

## Theta-Rhythmic Behavioral Oscillations Modulate Cross-Modal Stimulus and Response Conflict Processing

**Authors:** Xu Honghui, Xu Yiran, Yang Guochun, Norwich, Liu Xun, Liu Xun

**Date:** 2025-09-26T10:08:26+00:00

### Abstract

Theta oscillations are intimately associated with cognitive control. Previous research has demonstrated that theta oscillations are involved in the processing of cross-modal stimulus conflict and response conflict; however, the relationship between theta oscillations and the magnitude of cross-modal stimulus conflict and response conflict remains unclear. The present study employed an audiovisual Stroop task combined with a high-temporal-resolution behavioral sampling method to investigate this issue. Results indicate that the brain's processing rhythm for task-relevant channel stimuli is modulated by task-irrelevant channel stimuli. When task-irrelevant stimuli are identical to or in conflict with task-relevant stimuli, the brain processes task-relevant stimuli with a theta rhythm; when task-irrelevant stimuli are neutral, the brain processes task-relevant stimuli with an alpha rhythm. Furthermore, the study revealed that in auditory tasks, response conflict amplitude exhibits theta oscillations across audiovisual stimulus onset asynchrony (Stimulus Onset Asynchrony, SOA); in visual tasks, stimulus conflict amplitude exhibits theta oscillations across SOA. These results demonstrate that conflict processing exhibits rhythmicity at the behavioral level, revealing the relationship between theta oscillations and conflict magnitude, and extending the rhythmic theory of attention to the domain of conflict processing in cognitive control.

### Full Text

#### Preamble

#### The Influence of Theta Rhythms in Behavioral Oscillations on Cross-Modal Stimulus Conflict and Response Conflict Processing

Honghui Xu<sup>1,2,3</sup>, Yiran Xu<sup>4</sup>, Guochun Yang<sup>5</sup>, Weizhi Nan<sup>6</sup>, Xun Liu<sup>1,2</sup>

<sup>1</sup>CAS Key Laboratory of Behavioral Science, Institute of Psychology, Beijing 100101, China

<sup>2</sup>Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China

<sup>3</sup>School of Education, South-Central Minzu University, Wuhan 430074, China

<sup>4</sup>School of Art Education, Hubei Vocational College of Art, Wuhan 430079, China

<sup>5</sup>Guangdong Institute of Intelligence Science and Technology, Hengqin, Zhuhai 519031, Guangdong, China

<sup>6</sup>Department of Psychology and Center for Brain and Cognitive Sciences, School of Education, Guangzhou University, Guangzhou 510006, China

## Abstract

Theta oscillations are closely associated with cognitive control. Previous studies have found that theta oscillations participate in the processing of cross-modal stimulus conflict and response conflict. However, the relationship between theta oscillations and the magnitude of cross-modal stimulus and response conflicts remains unclear. The present study employed an audiovisual Stroop task combined with a high temporal resolution behavioral sampling method to investigate this issue. The results demonstrated that the processing rhythm of task-relevant channel stimuli is influenced by task-irrelevant channel stimuli. When task-irrelevant stimuli were congruent or incongruent with task-relevant stimuli, the brain processed task-relevant stimuli at a theta rhythm; when task-irrelevant stimuli were neutral, the brain processed task-relevant stimuli at an alpha rhythm. Additionally, the study found that in the auditory task, response conflict magnitude exhibited theta oscillations across stimulus onset asynchrony (SOA); in the visual task, stimulus conflict magnitude exhibited theta oscillations across SOA. These findings indicate that conflict processing manifests rhythmic patterns at the behavioral level, reveal the relationship between theta oscillations and conflict magnitude, and extend the rhythmic theory of attention to the domain of cognitive control in conflict processing.

**Keywords:** cross-modal, stimulus conflict, response conflict, behavioral oscillation, theta

## 1 Introduction

People process massive amounts of audiovisual information daily. When information from different modalities is inconsistent, it creates competition and induces conflict, thereby interfering with current information processing. Behaviorally, this manifests as prolonged reaction times and increased error rates. How the brain resolves conflict when facing interfering information has long been a central focus in cognitive control research. The Stroop task is a classic experimental paradigm for investigating cognitive control mechanisms (Stroop, 1935). In the color-word Stroop task, participants are required to respond to

the ink color of a word while ignoring its semantic meaning. Results show that when the word's meaning and ink color are incongruent, a conflict effect emerges, with longer reaction times and higher error rates in incongruent compared to congruent conditions. Numerous studies indicate that the Stroop task involves two processing stages: stimulus conflict and response conflict (Augustinova et al., 2018; Burca et al., 2021; De Houwer, 2003; van Veen & Carter, 2005; Zhang & Kornblum, 1998). During stimulus representation, the representation of task-irrelevant information (word meaning) and task-relevant information (ink color) fall within the same semantic category (color). When these two representations differ semantically, stimulus conflict arises. During response selection, the word's meaning and color activate corresponding responses; when these two responses differ, response conflict emerges. Cognitive control enables selective processing of stimuli and responses, allowing individuals to make optimal choices among numerous competing stimuli and responses (Nieuwenhuis & Yeung, 2005).

