golumbic E, Mesgarani N, Beauchamp MS, Schroeder CE, Mehta AD. 2020. Crossmodal phase reset and evoked responses provide complementary mechanisms for the influence of visual speech in auditory cortex. *J Neurosci* **40**:8530–8542. doi:10.1523/JNEUROSCI.0555-20.2020

Mercier MR, Foxe JJ, Fiebelkorn IC, Butler JS, Schwartz TH, Molholm S. 2013. Auditory-driven phase reset in visual cortex: human electrocorticography reveals mechanisms of early multisensory integration. *Neuroimage* **79**:19–29. doi:10.1016/j.neuroimage.2013.04.060

Mercier MR, Molholm S, Fiebelkorn IC, Butler JS, Schwartz TH, Foxe JJ. 2015. Neuro-oscillatory phase alignment drives speeded multisensory response times: an electro-corticographic investigation. *J Neurosci* **35**:8546–8557. doi:10.1523/JNEUROSCI.4527-14.2015

Milton A, Pleydell-Pearce CW. 2016. The phase of pre-stimulus alpha oscillations influences the visual perception of stimulus timing. *Neuroimage* **133**:53–61. doi:10.1016/j.neuroimage.2016.02.065

Neuling T, Rach S, Wagner S, Wolters CH, Herrmann CS. 2012. Good vibrations: oscillatory phase shapes perception. *Neuroimage* **63**:771–778. doi:10.1016/j.neuroimage.2012.07.024

Ng BSW, Schroeder T, Kayser C. 2012. A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. *J Neurosci* **32**:12268–12276. doi:10.1523/JNEUROSCI.1877-12.2012

Oeur RA, Palaniswamy M, Ha M, Fernandez-Corazza M, Margulies SS. 2023. Regional variations distinguish auditory from visual evoked potentials in healthy 4 week old piglets. *Physiol Meas* **44**. doi:10.1088/1361-6579/acb4da

Oostenveld R, Fries P, Maris E, Schoffelen J-M. 2011. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* **2011**:156869. doi:10.1155/2011/156869

Rohe T, Ehlis A-C, Noppeney U. 2019. The neural dynamics of hierarchical Bayesian causal inference in multisensory perception. *Nat Commun* **10**:1907. doi:10.1038/s41467-019-09664-2

Romei V, Gross J, Thut G. 2012. Sounds reset rhythms of visual cortex and corresponding human visual perception. *Curr Biol* **22**:807–813. doi:10.1016/j.cub.2012.03.025

Ronconi L, Oosterhof NN, Bonmassar C, Melcher D. 2017. Multiple oscillatory rhythms determine the temporal organization of perception. *Proc Natl Acad Sci USA* **114**:13435–13440. doi:10.1073/pnas.1714522114

Ruzzoli M, Torralba M, Morís Fernández L, Soto-Faraco S. 2019. The relevance of alpha phase in human perception. *Cortex* **120**:249–268. doi:10.1016/j.cortex.2019.05.012

Samaha J, Cohen MX. 2022. Power spectrum slope confounds estimation of instantaneous oscillatory frequency. *Neuroimage* **250**:118929. doi:10.1016/j.neuroimage.2022.118929

Samaha J, Postle BR. 2015. The Speed of Alpha-Band Oscillations Predicts the Temporal Resolution of Visual Perception. *Curr Biol* **25**:2985–2990. doi:10.1016/j.cub.2015.10.007

Schyns PG, Thut G, Gross J. 2011. Cracking the code of oscillatory activity. *PLoS Biol* **9**:e1001064. doi:10.1371/journal.pbio.1001064

Senkowski D, Engel AK. 2024. Multi-timescale neural dynamics for multisensory integration. *Nat Rev Neurosci* **25**:625–642. doi:10.1038/s41583-024-00845-7

Shah AS, Bressler SL, Knuth KH, Ding M, Mehta AD, Ulbert I, Schroeder CE. 2004. Neural dynamics and the fundamental mechanisms of event-related brain potentials. *Cereb Cortex* **14**:476–483. doi:10.1093/cercor/bhh009

Shams L, Kamitani Y, Shimojo S. 2002. Visual illusion induced by sound. *Brain Res Cogn Brain Res* **14**:147–152. doi:10.1016/S0926-6410(02)00069-1

Shen L, Han B, Chen L, Chen Q. 2019. Perceptual inference employs intrinsic alpha frequency to resolve perceptual ambiguity. *PLoS Biol* **17**:e3000025. doi:10.1371/journal.pbio.3000025

Shimojo S, Scheier C, Nijhawan R, Shams L, Kamitani Y, Watanabe K. 2001. Beyond perceptual modality: Auditory effects on visual perception. *Acoust Sci & Tech* **22**:61–67. doi:10.1250/ast.22.61

Shipley T. 1964. Auditory flutter-driving of visual flicker. *Science* **145**:1328–1330. doi:10.1126/science.145.3638.1328

Ten Oever S, Meierdierks T, Duecker F, De Graaf TA, Sack AT. 2020. Phase-Coded Oscillatory Ordering Promotes the Separation of Closely Matched Representations to Optimize Perceptual Discrimination. *iScience* **23**:101282. doi:10.1016/j.isci.2020.101282

Thorne JD, De Vos M, Viola FC, Debener S. 2011. Cross-modal phase reset predicts auditory task performance in humans. *J Neurosci* **31**:3853–3861. doi:10.1523/JNEUROSCI.6176-10.2011

VanRullen R. 2016a. How to Evaluate Phase Differences between Trial Groups in Ongoing Electrophysiological Signals. *Front Neurosci* **10**:426. doi:10.3389/fnins.2016.00426

VanRullen R. 2016b. Perceptual Cycles. *Trends Cogn Sci (Regul Ed)* **20**:723–735. doi:10.1016/j.tics.2016.07.006

van Diepen RM, Cohen MX, Denys D, Mazaheri A. 2015. Attention and temporal expectations modulate power, not phase, of ongoing alpha oscillations. *J Cogn Neurosci* **27**:1573–1586. doi:10.1162/jocn_a_00803

Vinck M, van Wingerden M, Womelsdorf T, Fries P, Pennartz CMA. 2010. The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization. *Neuroimage* **51**:112–122. doi:10.1016/j.neuroimage.2010.01.073

Walker JT, Scott KJ. 1981. Auditory-visual conflicts in the perceived duration of lights, tones and gaps. *J Exp Psychol Hum Percept Perform* **7**:1327–1339. doi:10.1037//0096-1523.7.6.1327

Watkins S, Shams L, Josephs O, Rees G. 2007. Activity in human V1 follows multisensory perception. *Neuroimage* **37**:572–578. doi:10.1016/j.neuroimage.2007.05.027

Watkins S, Shams L, Tanaka S, Haynes JD, Rees G. 2006. Sound alters activity in human V1 in association with illusory visual perception. *Neuroimage* **31**:1247–1256. doi:10.1016/j.neuroimage.2006.01.016

Welch RB, DuttonHurt LD, Warren DH. 1986. Contributions of audition and vision to temporal rate perception. *Percept Psychophys* **39**:294–300. doi:10.3758/bf03204939

Wutz A, Melcher D, Samaha J. 2018. Frequency modulation of neural oscillations according to visual task demands. *Proc Natl Acad Sci USA* **115**:1346–1351. doi:10.1073/pnas.1713318115

Wutz A, Weisz N, Braun C, Melcher D. 2014. Temporal windows in visual processing: “prestimulus brain state” and “poststimulus phase reset” segregate visual transients on different temporal scales. *J Neurosci* **34**:1554–1565. doi:10.1523/JNEUROSCI.3187-13.2014