Schizophrenia Research 281 (2025) 191–200

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Distinct processing stages of cross-modal conflict in schizophrenia: The role of auditory cortex underactivation

Heng Zhou [c,d], Xiangyong Yuan [c,d], Pei Xie [e], Aijun Wang [a,b,*], Yi Jiang [c,d,**]

[a] Department of psychology, Faculty of Education, Guangxi Normal University, Guilin, China
[b] Guangxi College and University Key Laboratory of Cognitive Neuroscience and Applied Psychology, Guilin, China
[c] State Key Laboratory of Cognitive Science and Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
[d] Department of Psychology, University of Chinese Academy of Sciences, Beijing, China
[e] Department of Psychology, Sichuan Normal University, Chengdu, China

ARTICLE INFO

Keywords:
Schizophrenia
Cross-modal conflict
Semantic conflict
Response conflict
Auditory cortex

ABSTRACT

Background: The cross-modal conflict deficit is a key feature of schizophrenia. However, it remains largely unknown whether cross-modal conflict in schizophrenia diverges at distinct processing stages and its potential association with the auditory cortex.

Methods: In Experiment 1, we divided cross-modal conflict into semantic and response stages, and we investigated the cross-modal conflict between schizophrenia patients ($n = 30$) and health individuals ($n = 32$). In Experiment 2, we utilized tDCS (transcranial direct current stimulation) to inhibit the activity of the auditory cortex in healthy individuals ($n = 20$), and we substituted auditory sounds with visual words in healthy individuals ($n = 34$) in Experiment 3, exploring the association between the patients' cross-modal conflict patterns and the auditory cortex. Furthermore, we employed machine learning techniques to further validate the stability of the distinct pattern.

Results: We found that schizophrenia patients exhibited auditory dominance at the semantic conflict stage and visual dominance at the response conflict stage, contrary to healthy individuals. By causally interfering with the normal function of the auditory cortex in healthy individuals, we observed behavioral similarities to those with schizophrenia, supporting the critical role of insufficient auditory cortex activation in the early development of schizophrenia. The classification analysis further confirmed the double dissociation of cross-modal conflicts in schizophrenia and the role of auditory cortex underactivation.

Conclusions: These findings not only demonstrate a unique mechanism and its neural correlate in how schizophrenia patients cope with cross-modal conflicts but also provide potential early diagnostic markers or therapeutic targets for schizophrenia.

1. Introduction

Schizophrenia is a severe and persistent mental disorder, and significant resources have been invested in its diagnosis, treatment, and research (Cloutier et al., 2016; Donde et al., 2023). While numerous studies have focused on cognitive dysfunction in schizophrenia patients, less attention has been given to their deficits in multisensory integration (Donde et al., 2019; Wallace et al., 2020). In our complex world, consistent information from different sensory modalities should be integrated into a unified and coherent whole (Stein and Stanford, 2008; Tang et al., 2016). However, when information from multiple sensory modalities is inconsistent, it can result in cross-modal conflicts. Typically, if such interferences between sensory modalities are asymmetric, one sensory modality will dominate the other during cross-modal conflicts (Callan et al., 2015; Hirst et al., 2018; Zhou et al., 2010).

Cross-modal conflicts and sensory dominance in these conflicts are observed in schizophrenia patients. Compared to healthy individuals, schizophrenia patients exhibit unique pattern in cross-modal conflicts. For example, when asked to watch lip movements while listening to speech, their speech perception is less affected by the lip movements (de Gelder et al., 2003; Pearl et al., 2009; White et al., 2014), especially in noisy environments (Ross et al., 2007; Wallace and Stevenson, 2014).

* Correspondence to: A. Wang, Department of Psychology, Guangxi Normal University, 15 Yucai Road, Guilin 541004, China.
** Correspondence to: Y. Jiang, Institute of Psychology, Chinese Academy of Sciences, 16 Lincui Road, Beijing 100101, China.
E-mail addresses: ajwang@suda.edu.cn (A. Wang), yijiang@psych.ac.cn (Y. Jiang).

https://doi.org/10.1016/j.schres.2025.05.014
Received 4 January 2025; Received in revised form 12 May 2025; Accepted 14 May 2025
Available online 19 May 2025
0920-9964/© 2025 Published by Elsevier B.V.

Similarly, their categorization of emotional voices is not profoundly influenced by simultaneously presented emotional faces (de Jong et al., 2009). These studies suggest that vision may dominate audition in auditory discrimination tasks to a larger extent in healthy controls than in schizophrenia patients. Conversely, audition may strongly dominate vision in visual discrimination tasks in schizophrenia patients compared to healthy controls. For instance, when there are mismatched flashes and beeps simultaneously, healthy controls tend to judge the number of flashes based on the number of beeps, but schizophrenia patients and individuals with high schizotypal traits are more prone to illusorily report the number of flashes (Ferri et al., 2018; Haß et al., 2017).

Although previous research suggests that cross-modal conflict patterns in schizophrenia patients differ from healthy individuals, no studies have directly evaluated the asymmetry in cross-modal conflicts within a single paradigm. This requires a modality-general task that simultaneously tests the extent of vision and audition dominance in cross-modal conflicts in schizophrenia patients and compare it to healthy controls. Chen and Zhou (2013) designed a categorization task that asks participants to attend to one modality (e.g., visual) and classify stimuli in this modality into two categories (e.g., politicians or movie stars) while ignoring simultaneous stimuli in the other unattended modality. By exchanging the attended modality, they could simultaneously measure visual and auditory dominance in cross-modal conflicts. Furthermore, this paradigm allows for the division of cross-modal conflicts into semantic and response stages (see Fig. 1A in Chen and Zhou, 2013, and Fig. 1B here), which depend on distinct neural mechanisms (Xu et al., 2024). Given the advantage of this paradigm, the current study adopted it to elucidate sensory dominance at both the semantic and response conflict stages in schizophrenia patients and compared them to healthy controls in order to more precisely characterize distinct processing stages of cross-modal conflicts in schizophrenia.