Theta oscillations (4–8 Hz) play a crucial role in conflict processing, with midfrontal theta oscillations considered closely related to cognitive control functions. Research indicates that midfrontal theta power is sensitive to cognitive control demands and can reflect conflict monitoring and resolution processes (Cohen & Cavanagh, 2011; Cohen & Donner, 2013). Using unimodal visual stimuli, some studies have found that both stimulus conflict and response conflict enhance midfrontal theta power (Haciahmet et al., 2021; Jiang et al., 2015; Nigbur et al., 2012), while other studies have observed modulatory effects of response conflict on theta power only (Haciahmet et al., 2023). Using cross-modal audiovisual stimuli, research has found that in auditory tasks (where participants respond to auditory stimuli while ignoring visual stimuli), only response conflict enhances midfrontal theta power (Xu, Yang, Göschl, et al., 2024); in visual tasks (where participants respond to visual stimuli while ignoring auditory stimuli), only stimulus conflict enhances midfrontal theta power (Xu, Yang, Wu, et al., 2024). Although cross-modal conflict results differ somewhat from unimodal visual conflict findings, these studies consistently demonstrate that midfrontal theta oscillations are closely related to stimulus conflict and response conflict processing. However, EEG studies have primarily examined whether theta oscillations are involved in stimulus and response conflict processing by comparing theta power between conditions, lacking in-depth investigation into the relationship between theta oscillations and behavioral performance. Although some studies have explored correlations between theta oscillations and behavioral performance through correlation analyses (Jiang et al., 2015; Xu, Yang, Göschl, et al., 2024), no study has systematically examined the relationship between theta oscillations and the magnitude of cross-modal stimulus and response conflicts.

Research on behavioral oscillations provides a new perspective for investigating the relationship between neural oscillations and behavioral performance. Behavioral oscillations refer to rhythmic activities of psychological phenomena manifested at the behavioral level (Yang et al., 2019; Zhang et al., 2021), considered to result from rhythmic processing of neuronal ensembles in the brain (Helfrich et al., 2018; Kienitz et al., 2022). To explore the relationship between neural

oscillations and behavioral performance, researchers typically employ cue-target paradigms (Huang et al., 2015; Landau & Fries, 2012). This paradigm resets the phase of internal neural oscillations by presenting a cue stimulus, followed by a target stimulus after a short interval. Because target stimuli presented at different time points fall on different phases of neural oscillations, participants' behavior exhibits periodic fluctuations. By manipulating the stimulus onset asynchrony (SOA) between cue and target stimuli and densely sampling behavioral performance across different SOAs, a time series of behavioral fluctuations can be characterized. Fourier transformation of this time series yields the oscillation frequency of behavioral performance, establishing a relationship between neural oscillations and behavioral performance. Using this method, studies have found that visual and auditory attention processing of spatial locations and features exhibits periodic changes at theta or alpha rhythms in detection performance (Fiebelkorn et al., 2013; Landau & Fries, 2012; Mo et al., 2019; Plöchl et al., 2022; Song et al., 2014). Based on these findings, Fiebelkorn and Kastner (2019) proposed the rhythmic theory of attention. This theory posits that theta oscillations in attention networks periodically modulate functional connectivity between higher brain regions and sensory or motor cortices, enabling rhythmic sampling and shifting of attention to task-relevant locations. Notably, theta rhythms in behavioral oscillations are not limited to attentional sampling of the external environment but also manifest in internal cognitive processing (Huang et al., 2015; Pomper & Ansorge, 2021). Research has shown that perceptual predictions for congruent and incongruent conditions exhibit theta oscillations at the behavioral level (Huang et al., 2015). Particularly, during working memory maintenance, studies have found that hit rates for two visual template representations dynamically change over time at a theta rhythm (Pomper & Ansorge, 2021). Since working memory is a core component of cognitive control (Miller & Cohen, 2001), this result suggests that theta rhythms in behavioral oscillations may not be limited to attention networks but could involve higher-level cognitive control systems. Currently, it remains unclear whether theta rhythms in behavioral oscillations exist in cognitive control conflict processing.

In summary, the present study investigates the relationship between theta oscillations and the magnitude of cross-modal stimulus and response conflicts from the perspective of behavioral oscillations. Previous EEG studies have found that selective attention modality influences the stage at which theta oscillations occur in cross-modal stimulus and response conflict processing (Xu, Yang, Göschl, et al., 2024; Xu, Yang, Wu, et al., 2024). Based on this, we designed two behavioral experiments to examine the relationship between theta oscillations and conflict magnitude in auditory tasks (Experiment 1) and visual tasks (Experiment 2), respectively. Previous studies using behavioral oscillation methods in attention (Mo et al., 2019; Senoussi et al., 2019), perceptual prediction (Huang et al., 2015), and working memory (Pomper & Ansorge, 2021) have found that individual behavioral performance exhibits rhythms identical to neural oscillations, supporting the view that "behavioral oscillations are behavioral manifestations of internal neural oscillations" (Helfrich et al., 2018; Kienitz et al., 2022). Ac-

cordingly, we propose the following hypotheses: At the behavioral level, the stage at which theta oscillations occur should be consistent with that observed in EEG studies. Specifically, in auditory tasks, theta oscillations should appear during cross-modal response conflict processing (Xu, Yang, Göschl, et al., 2024), manifesting as rhythmic changes in cross-modal response conflict magnitude across SOA at a theta frequency. In visual tasks, theta oscillations should appear during cross-modal stimulus conflict processing (Xu, Yang, Wu, et al., 2024), manifesting as rhythmic oscillations in cross-modal stimulus conflict magnitude across SOA at a theta frequency. The present study is innovative in both research perspective and theory. In terms of research perspective, this is the first study to apply the behavioral oscillation method to cognitive control conflict processing. Compared with previous EEG studies that focused primarily on whether theta oscillations are involved in stimulus or response conflict processing (Nigbur et al., 2012; Xu, Yang, Göschl, et al., 2024; Xu, Yang, Wu, et al., 2024), the current study reveals the dynamic relationship between theta oscillations and stimulus and response conflict magnitude by examining dynamic changes in conflict magnitude across SOA. Theoretically, this study provides empirical evidence for the rhythmic theory of attention in cross-modal conflict processing, extending the theory's applicability from attention and perceptual prediction to higher-level cognitive control conflict processing.

## 2 Experiment 1: The Relationship Between Theta Oscillations and Cross-Modal Stimulus and Response Conflict Magnitude in Auditory Tasks

### 2.1.1 Participants

Since behavioral oscillations are considered behavioral manifestations of neural oscillations (Helfrich et al., 2018), to detect theta oscillations behaviorally, we planned to recruit 40 participants based on the sample size used in EEG studies with the same paradigm ( $N = 40$ ) (Xu, Yang, Göschl, et al., 2024). Experiment 1 recruited 43 participants (age:  $23 \pm 2.4$ , 17 males). Three participants were excluded because their accuracy in at least one condition was below 55%, resulting in a valid sample of 40 participants (age:  $23 \pm 2.4$ , 15 males). All participants were healthy, with no psychiatric disorders, normal or corrected-to-normal vision, normal hearing, and right-handed. Participants signed informed consent before the experiment and received compensation upon completion.

### 2.1.3 Experimental Design and Procedure

In Experiment 1, participants responded to auditory stimuli while ignoring visual stimuli (e.g., “hong” and “lan” required pressing the F key, “huang” and “lv” required pressing the J key). Based on the mapping relationship between stimuli and responses, trials were divided into four types: Congruent (CO), where visual and auditory stimuli were identical (e.g., seeing the character “蓝” while hearing “lan”); Stimulus Incongruent (SI), where visual and auditory stimuli

were semantically different but mapped to the same response key (e.g., seeing “红” while hearing “lan” ); Response Incongruent (RI), where visual and auditory stimuli were different and mapped to different response keys (e.g., seeing “绿” while hearing “lan” —the other half of stimuli involved seeing “黄” while hearing “lan” ); and Neutral (NEU), where visual stimuli were “莫” or “结” and auditory stimuli were one of “hong,” “huang,” “lan,” or “lv” (see Figure 1 [Figure 1: see original paper]A). Stimulus conflict magnitude was calculated as the difference between SI and CO conditions, while response conflict magnitude was calculated as the difference between RI and SI conditions. The SOA between audiovisual stimuli had 25 levels, ranging from -200 ms to 200 ms in 16.67 ms steps. Negative values indicated that the task-irrelevant stimulus was presented first, while positive values indicated that the task-relevant stimulus was presented first. This SOA setting served two purposes: First, previous research has shown that distractor stimuli presented within 200 ms before or after the task-relevant stimulus can interfere with processing and generate cross-modal conflict (Donohue et al., 2013). Therefore, SOAs within the -200 ms to 200 ms range could detect conflict. Second, this SOA range (-200 ms to 200 ms) covers a complete theta oscillation cycle (4–7 Hz, period 143–250 ms). Consequently, we used 25 SOAs to densely sample behavioral performance and track changes in conflict magnitude over time.

To familiarize participants with the response mapping rules, they completed three practice blocks of 20 trials each before the formal experiment. Practice Block 1 involved only visual stimuli, requiring participants to respond to the semantic meaning of characters on screen: press “F” for “红” or “蓝,” and “J” for “黄” or “绿.” Practice Block 2 involved only auditory stimuli, requiring participants to respond to heard sounds: press “F” for “hong” or “lan,” and “J” for “huang” or “lv.” Practice Block 3 presented audiovisual stimuli with SOA, and participants were instructed to respond to auditory stimuli while ignoring visual stimuli, using the same response rules as in Blocks 1 and 2. Accuracy in each practice block had to reach 85% before proceeding to the formal experiment. In the formal experiment, each participant completed 17 blocks: the first 16 blocks contained 120 trials each, and the 17th block contained 80 trials. Each SOA condition included 20 trials for each trial type, with trial types equally distributed (congruent:neutral:stimulus incongruent:response incongruent = 1:1:1:1). Each participant completed 2,000 trials.

The experimental procedure is illustrated in Figure 1C. A fixation point was presented for 100–300 ms, followed by visual and auditory stimuli for 450 ms with an SOA between them, and finally a fixation point for 1000–1200 ms. Responses made within 1550 ms after auditory stimulus onset were recorded; trials without responses within 1550 ms were marked as errors (Xu, Yang, Göschl, et al., 2024).

### 2.1.4 Data Analysis

For each participant, practice trials, error trials, reaction times (RTs) beyond three standard deviations, and RTs less than 200 ms were excluded. The mean

number of valid trials per participant for each trial type and SOA was 18 (SD = 0.36). To test whether stimulus and response conflict magnitudes were significant, SOAs were collapsed across conditions, and a one-way repeated measures ANOVA on trial type (congruent, neutral, stimulus incongruent, response incongruent) was conducted.

Next, for RT data, trials were standardized within each participant by calculating Z-scores based on the distribution of all trials within that participant. Standardization removed large between-subject variability due to motor responses (Huang et al., 2015). The NoiseTools toolbox was then used to detrend each trial type using a second-order polynomial method (De Cheveigné & Arzounian, 2018), yielding trend data for each trial type. Each trial's Z-score was then subtracted by the corresponding trend value for that SOA condition to obtain detrended data. Detrending removed classic priming and expectancy effects (Fiebelkorn et al., 2013; Huang et al., 2015; Wang et al., 2020). Subsequent analyses were based on these detrended data.

**Trial Type Spectral Analysis:** An aggregate subject data analysis approach was used for spectral analysis of each trial type's time series. For each trial type and SOA, all participants' trials were pooled to calculate the mean for each SOA. Fourier transformation was then performed on each trial type's 400 ms time series (Liu & Luo, 2019; Tomassini et al., 2017). To test whether fluctuations in amplitude across frequencies were random, we conducted permutation tests. On the detrended data, for each trial type, SOAs were randomly shuffled for each participant. Trials were then pooled across participants by SOA to calculate means and perform Fourier transformation. This procedure was iterated 1,000 times, yielding amplitude distributions for each frequency. The actual observed value's position within the 1,000-iteration distribution was tested for significance. Actual p-values were then FDR-corrected across frequencies.