Based on previous findings, we expect that when facing cross-modal conflicts, visual interference with auditory discrimination will decrease (White et al., 2014), while auditory interference with visual discrimination will increase in schizophrenia patients (Ferri et al., 2018). Regarding the conflict stage, studies using the cross-modal conflict paradigm have shown that visual dominance occurs at the semantic conflict stage, while auditory dominance occurs at the response conflict stage for healthy individuals (Chen and Zhou, 2013; Li et al., 2019). However, although some researchers have noticed the significance of separating the two conflict stages in schizophrenia patients (Ettinger et al., 2018; Westerhausen et al., 2011), no clear conclusion has been reached about which conflict stage auditory and visual dominance occurs in schizophrenia.

In addition to characterizing the distinct cross-modal conflict

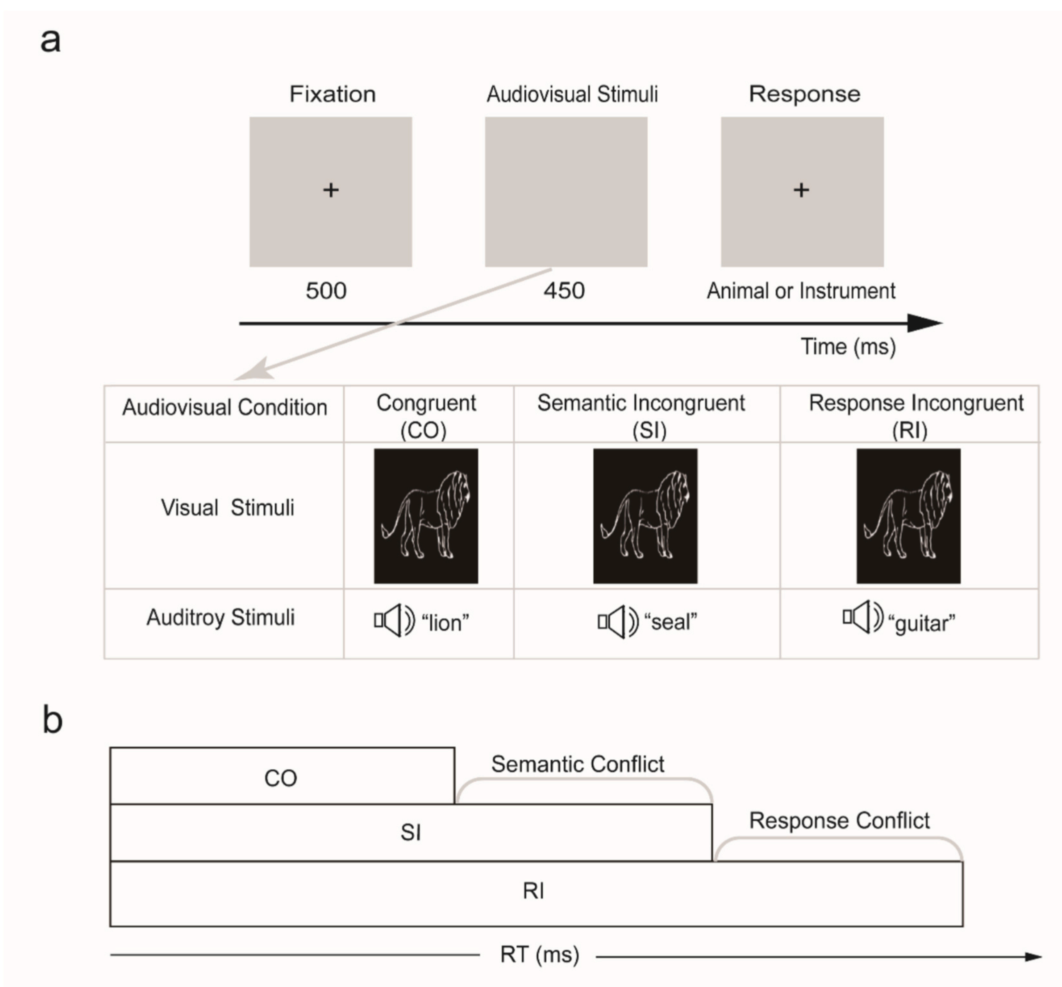

Fig. 1. Illustration of stimuli and procedure for Experiment 1. (a) At the start of each trial, the cross fixation was presented in the middle of the screen for 500 ms, followed by the audiovisual stimuli which lasted 450 ms. Participants were required to categorize the stimuli (animal or musical instrument) at their attended modality (visual or auditory) as quickly as possible while ensuring accuracy. The audiovisual stimuli constituted three conditions: the congruent (CO), semantic incongruent (SI) and response incongruent (RI) condition. (b) Illustration of the conflict effects in the semantic and response stages. The semantic conflict effect is calculated by subtracting RT_{CO} from RT_{SI}, while the response conflict effect is calculated by subtracting RT_{SI} from RT_{RI}.

patterns in schizophrenia patients relative to healthy individuals, the current study aimed to further explore its underlying mechanism. Schizophrenia patients are usually accompanied by typical hearing impairment (Schneider et al., 2016), and patients with hearing impairment have a high probability of suffering from symptoms such as auditory hallucinations (Zhuo et al., 2020). Neurobiological studies have shown that hearing impairment or impaired auditory processing in patients (Koenig et al., 2012; Zhuo et al., 2020) is likely due to significantly weakened activation of the auditory cortex caused by a reduction of neurotransmitters (Moyer et al., 2012). As hearing impairment may bias the weight of audition in perceptual decision-making during cross-modal conflicts, we speculate that insufficient activation of the auditory cortex underlies the distinct cross-modal conflict patterns in schizophrenia patients compared to healthy individuals.