## 2.2 Results

**Collapsed SOA RT Results:** After collapsing across SOAs, RT results for each trial type are shown in Figure 2A. A one-way repeated measures ANOVA on trial type (congruent, neutral, stimulus incongruent, response incongruent) revealed a significant main effect,  $F(3, 117) = 32.47$ ,  $p < 0.001$ ,  $\eta^2 = 0.45$ . Pairwise comparisons showed that the congruent condition ( $M = 512$  ms,  $SE = 7.95$ ) was significantly faster than the neutral condition ( $M = 523$  ms,  $SE = 8.21$ ),  $t(39) = -5.42$ ,  $p < 0.001$ , Cohen's  $d = -0.18$ , 95% CI = [-15.69, -5.38]; stimulus incongruent condition ( $M = 534$  ms,  $SE = 8.56$ ),  $t(39) = -10.73$ ,  $p < 0.001$ , Cohen's  $d = -0.35$ , 95% CI = [-27.64, -15.58]; and response incongruent condition ( $M = 559$  ms,  $SE = 9.08$ ),  $t(39) = -16.86$ ,  $p < 0.001$ , Cohen's  $d = -0.65$ , 95% CI = [-53.73, -38.56]. The neutral condition was significantly faster than the stimulus incongruent condition,  $t(39) = -5.09$ ,  $p < 0.001$ , Cohen's  $d = -0.18$ , 95% CI = [-16.09, -4.78]; and response incongruent condition,  $t(39) = -13.77$ ,  $p < 0.001$ , Cohen's  $d = -0.34$ , 95% CI = [-41.76, -27.74]. Overall, the collapsed SOA RT results indicated that both stimulus and response conflicts

were significant.

**Trial Type Spectral Analysis Results:** The curves for each trial type across SOAs are shown in Figure 3 Figure 3: see original paper. Raw RT values increased with SOA (Figure 3A). Figure 3B shows detrended values for each trial type fluctuating around zero. To test whether these fluctuations contained theta oscillation information, we performed spectral analysis using both aggregate subject data analysis and single-trial linear modeling methods. Aggregate subject data analysis results are shown in Figure 3C. Significant peak frequencies were 3.8 Hz for congruent and stimulus incongruent conditions ( $p < 0.001$ ), 7.5 Hz for response incongruent condition ( $p < 0.001$ ), and 9.4 Hz for neutral condition ( $p < 0.001$ ). All p-values were FDR-corrected.

Single-trial linear modeling yielded similar results to aggregate subject data analysis (Figure 3D). Group-level Hotelling's T-square statistics revealed significant peak frequencies of 5 Hz for congruent and stimulus incongruent conditions ( $p = 0.004$ ,  $p < 0.001$ , respectively), 7.5 Hz for response incongruent condition ( $p < 0.001$ ), and 7.5 Hz for neutral condition ( $p = 0.038$ ).

**Stimulus and Response Conflict Spectral Analysis Results:** Differences between stimulus incongruent and congruent conditions (stimulus conflict) and between response incongruent and stimulus incongruent conditions (response conflict) after standardization and detrending are shown in Figure 4 [Figure 4: see original paper]A. Spectral analysis revealed that stimulus conflict oscillation frequency was not significant,  $p > 0.05$ ; response conflict showed a significant peak frequency of 3.8 Hz ( $p = 0.048$ , FDR corrected) (see Figure 4B).

## 2.3 Discussion

Experiment 1 used a 2-to-1 audiovisual Stroop task combined with high temporal resolution behavioral sampling to investigate the relationship between theta oscillations and cross-modal stimulus and response conflict magnitude in auditory tasks. Collapsed SOA RT results showed significant stimulus and response conflicts, replicating previous findings (Xu, Yang, Göschl, et al., 2024). In terms of trial types, spectral analysis using both aggregate subject data analysis and single-trial linear modeling revealed that congruent, stimulus incongruent, and response incongruent conditions oscillated at theta (4-7.5 Hz) rhythms, whereas the neutral condition oscillated at an alpha (9.4 Hz) rhythm.

Previous behavioral oscillation research has defined the theta band as 4-8 Hz (Fiebelkorn et al., 2013). Therefore, we defined 7.5 Hz as within the theta frequency range. Critically, spectral analysis of stimulus and response conflicts showed that response conflict magnitude oscillated at theta frequency (3.8 Hz), while no significant oscillation frequency was found for stimulus conflict. This result supports our hypothesis for Experiment 1: in auditory tasks, response conflict magnitude oscillates at theta frequency.

In auditory tasks, only cross-modal response conflict magnitude exhibited theta

rhythmic oscillations. The occurrence of theta oscillations during cross-modal response conflict processing is consistent with previous EEG findings (Xu, Yang, Göschl, et al., 2024). Compared with EEG studies (Xu, Yang, Göschl, et al., 2024), the current study employed high temporal resolution behavioral sampling to further reveal the relationship between theta oscillations and conflict magnitude. Since behavioral oscillations are considered behavioral manifestations of neural oscillations (Kienitz et al., 2022), we speculate that the theta rhythm observed behaviorally may reflect periodic modulation of conflict magnitude by cognitive control-related theta oscillations. The current study demonstrates that cognitive control conflict processing exhibits rhythmic patterns at the behavioral level. If this phenomenon reflects a general processing mode of cognitive control, theta oscillations should also be observable in behavioral performance in visual tasks. In visual tasks, previous EEG research has found theta oscillations involved in cross-modal stimulus conflict processing at the neural level (Xu, Yang, Wu, et al., 2024). However, at the behavioral level, whether cross-modal stimulus conflict magnitude exhibits theta rhythmic oscillation patterns remains unclear. Experiment 2 will investigate the relationship between theta oscillations and cross-modal stimulus conflict magnitude in visual tasks.