Three experiments, all adopting the cross-modal conflict paradigm, were conducted to address these questions. Experiment 1 investigated cross-modal conflicts at the semantic and response stages in schizophrenia patients and healthy individuals (Chen and Zhou, 2013; Xu et al., 2024). The following two experiments investigated whether the distinct cross-modal conflict pattern in schizophrenia patients is related to insufficient activation of the auditory cortex. Specifically, Experiment 2 used transcranial direct current stimulation (tDCS) to inhibit auditory activities in healthy individuals, while Experiment 3 replaced the auditory stimuli presented to healthy individuals with visual stimuli of identical semantic meaning, which also maximally deactivated the auditory cortex. If our hypothesis is corrected, these operations would cause the healthy group to replicate a cross-modal conflict pattern similar to that of the schizophrenia patients, confirming that insufficient auditory activation is the main cause of the distinct cross-modal conflict pattern in schizophrenia patients. Additionally, we further explored whether the cross-modal conflict pattern can be used as a behavioral diagnosis to distinguish schizophrenia patients from healthy individuals.

2. Methods

2.1. Participants

Based on the effect size ($\eta_p^2 = 0.29$) reported by Xu et al. (2024), with α set at 0.05 and power (1-β) of 0.80, we conducted an a priori power analysis using G*Power 3.1. The *F*-test was selected as the appropriate statistical test for our 2 (attended modality) by 3 (congruency) within-subjects design. The analysis indicated that a minimum sample size of 14 participants would provide adequate statistical power to detect cross-modal conflict effects in healthy individuals. To investigate potential cross-modal conflict deficits in schizophrenia patients, we increased the sample size to ~30. In Experiment 1, we recruited 30 schizophrenia patients (9 females, mean age = 29.86 years, $SD = 7.30$) and 32 healthy controls (19 females, mean age = 20.34 years, $SD = 1.93$). Experiment 2 recruited 20 healthy controls (10 females, mean age = 22.35 years, $SD = 1.81$), and Experiment 3 recruited 34 healthy controls (19 females, mean age = 21.15 years, $SD = 1.88$).

All schizophrenia patients were diagnosed by two psychiatrists, meeting the DSM-5 diagnostic criteria, and had no history of substance abuse, intellectual disability, or other organic diseases. Patients had an average illness duration of 6.6 years. The Scale for the Assessment of Positive Symptoms (SAPS) score was 1.57 ($SD = 2.43$) and the Scale for the Assessment of Negative Symptoms (SANS) score was 13.9 ($SD = 11.56$). All patients exhibited minimal positive and negative symptoms. During the experimental phase, all patients were taking atypical antipsychotic medications such as risperidone, quetiapine, and clozapine. All participants had normal or corrected-to-normal vision and hearing, and had not previously participated in similar experiments. They received compensation for their participation and provided written informed consent. This study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of the Academic Committee of the Department of Psychology, Soochow University.

2.2. Stimuli and procedure

Participants sat comfortably in a dimly lit room, approximately 60 cm away from an LCD monitor with a refresh rate of 60 Hz and a resolution of 1920 × 1080. All experimental programs were presented on the monitor using Matlab 2014b (The MathWorks) equipped with Psychtoolbox (Brainard, 1997). The visual stimuli were selected from a standard outline gallery by Snodgrass and Vanderwart (1980) and included three animal pictures (elephant, lion, seal) and three musical instrument pictures (guitar, piano, flute) with a visual angle 5° horizontally × 6° vertically. The auditory stimuli were the Chinese names of the pictures, articulated at approximately 70 dB through a head-mounted iron triangle headset (ATH-WS99).

Experiment 1 examined cross-modal conflicts in schizophrenia patients and its distinctive pattern compared to healthy individuals. At the beginning of each block, an instruction was presented for 2000 ms to inform participants which stimuli, the visual or the auditory, they should focus on. Then, each trial started with a 500-ms black fixation cross (1.5° × 1.5°), followed by a picture and a sound simultaneously presented for 450 ms (Fig. 1). Afterward, participants indicated whether their attended stimuli were animals or musical instruments within a time limit of 3000 ms by pressing two separate buttons on the computer keyboard. The mapping between these two buttons and their indications was counterbalanced across participants. The next trial began in an interval of 500 ms. There were three audiovisual conditions, the congruent (CO), the semantic incongruent (SI) and the response incongruent (RI). In the CO condition, the simultaneously presented pictures and sounds were completely matched. e.g., a picture of a lion is seen and a word 'lion' is heard. In the SI condition, the pictures and sounds did not correspond to the same object but were assigned to the same category. e.g., a picture of a lion is seen but another animal's name is heard. In the RI condition, the pictures and sounds did not correspond to the same object either, but came from different categories which would lead to different key responses, e.g., a picture of a lion is seen but the word 'guitar' is heard. Each condition consisted of 96 trials, resulting in a total of 288 trials distributed across 24 blocks. Half the participants attended the visual stimuli first, while the other half attended the auditory stimuli first.

To determine whether decreased activities in the auditory cortex would lead to a schizophrenia-like cross-modal conflicts in healthy individuals, Experiment 2 tested them in the same cross-modal conflict task but applied brain stimulation on their left auditory cortex. A TES2001 Transcranial Direct Current Stimulation Device (SOTERIX, USA), together with two 5 × 7 cm sponge electrodes covered with saline-soaked sponges was used. The electrode placement followed the international EEG 10–20 system standards. The cathodal electrode was positioned between C5 and T7, on the scalp surface of the left auditory cortex, while the anodal electrode was placed approximately five centimeters above the right eye (Fig. 2c), just over the right orbital area. The cathodal stimulation was delivered at an intensity of 2 mA, lasting for 10 min, with 30-s fade-in and fade-out phases. A sham stimulation was applied to the same location serving as a control. It was also delivered at an intensity of 2 mA but for 30 s only. Experiment 2 shared the same stimuli and procedures as Experiment 1 except that the participants underwent the transcranial direct current stimulation (tDCS) immediately before the task. All participants completed the task twice within two days and were unaware of the type of stimulation they received. Half of them received the sham stimulation first, while the other half received the cathodal stimulation first.