### **3 Experiment 2: The Relationship Between Theta Oscillations and Cross-Modal Stimulus and Response Conflict Magnitude in Visual Tasks**

#### **3.1.1 Participants**

The sample size estimation for Experiment 2 was identical to Experiment 1. Based on EEG studies using the same paradigm (Xu, Yang, Wu, et al., 2024), we planned to recruit 40 participants. Experiment 2 recruited 40 participants (age:  $22 \pm 2.4$ , 17 males). All participants were healthy, with no psychiatric disorders, normal or corrected-to-normal vision, normal hearing, and right-handed. Participants signed informed consent before the experiment and received compensation upon completion.

#### **3.1.3 Experimental Design and Procedure**

In Experiment 2, participants responded to visual stimuli while ignoring auditory stimuli. Similar to Experiment 1, trials were divided into four types based on stimulus-response mapping: congruent condition (visual and auditory stimuli congruent), stimulus incongruent condition (visual and auditory stimuli semantically different but mapping to the same response), response incongruent condition (visual and auditory stimuli different and mapping to different responses), and neutral condition (auditory stimuli were “mo” or “jie,” visual stimuli were one of “红,” “黄,” “蓝,” or “绿”) (see Figure 1B). SOA had 25 levels, ranging from -200 ms to 200 ms in 16.67 ms steps. Negative values indicated that the task-irrelevant stimulus (auditory) was presented first, while positive values indicated that the task-relevant stimulus (visual) was presented first. The

experimental procedure was similar to Experiment 1 (see Figure 1C), with the difference that responses made within 1550 ms after visual stimulus onset were recorded; trials without responses within 1550 ms were marked as errors (Xu, Yang, Wu, et al., 2024).

### 3.1.4 Data Analysis

For each participant, practice trials, error trials, RTs beyond three standard deviations, and RTs less than 200 ms were excluded. The mean number of valid trials per participant for each trial type and SOA was 19 (SD = 0.29). The subsequent data analysis procedure was identical to Experiment 1. We first calculated whether stimulus and response conflict magnitudes were significant after collapsing across SOAs, then performed spectral analysis on trial types and conflict magnitudes using standardized and detrended data.

## 3.2 Results

**Collapsed SOA RT Results:** After collapsing across SOAs, RT results for each trial type are shown in Figure 2B. A one-way repeated measures ANOVA on trial type (congruent, neutral, stimulus incongruent, response incongruent) revealed a significant main effect,  $F(3, 117) = 38.85$ ,  $p < 0.001$ ,  $p^2 = 0.50$ . Pairwise comparisons showed that the congruent condition ( $M = 513$  ms,  $SE = 8.27$ ) was significantly faster than the stimulus incongruent condition ( $M = 521$  ms,  $SE = 8.67$ ),  $t(39) = -7.71$ ,  $p < 0.001$ , Cohen's  $d = -0.15$ , 95% CI =  $[-10.79, -5.07]$ , indicating significant stimulus conflict; the stimulus incongruent condition was significantly faster than the response incongruent condition ( $M = 527$  ms,  $SE = 8.93$ ),  $t(39) = -5.32$ ,  $p < 0.001$ , Cohen's  $d = -0.11$ , 95% CI =  $[-9.04, -2.83]$ , indicating significant response conflict. Additionally, the congruent condition was significantly faster than the neutral condition ( $M = 522$  ms,  $SE = 8.22$ ),  $t(39) = -6.78$ ,  $p < 0.001$ , Cohen's  $d = -0.15$ , 95% CI =  $[-11.51, -4.82]$ , and the response incongruent condition,  $t(39) = -8.69$ ,  $p < 0.001$ , Cohen's  $d = -0.26$ , 95% CI =  $[-18.30, -9.43]$ ; the difference between neutral and stimulus incongruent conditions was not significant,  $t(39) = 0.20$ ,  $p = 1.000$ , Cohen's  $d = 0.004$ , 95% CI =  $[-3.03, 3.50]$ ; the neutral condition was significantly faster than the response incongruent condition,  $t(39) = -3.72$ ,  $p = 0.004$ , Cohen's  $d = -0.11$ , 95% CI =  $[-9.96, -1.44]$ . Overall, the collapsed SOA RT results indicated that both stimulus and response conflicts were significant.

**Trial Type Spectral Analysis Results:** As in Experiment 1, raw RT values and standardized, detrended RT values for each trial type across SOAs are shown in Figure 5: see original paper. In Figure 5A, RTs for all trial types gradually increased as the task-irrelevant stimulus was presented first, then remained constant. In Figure 5B, standardized and detrended RT values for each trial type fluctuated around zero. To test whether these fluctuations contained oscillatory information, we performed spectral analysis using both aggregate subject data analysis and single-trial linear modeling methods (Figure 5C-D). Aggregate subject data analysis results (Figure 5C) showed significant

peak frequencies of 5.6 Hz for congruent and response incongruent conditions ( $p < 0.001$ ), 9.4 Hz for neutral condition ( $p < 0.001$ ), but no significant frequency for stimulus incongruent condition,  $p > 0.05$ . All  $p$ -values were FDR-corrected. Single-trial linear modeling results (Figure 5D) showed significant peak frequencies of 5 Hz for congruent ( $p < 0.001$ ), stimulus incongruent ( $p = 0.034$ ), and response incongruent conditions ( $p < 0.001$ ), and 10 Hz for neutral condition ( $p < 0.001$ ).