To determine whether healthy individuals would behave like the schizophrenia patients when exposed to a conflict that does not strongly activate the auditory cortex, Experiment 3 presented the object names in written words on top of the pictures, instead of displaying them aurally.

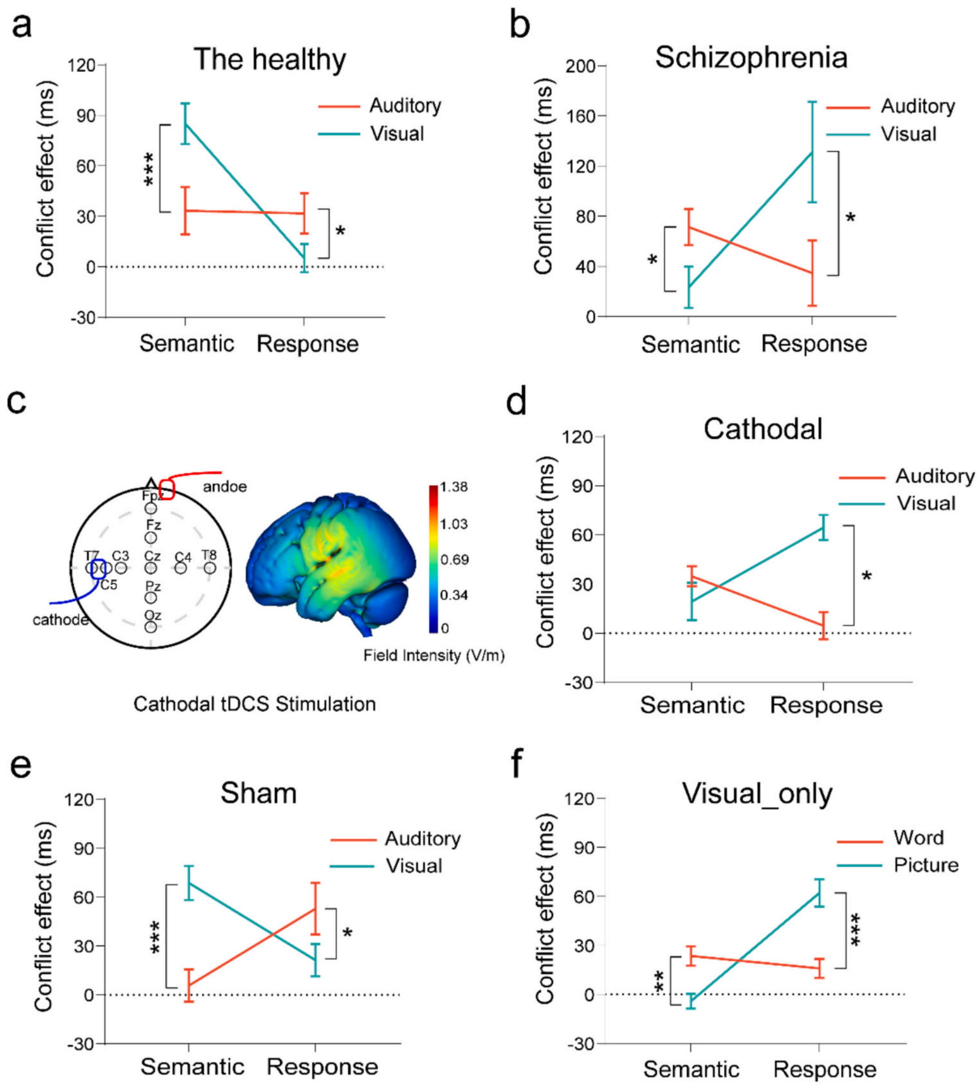

Fig. 2. Results from Experiments 1–3. Conflict effects in the healthy (a) and schizophrenia group (b) in Experiment 1. The field intensity of the cathodal tDCS stimulation placed in the left auditory cortex is drawn in (c). The conflict effects in the cathodal (d) and the sham (e) group in Experiment 2. The conflict effects in the visual-only group in Experiment 3 (f). Asterisks indicate significant differences between the semantic and response conflict effects when different modalities were attended (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

Before each block, participants were instructed to attend to either the pictures or the words. The remaining procedures were the same as in Experiment 1. Therefore, the congruent and congruent conditions remained consistent in Experiment 3, but the resulting conflicts are restricted in the visual domain only. If disturbed auditory activities is the key determinant for the cross-modal conflict pattern in the schizophrenia, we would expect that the conflicts in healthy individuals, as long as disengage the auditory cortices, are similar to cross-modal conflicts in schizophrenia patients.

2.3. Data analysis

As the task accuracies of healthy individuals were close to a ceiling (~95 %), we only performed a 2 attended modality (Visual vs. Auditory) × 3 condition (CO vs. SI vs. RI) repeated measures ANOVA on the accuracy of schizophrenia patients (78 %).

The RT analysis mainly focused on cross-modal conflicts at the semantic and response conflict stages. First, incorrect trials and correct trials with RTs beyond 3 standard deviations of the mean were exclude. Next, we calculated the semantic conflict effect by subtracting the RT in the CO condition from the RT in the SI condition (RT_{SI} - RT_{CO}), and the response conflict effect by subtracting the RT in the SI condition from the RT in the RI condition (RT_{RI} - RT_{SI}). Lastly, we conducted a repeated measures ANOVA of 2 attended modality (Visual vs. Auditory) × 2 conflict stage (Semantic vs. Response) on the conflict effects. All the post-hoc tests were corrected by Bonferroni method.

After obtaining the semantic and response conflict effects for each attended modality in the three experiments, we further investigated whether they are informative enough to accurately classify participants into their belonging group, the healthy or schizophrenia. Here, we trained a support vector machine (SVM) as a classifier model (Chen et al., 2020; Fung and Mangasarian, 2005), and compressed the four conflict effects into two features by subtracting the semantic and responses conflict effects when visual modality was attended from their counterparts when auditory modality was attended. We performed the commonly used leave-one-out cross-validation approach to validate the performance of the classifier, and the accuracy (ACC) of the classifier was measured by the proportion of observations that were correctly classified into their belonging group. These processes were completed using the fitcsvm function in Matlab.

In addition, we calculated the sensitivity, specificity, d-prime and the area under the receiver operating characteristic (ROC) curve (AUC) to further evaluate the classification performance (Table 2 and Fig. 3C–E):

$$\text{Sensitivity} = \text{TP}/(\text{TP}+\text{FN}),$$

$$\text{Specificity} = \text{TN}/(\text{TN}+\text{FP}),$$

$$\text{d-prime} = Z\,(\text{sensitivity}) - Z\,(1-\text{specificity}),$$

where the number of true positives (TP) indicates the number of patients correctly classified, the number of true negatives (TN) indicates the number of healthy individuals correctly classified, the number of false positives (FP) referred to the number of healthy individuals mistakenly classified as patients, and the number of false negatives (FN) referred to the number of patients wrongly classified as healthy individuals.

Permutation tests were employed to assess whether the model evaluation indices (ACC, d-prime and AUC) would be significantly higher than the chance level. Specifically, we shuffled the labels to which each participant belongs 1000 times, and then repeated the aforementioned SVM process to obtain the null distribution for each index. Consequently, we could compute their corresponding *p*-values

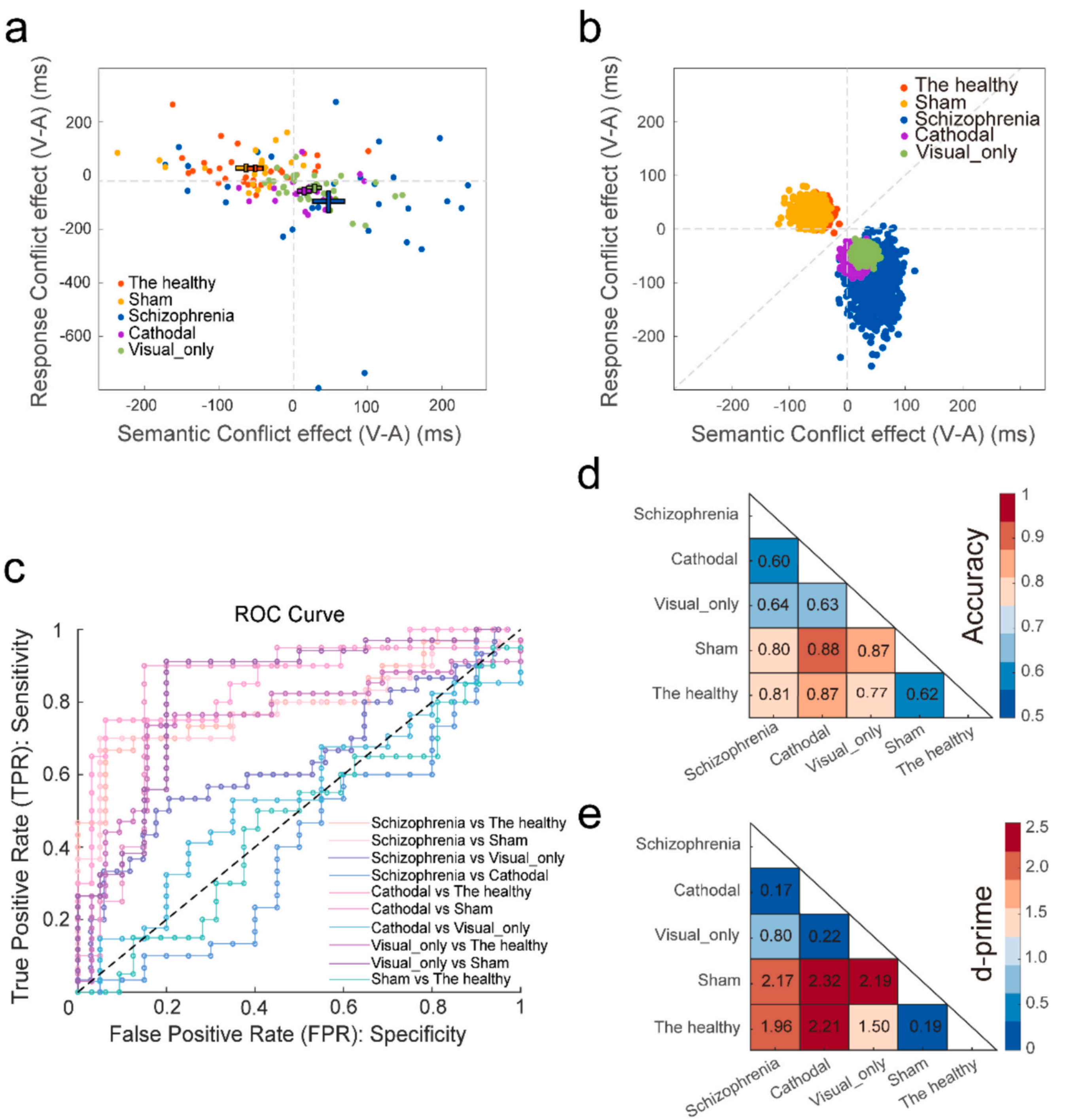

Fig. 3. The individual semantic and response conflict effects (A-V) in all experiments. (a) Each colored dot represents one participant, and each colored cross represents the standard error for each conflict effect (A-V). (b) Bootstrapped semantic and response conflict effects (A-V) for each group, separately. Receiver-operating characteristic (ROC) curve (c) and the heatmaps of the classification accuracy (d) and d-prime (e) were plotted for each classification.

from these null distributions.