**Stimulus and Response Conflict Spectral Analysis Results:** Differences between stimulus incongruent and congruent conditions (stimulus conflict) and between response incongruent and stimulus incongruent conditions (response conflict) after standardization and detrending are shown in Figure 4C. Spectral analysis results are presented in Figure 4D. Stimulus conflict showed a significant peak frequency of 5.6 Hz ( $p < 0.001$ , FDR corrected). No significant frequency was found for response conflict,  $p_s > 0.05$ .

### 3.3 Discussion

Experiment 2 investigated the relationship between theta oscillations and cross-modal stimulus and response conflict magnitude in visual tasks. Analysis of collapsed SOA RT data revealed significant cross-modal stimulus and response conflict magnitudes. In terms of behavioral oscillations, consistent with Experiment 1, congruent, stimulus incongruent, and response incongruent conditions oscillated at theta (5–5.6 Hz) rhythms, while the neutral condition oscillated at alpha (9.4–10 Hz) rhythms. Notably, Experiment 2 found that cross-modal stimulus conflict magnitude exhibited theta (5.6 Hz) rhythmic oscillations in visual tasks. The occurrence of theta oscillations during cross-modal stimulus conflict processing is consistent with previous EEG findings in visual tasks (Xu, Yang, Wu, et al., 2024), supporting our hypothesis. Experiments 1 and 2 used identical audiovisual stimuli and both found theta oscillations in cross-modal conflict magnitude, suggesting that cognitive control conflict processing may universally exhibit rhythmic processing patterns.

Previous research has found asymmetry in cross-modal conflict effects, with visual interference on audition being greater than auditory interference on vision (Donohue et al., 2013). The present study found this asymmetry across different conflict processing stages, with both cross-modal stimulus and response conflict magnitudes being significantly larger in auditory than visual tasks (see Figure A1). This differs from Chen and Zhou (2013), who found that visual interference on audition was greater than auditory interference on vision in cross-modal stimulus conflict, but the opposite pattern in cross-modal response conflict. Differences in experimental design may account for these divergent results. Chen and Zhou (2013) used only simultaneous audiovisual presentation (SOA = 0), whereas the current study employed SOAs ranging from -200 ms to 200 ms. We speculate that different SOAs may affect the asymmetry of conflict effects across processing stages. Future research should investigate this issue more deeply.

The theta frequency for response conflict in Experiment 1 (3.8 Hz) was lower than that for stimulus conflict in Experiment 2 (5.6 Hz). Response conflict magnitude was larger than stimulus conflict magnitude (see Appendix), indicating greater difficulty in resolving response conflict and higher cognitive control demands. We speculate that the lower theta frequency in response conflict compared to stimulus conflict may reflect a decrease in theta frequency with increasing cognitive control demands. This speculation aligns with previous findings (Senoussi et al., 2022). Based on computational modeling and previous behavioral and EEG results, Senoussi et al. (2022) proposed that the theta rhythm of cognitive control shifts to an optimal frequency according to task demands. Specifically, when task difficulty is high, lower theta frequencies yield better behavioral performance; when task difficulty is low, higher theta frequencies yield better performance. Thus, the current study supports the view that theta frequency shifts toward lower frequencies as task demands increase.

## 4 General Discussion

The present study used a 2-to-1 audiovisual Stroop task combined with high temporal resolution behavioral sampling to investigate the relationship between theta oscillations and cross-modal stimulus and response conflict magnitude. The results showed that in both auditory and visual tasks, when task-irrelevant stimuli were congruent or incongruent with task-relevant stimuli (congruent, stimulus incongruent, and response incongruent conditions), processing of task-relevant stimuli exhibited theta (4-7.5 Hz) oscillations at the behavioral level; when task-irrelevant stimuli were neutral, processing exhibited alpha (9.4-10 Hz) oscillations. Furthermore, the study found that in auditory tasks, cross-modal response conflict magnitude exhibited theta (3.8 Hz) oscillations across SOA; in visual tasks, cross-modal stimulus conflict magnitude exhibited theta (5.6 Hz) oscillations across SOA. To our knowledge, this is the first study to demonstrate rhythmic patterns in cognitive control conflict processing at the behavioral level. These findings reveal the relationship between theta oscillations and cross-modal stimulus and response conflict magnitude, extending the rhythmic theory of attention and deepening our understanding of the theta oscillation mechanisms underlying cognitive control.

### 4.1 The Rhythmic Processing of Task-Relevant Channel Stimuli Is Modulated by Task-Irrelevant Channel Stimuli

Both auditory and visual tasks in this study found theta oscillations in behavioral processing for congruent, stimulus incongruent, and response incongruent conditions. This result is consistent with previous behavioral oscillation research on perceptual prediction. Huang et al. (2015) used a visual priming paradigm and found that RTs for congruent and incongruent conditions exhibited theta rhythmic oscillation patterns across SOA. In that experiment, task-relevant stimuli were always presented after task-irrelevant stimuli, so the theta oscillations primarily reflected perceptual prediction processes. In the present study,

task-relevant stimuli could be presented before or after task-irrelevant stimuli. Therefore, the theta rhythm observed behaviorally involves not only perceptual prediction but also conflict processing, demonstrating that theta rhythms in behavioral oscillations exist in higher-level cognitive control processing.

Additionally, the study found that the neutral condition oscillated in the alpha (9 Hz) band. Alpha rhythm in behavioral oscillation research is typically interpreted as the inherent perceptual sampling frequency of primary sensory cortex (Dugué & VanRullen, 2017; Michel et al., 2022), widely present in visual (Jia et al., 2022) and auditory processing (Ho et al., 2019). Consistently, in the neutral condition, task-irrelevant stimuli contained no color information and did not compete with task-relevant stimuli. Therefore, the neutral condition did not involve attentional resource allocation but rather perceptual processing. We speculate that alpha oscillations may reflect perceptual processing of task-relevant channel stimuli in auditory or visual tasks.