3. Results

3.1. Distinct cross-modal conflict pattern in schizophrenia patients compared to healthy individuals

Experiment 1 examined and compared cross-modal conflicts in schizophrenia patients and healthy controls at both the semantic and response conflict stages. Participants were required to attend to and respond to visual (or auditory) stimuli while ignoring simultaneously presented auditory (or visual) distractors. There were three conditions: congruent (CO), where the visual and auditory stimuli completely matched; semantic incongruent (SI), where the visual and auditory stimuli did not correspond to the same object, but their corresponding keys matched; and response incongruent (RI), where the visual and auditory stimuli neither corresponded to the same object nor were assigned to the same key (Fig. 1a). Overall, schizophrenia patients exhibited lower accuracy compared to healthy controls (78 % vs. 95 %, $t(60) = -5.005$, $p < 0.001$, Cohen's $d = -1.272$, 95 % CI = [−0.228, −0.098]). However, the main effect of condition was significant ($F(2,58) = 27.226$, $p < 0.001$, $\eta_p^2 = 0.484$), neither the main effect of attended modality nor their interact were significant ($Fs < 1$). Post hoc tests revealed that the accuracy of CO (84 %) was significantly higher than that of RI (66 %), and the accuracy of SI (84 %) was significantly higher than that of RI ($ps < 0.001$), suggesting a general conflict impairment in schizophrenia patients independent of modality.

Our analysis primarily focused on the reaction time (RT) in cross-modal conflicts (The raw RTs and the number of error trials were shown in Table 1). By computing the RT difference between conditions, we obtained the semantic conflict effect (RT_{SI} - RT_{CO}) and the response conflict effect (RT_{RI} - RT_{SI}), respectively (Fig. 1b). We conducted a 2 (group) by 2 (attended modality) by 2 (conflict stage) repeated measures ANOVA. The analysis revealed a significant group by conflict stage interaction ($F(1,60) = 5.723$, $p = 0.020$, $\eta_p^2 = 0.087$). More importantly, we observed a significant three-way interaction among group, attended modality, and conflict stage ($F(1,60) = 19.664$, $p < 0.001$, $\eta_p^2 = 0.247$). To further investigate these interaction effects, we proceeded with separate analyses of cross-modal conflict patterns for each group. For healthy individuals, a 2 (attended modality) by 2 (conflict stage) repeated measures ANOVA showed a significant main effect of conflict stage ($F(1,31) = 9.174$, $p = 0.005$, $\eta_p^2 = 0.228$) and a significant interaction ($F(1,31) = 20.459$, $p < 0.001$, $\eta_p^2 = 0.398$), while the main effect of attended modality was not significant ($F(1,31) = 3.078$, $p = 0.089$, $\eta_p^2 = 0.090$), as illustrated in Fig. 2a. Post hoc tests revealed that the semantic conflict effect was significantly larger when the auditory modality was attended than when the visual modality was attended (85 ms vs. 33 ms, $t(31) = 4.752$, $p < 0.001$, Cohen's $d = 0.840$, 95 % CI = [29.525, 73.933]). Conversely, the response conflict effect showed an opposite pattern (5 ms vs. 32 ms, $t(31) = -2.296$, $p = 0.029$, Cohen's $d = 0.406$, 95 % CI = [−50.198, −2.970]). These results align with previous studies (Chen and Zhou, 2013; Li et al., 2019), indicating that vision dominated in the semantic conflict stage, while audition dominated in the response stage.

Intriguingly, schizophrenia patients exhibited opposite behavioral patterns to healthy individuals in the cross-modal conflict task, as shown in Fig. 2b. The same 2 (attended modality) by 2 (conflict stage) repeated measures ANOVA revealed a significant interaction ($F(1,29) = 8.872$, $p = 0.006$, $\eta_p^2 = 0.234$), while the main effects of attended modality ($F(1,29) = 1.297$, $p = 0.264$, $\eta_p^2 = 0.043$) and conflict stage ($F(1,29) = 1.438$, $p = 0.240$, $\eta_p^2 = 0.047$) were not significant. The semantic conflict effect was significantly larger when the visual modality was attended compared with when the auditory modality was attended (23 ms vs. 71 ms, $t(29) = -2.237$, $p = 0.033$, Cohen's $d = 0.408$, 95 % CI = [−92.024, −4.123]), while the response conflict effect depicted an opposite pattern (131 ms vs. 35 ms, $t(29) = 2.397$, $p = 0.023$, Cohen's $d = 0.438$, 95 % CI = [14.171, 178.867]). Unlike healthy individuals, schizophrenia patients exhibited auditory dominance at the semantic conflict stage and visual dominance at the response conflict stage.

To quantitatively compare group differences, we contrasted the semantic and response conflict effect (A-V) between groups using independent samples t-tests. Schizophrenia patients exhibited significantly larger semantic conflict effect (A-V) than healthy individuals (−52 ms vs. 48 ms, $t(60) = -4.221$, $p < 0.001$, Cohen's $d = 1.073$, 95 % CI = [−147.096, −52.509]). Conversely, healthy individuals showed significantly greater response conflict effect (A-V) than schizophrenia patients (27 ms vs. −97 ms, $t(60) = 3.020$, $p = 0.004$, Cohen's $d = 0.768$, 95 % CI = [41.573, 204.633]). These findings provide compelling evidence for distinct cross-modal conflict processing patterns between schizophrenia patients and healthy individuals.

3.2. Healthy individuals showed a schizophrenia-like cross-modal conflict pattern when their auditory cortex were inhibited

To investigate the potential link between underactivation of the auditory cortex and the distinct cross-modal conflict pattern observed in schizophrenia patients, we conducted Experiment 2. Prior to the cross-modal conflict task, we administered tDCS cathodal stimulation to the left auditory cortex of healthy individuals (Fig. 2c) to inhibit the neuronal activity. We performed a one-way ANOVA comparing accuracy between healthy individuals and participants receiving tDCS stimulation, which yielded no significant group difference ($F < 1$). This null result suggested that the tDCS did not produce generalized effects on response accuracy. Then the same 2 (attended modality) by 2 (conflict stage) repeated measures ANOVA showed a significant main effect of attended modality ($F(1,19) = 7.846$, $p = 0.011$, $\eta_p^2 = 0.292$) and a significant interaction ($F(1,19) = 28.343$, $p < 0.001$, $\eta_p^2 = 0.599$), whereas the main effect of conflict stage was not significant ($F(1,19) = 0.595$, $p = 0.450$, $\eta_p^2 = 0.030$), as illustrated in Fig. 2d. The semantic conflict effect was not significantly different between when the visual modality or the auditory modality was attended (35 ms vs. 19 ms, $t(19) = 1.653$, $p = 0.115$, Cohen's $d = 0.370$, 95 % CI = [−4.085, 34.736]). However, the semantic conflict effect was significantly larger when the auditory modality was attended than when the visual modality was attended (64 ms vs. 5 ms, $t(19) = 5.055$, $p < 0.001$, Cohen's $d = 0.426$, 95 % CI = [35.095, 84.694]). It is evident that the cross-modal conflict pattern of healthy individuals following cathodal stimulation was similar to that of schizophrenia patients.

Table 1
Mean RT (ms) and number of error trials for each group (±SE).

Group	Attended visual			Attended auditory			Error trials
	CO	SI	RI	CO	SI	RI	
Schizophrenia	872.16 ± 41.24	943.48 ± 47.56	978.10 ± 59.83	1088.97 ± 49.50	1112.22 ± 47.87	1243.35 ± 60.10	62.73 ± 9.51
The healthy	495.98 ± 30.31	529.30 ± 36.72	561.03 ± 44.65	776.55 ± 26.94	861.60 ± 29.59	866.76 ± 30.22	15.81 ± 1.71
Sham	508.81 ± 21.60	514.54 ± 21.10	567.49 ± 31.10	663.92 ± 24.23	732.57 ± 27.17	753.88 ± 23.72	16.30 ± 2.30
Cathodal	475.12 ± 17.72	509.83 ± 18.63	514.43 ± 23.87	656.52 ± 22.78	675.90 ± 21.27	740.40 ± 20.28	16.20 ± 2.92
Visual-only	467.02 ± 12.70	489.24 ± 14.92	505.23 ± 15.81	516.95 ± 13.83	512.83 ± 13.80	575.35 ± 17.54	13.76 ± 1.49

Under the sham stimulation condition, we replicated the semantic and response conflict effects found in the healthy group of Experiment 1 (Fig. 2e). The main effects of attended modality ($F(1,19) = 2.772$, $p = 0.112$, $\eta_p^2 = 0.127$) and conflict stage ($F(1,19) = 1.352 \times 10^{-5}$, $p = 0.997$, $\eta_p^2 = 7.115 \times 10^{-7}$), were not significant; however, the interaction was significant ($F(1,19) = 22.028$, $p < 0.001$, $\eta_p^2 = 0.537$). At the semantic conflict stage, the conflict effect was significantly larger when the auditory modality was attended than when the visual modality was attended (69 ms vs. 6 ms, $t(19) = 4.397$, $p < 0.001$, Cohen's $d = 0.983$, 95 % CI = [32.970, 92.880]), indicating visual dominance, while at the response conflict stage, audition dominated (21 ms vs. 53 ms, $t(19) = -2.394$, $p = 0.027$, Cohen's $d = 0.535$, 95 % CI = [−59.308, −3.982]). In short, Experiment 2 replicated the cross-modal conflict pattern of schizophrenia patients in healthy individuals by inhibiting their auditory cortex activity, causally verifying that insufficient auditory cortex activation is a key cause of the distinct cross-modal conflict in schizophrenia patients.

3.3. Healthy individuals showed a similar conflict pattern as schizophrenia patients when auditory sounds were replaced by visual words

The incongruent semantic meanings were distributed into two different modalities to create cross-modal conflicts. However, such conflict effects would be visual only if we presented the object name through visual words instead of auditory sounds. Previous studies have shown that visual words and auditory sounds share some commonalities in processing. Compared to auditory sounds, visual words can also activate the auditory cortex, but to a lesser extent (Haist et al., 2001). Additionally, this design might be closer to the cross-modal deficits observed in schizophrenia patients (Sass et al., 2013; Surguladze et al., 2002). For patients, according to the dysregulation of information encapsulation, although sensory information was input through different modalities, the processing of sensory information did not differentiate between within-modal and cross-modal information (Sass et al., 2013). Therefore, we could examine the conflict effects using the same logic but without substantially activating the auditory cortex. We tested whether healthy individuals, when their auditory cortex was minimally engaged in the visual-only conflict task, would largely approximate the pattern of cross-modal conflict observed in schizophrenia patients. The 2 (attended mode) by 2 (conflict stage) repeated measures ANOVA showed significant main effects of attended mode ($F(1,33) = 4.365$, $p = 0.044$, $\eta_p^2 = 0.117$) and conflict stage ($F(1,33) = 15.657$, $p < 0.001$, $\eta_p^2 = 0.322$), as well as a significant interaction ($F(1,33) = 26.409$, $p < 0.001$, $\eta_p^2 = 0.445$), as shown in Fig. 2f. The semantic conflict effect was significantly larger when the picture was attended than when the word was attended (24 ms vs. −4 ms, $t(33) = 3.485$, $p = 0.001$, Cohen's $d = 0.598$, 95 % CI = [11.543, 43.914]), while the response conflict effect showed an opposite pattern (16 ms vs. 62 ms, $t(33) = -5.192$, $p < 0.001$, Cohen's $d = 0.890$, 95 % CI = [−64.266, −28.081]). As assumed, the visual-only conflict pattern in Experiment 3 closely resembled that of schizophrenia patients. Taking into account the results of Experiments 1–3, we concluded that the hypoactivation of the auditory cortex in schizophrenia patients was a major cause of the distinct cross-modal conflict pattern.