Notably, only the neutral condition exhibited alpha rhythmic oscillations across SOA, while other conditions showed theta rhythmic oscillations. Compared with the neutral condition, processing of congruent, stimulus incongruent, and response incongruent conditions all involved attentional resource allocation. These results support the view that perception samples at alpha rhythm while attention samples at theta rhythm (Dugué & VanRullen, 2017; Michel et al., 2022). Dugué and VanRullen (2017) proposed that visual cortex samples environmental information at its inherent alpha rhythm and transmits it to higher brain regions; when attentional modulation is required, higher brain regions feed back to visual cortex via theta oscillations, resetting sensory cortex alpha rhythm. According to task relevance, behavioral alpha rhythms may be masked by theta rhythms, resulting in theta oscillations at the behavioral level. The rhythmic theory of attention further proposes that theta rhythms may reflect periodic modulation of sensory or motor cortex by higher brain regions (Fiebelkorn & Kastner, 2019). Based on this theory, we speculate that the presentation of the first stimulus resets neural oscillations in the corresponding sensory cortex, which then samples at alpha rhythm. When the second stimulus appears, if it is irrelevant to the first stimulus (neutral condition), no attentional resource allocation is involved, higher brain regions do not feed back to sensory cortex, and sensory cortex continues sampling at alpha rhythm, resulting in alpha rhythmic oscillations in the neutral condition. If the second stimulus is relevant to the first stimulus (congruent, stimulus incongruent, or response incongruent conditions), attentional resource allocation between different stimuli or response options is required (Nigbur et al., 2012). The midfrontal region modulates sensory or motor cortex via theta oscillations, resetting sensory cortex alpha oscillations and resulting in theta oscillations at the behavioral level.

## 4.2 Cross-Modal Stimulus and Response Conflict Magnitudes Exhibit Theta Oscillation Patterns Across SOA

Consistent with our hypotheses, cross-modal response conflict magnitude exhibited theta oscillation patterns across SOA in auditory tasks, while cross-modal stimulus conflict magnitude exhibited theta oscillation patterns across SOA in visual tasks. The stage at which theta oscillations occur in cross-modal conflict processing aligns with previous EEG findings (Xu, Yang, Göschl, et al., 2024; Xu, Yang, Wu, et al., 2024), supporting the view that behavioral oscillations are behavioral manifestations of internal neural oscillations (Helfrich et al., 2018; Kienitz et al., 2022). Numerous neural oscillation results indicate that conflict processing is accompanied by enhanced midfrontal theta power (Cohen & Cavanagh, 2011; Cohen & Donner, 2013; Nigbur et al., 2012; Nigbur et al., 2011). Specifically, resolution of response conflict is accompanied by enhanced theta phase synchronization between midfrontal and lateral frontal regions, as well as between midfrontal and motor regions (Nigbur et al., 2012; Xu, Yang, Göschl, et al., 2024), whereas resolution of stimulus conflict may be associated with enhanced theta functional connectivity between midfrontal and posterior parietal regions (Nigbur et al., 2012). Based on the relationship between behavioral and neural oscillations (Helfrich et al., 2018; Kienitz et al., 2022), we speculate that theta oscillations in cross-modal response conflict magnitude across SOA in auditory tasks may reflect periodic modulation of lateral frontal and motor regions by midfrontal theta oscillations; theta oscillations in cross-modal stimulus conflict magnitude across SOA in visual tasks may reflect periodic modulation of posterior parietal regions by midfrontal theta oscillations.

This study is the first to reveal rhythmic characteristics of theta oscillations in cognitive control conflict processing at the behavioral level. Previous research has shown that attentional processes exhibit theta rhythmic oscillations at the behavioral level (Fiebelkorn et al., 2013; Jia et al., 2017; Landau & Fries, 2012; Re et al., 2019; Song et al., 2014). Based on these findings, Fiebelkorn and Kastner (2019) proposed the rhythmic theory of attention, which posits that the attention system periodically samples information via theta rhythms, with behavioral theta rhythms reflecting top-down periodic modulation of sensory or motor cortex by higher brain regions (Fiebelkorn & Kastner, 2019). Recent evidence further demonstrates that working memory processing also exhibits theta rhythms at the behavioral level (Pomper & Ansorge, 2021), suggesting that theta rhythms in behavioral oscillations may be universally present across multiple cognitive processes. Using a cross-modal conflict processing paradigm, the present study found theta rhythmic oscillations in cross-modal conflict magnitude across different tasks. This discovery provides direct behavioral evidence that cognitive control conflict processing is rhythmic. In conflict processing research, larger conflict magnitude represents poorer cognitive control ability (Egner, 2008). This study demonstrates from a behavioral perspective that cognitive control is not fixed but dynamically variable, offering a new research perspective for investigating the dynamic mechanisms of cognitive control. Over-

all, this study provides behavioral empirical evidence for rhythmic processing in cognitive control, theoretically extending the rhythmic theory of attention to the domain of cognitive control conflict processing.

Different selective attention tasks exhibit theta oscillations at different stages of cross-modal conflict processing. Specifically, theta oscillations occur during cross-modal response conflict processing in auditory tasks and during cross-modal stimulus conflict processing in visual tasks. Previous EEG studies have used event coding theory to explain the stage at which theta oscillations occur in stimulus and response conflicts (Haciahmet et al., 2023; Xu, Yang, Göschl, et al., 2024). This theory proposes that the brain uses the same coding scheme to encode perceptual, motor, and context-related features of information, representing these features in a distributed manner in event files. Selectively enhancing stimulus or response features can strengthen the encoding of corresponding features in event files, thereby influencing attentional selection of stimulus or response features (Hommel, 2019) and ultimately affecting the stage at which theta oscillations occur. However, in the present study, task-irrelevant stimuli could be presented before or after task-relevant stimuli in both visual and auditory tasks. Therefore, it is difficult to infer from the experimental design whether the brain selectively enhanced stimulus or response features, which would affect the stage of theta oscillations in conflict processing. Consequently, event coding theory cannot adequately explain the current findings. Future research should further investigate the specific mechanisms by which different selective attention tasks influence the stage of theta oscillations in cross-modal stimulus and response conflicts.

### 4.3 Limitations and Future Directions

Although the current study reveals the relationship between theta oscillations and conflict magnitude, extending the rhythmic theory of attention to cognitive control conflict processing, several limitations remain. While we speculate that theta oscillations may reflect periodic modulation of sensory or motor cortex by midfrontal regions, the current study cannot clarify how midfrontal regions dynamically modulate sensory or motor cortex to produce theta rhythmic conflict effects at the behavioral level. Future research could combine intracranial EEG techniques with high temporal resolution behavioral sampling to explore this question more deeply.

Furthermore, the SOA settings in this study included both task-irrelevant stimulus-first ( $SOA < 0$ ) and task-relevant stimulus-first ( $SOA > 0$ ) conditions, which may yield different results than designs with only a single presentation order ( $SOA > 0$  or  $SOA < 0$ ). Previous research using EEG technology to investigate cognitive control mechanisms across different SOAs (-200 ms to 200 ms) found that conflict monitoring (incongruency negativity, Ninc) and conflict resolution (late positive component, LPC) indices were present across SOAs (Appelbaum et al., 2014), suggesting that conflict processing across different SOAs may share similar cognitive control mechanisms. Based on

this, we speculate that conflict magnitude across SOA may exhibit theta rhythmic oscillations regardless of whether task-irrelevant stimuli are presented first or second. However, previous research shows that task-irrelevant stimuli presented first produce larger conflict magnitude than when presented second (Donohue et al., 2013), indicating differences in cognitive control demands. Studies have found that cognitive control theta rhythms shift toward lower frequencies as cognitive control demands increase (Senoussi et al., 2022). Therefore, we speculate that task-irrelevant stimulus-first and stimulus-second conditions may differ in theta frequency magnitude. Since the SOA range for task-irrelevant stimulus-first and stimulus-second conditions in this study was only 200 ms, which does not fully cover one theta oscillation cycle (~250 ms), the current study cannot verify this speculation. Future research could separately investigate behavioral oscillations for task-irrelevant stimulus-first and stimulus-second conditions to further explore potential differences in oscillation patterns.

## 5 Conclusion

The present study used a 2-to-1 audiovisual Stroop task combined with high temporal resolution behavioral sampling and found that in auditory tasks, cross-modal response conflict magnitude exhibited theta oscillation patterns across SOA; in visual tasks, cross-modal stimulus conflict magnitude exhibited theta oscillation patterns across SOA. These findings demonstrate that cognitive control processing exhibits rhythmic patterns at the behavioral level, revealing the relationship between theta oscillations and cross-modal stimulus and response conflict magnitude and extending the rhythmic theory of attention.

## Appendix: Comparison Between Experiment 1 and Experiment 2

To compare cross-modal stimulus and response conflict magnitudes across task conditions, we calculated stimulus and response conflict magnitudes from Experiments 1 and 2 and conducted a 2 (task condition: auditory, visual)  $\times$  2 (conflict effect: stimulus conflict, response conflict) mixed-design ANOVA, with task condition as a between-subjects factor and conflict effect as a within-subjects factor. Results showed significant main effects of task condition,  $F(1, 78) = 158.97$ ,  $p < 0.001$ ,  $\eta^2 = 0.67$ , and conflict effect,  $F(1, 78) = 36.87$ ,  $p < 0.001$ , and a significant interaction,  $F(1, 78) = 29.13$ ,  $p < 0.001$ . Simple effects analysis revealed that cross-modal stimulus conflict magnitude was significantly larger in auditory tasks ( $M = 59$  ms,  $SE = 3.54$ ) than in visual tasks ( $M = 8$  ms,  $SE = 3.54$ ),  $t(39) = 12.60$ ,  $p < 0.001$ , Cohen's  $d = 2.82$ , 95% CI = [41.40, 61.34]. Cross-modal response conflict magnitude was also significantly larger in auditory tasks ( $M = 25$  ms,  $SE = 2.02$ ) than in visual tasks ( $M = 6$  ms,  $SE = 2.02$ ),  $t(39) = 4.77$ ,  $p < 0.001$ , Cohen's  $d = 1.07$ , 95% CI = [13.77, 25.16]. Results are shown in Figure A1. These findings indicate that both cross-modal stimulus and response conflict magnitudes were larger in auditory than visual tasks, with

greater differences in stimulus conflict than response conflict between tasks.

Regarding theta oscillations of stimulus and response conflict magnitudes, since conflict magnitudes can only be calculated at the subject level, single-trial fitting was not possible. Following previous research (Fiebelkorn et al., 2013), we used aggregate subject data analysis. As this method yields only one theta value for stimulus and response conflict magnitudes, statistical analysis between experiments was not possible. We discussed differences in theta values for conflict magnitudes between Experiments 1 and 2 in the Discussion section of Experiment 2.

**Figure A1.** Comparison of cross-modal stimulus conflict magnitude and cross-modal response conflict magnitude between Experiment 1 (auditory task) and Experiment 2 (visual task)

*Note: Figure translations are in progress. See original paper for figures.*

*Source: ChinaXiv –Machine translation. Verify with original.*