3.4. The cross-modal conflict pattern serves as a potential biomarker to distinguish schizophrenia from healthy individuals

To further verify whether the cross-modal conflict patterns can distinguish schizophrenia from healthy individuals at the individual level, we first respectively subtracted the semantic and response conflict effects when the visual modality was attended from their counterparts when the auditory modality was attended, and plotted the individual semantic and response conflict effect (A-V) on two axes in Fig. 3a. It can be seen that healthy individuals and the sham group mainly clustered in the top left quadrant, while schizophrenia patients, the cathodal group, and the visual-only group mostly clustered in the bottom right quadrant. To illustrate this separation more clearly, we constructed and drew a distribution of the semantic and responses conflict effects (A-V) from 1000 bootstrapped samples of the raw data for each group in Fig. 3b. The bootstrapped samples suggest that the cross-modal conflict pattern could largely distinguish schizophrenia and schizophrenia-like individual from healthy individuals.

We then empirically tested this possibility using state-of-the-art machine learning algorithms to classify participants from all the three experiments into their respective groups. Specifically, we employed Support Vector Machine (SVM) as the classifier, using the semantic and responses conflict effects (A-V) as features. A leave-one-out cross-validation loop was run during the training process. We used sensitivity, specificity, d-prime, AUC, and ACC to evaluate the performance of these models. The results are shown in Fig. 3c–e and Table 2. We found that schizophrenia patients can be distinguished from healthy individuals with an accuracy of 0.81, an AUC of 0.77, and a d-prime of 1.96 ($ps < 0.001$), and from the sham with an accuracy of 0.80, an AUC of 0.79, and a d-prime of 2.17 ($ps < 0.001$). All these indicate good classification performance using only the two conflict effect features. Moreover, the classification demonstrated that schizophrenia patients and the cathodal group could not be distinguished from each other, although schizophrenia patients could be separated from the visual-only group with an accuracy slightly above chance level (0.64, $p < 0.05$). The SVM results further validate that the distinct cross-modal conflict pattern in schizophrenia patients is likely related to insufficient activation in the auditory cortex, and emphasize the effectiveness of cross-modal conflicts as a promising clinical diagnostic tool for distinguishing between schizophrenia patients and healthy individuals.

4. Discussion

Using the cross-modal conflict paradigm (Chen and Zhou, 2013; Xu et al., 2024), the current study found that schizophrenia patients exhibited auditory dominance at the semantic conflict stage and visual dominance at the response conflict stage, in contrast to healthy individuals. This distinct cross-modal conflict pattern in schizophrenia

Table 2
Each model performs on an internal test set.

Model	Sensitivity	Specificity	d-prime	AUC	ACC
Schizophrenia vs. The healthy	0.67	0.94	1.96***	0.77***	0.81***
Schizophrenia vs. Sham	0.70	0.95	2.17***	0.79***	0.80***
Schizophrenia vs. Cathodal	1.00	0	0.17	0.36	0.60
Schizophrenia vs. Visual-only	0.40	0.85	0.80**	0.61*	0.64*
Cathodal vs. The healthy	0.75	0.94	2.21***	0.83***	0.87***
Cathodal vs. Sham	0.85	0.90	2.32***	0.82***	0.88***
Cathodal vs. Visual-only	0	1.00	0.22	0.48	0.63
Visual-only vs. The healthy	0.76	0.78	1.50***	0.75***	0.77***
Visual-only vs. Sham	0.91	0.80	2.19***	0.78***	0.87***
Sham vs. The healthy	0	1.00	0.19	0.47	0.62

AUC: area under the receiver operating characteristic curve.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

could be replicated in healthy individuals by applying cathodal tDCS stimulation to inhibit the activity of their left auditory cortex. Similarly, replacing auditory sounds with visual words in healthy individuals also replicated this conflict pattern. Moreover, we found that schizophrenia patients, healthy individuals receiving cathodal tDCS, and healthy individuals encountering conflicts only in the visual domain were more likely to be classified into the same group and separated from the untreated healthy group by trained SVM classifiers. These results suggest that the distinct cross-modal conflict pattern in schizophrenia patients may develop when the auditory cortex is insufficiently activated. In short, these findings largely support that cross-modal conflicts in schizophrenia patients may arise from abnormal activation of the auditory cortex.

The cross-modal conflict pattern reflects a flexible coping strategy in the face of inconsistent sensory information. Previous studies have shown that when schizophrenia patients experience cross-modal conflicts, visual interference with auditory targets decreases, while auditory interference with visual targets increases (de Gelder et al., 2003; Vanes et al., 2016; White et al., 2014), which is the exact opposite of healthy individuals (Donohue et al., 2013; Hutmacher, 2019). This implies that schizophrenia patients were more likely to exhibit auditory dominance in cross-modal conflict situations. In line with previous findings, schizophrenia patients in the current study also showed an auditory dominance, but only at the semantic conflict stage. Researchers have found that the temporal binding window of these patients is unusually wide, which well explains why schizophrenia patients have greater auditory dominance than healthy individuals in the double-flash illusion. In other words, schizophrenia patients illusorily perceive two flashes when the two beeps are separated by a longer time interval than healthy individuals because each beep resides in a larger temporal window and illusorily binds a flash (Ferri et al., 2018; Haß et al., 2017). However, the widened temporal binding window could not explain the auditory dominance found in the current study, as all the audiovisual stimuli were presented simultaneously rather than successively. We noticed that schizophrenia patients are typically characterized by early auditory processing deficits (Donde et al., 2020), while their abilities to process receptive speech are relatively intact (Ross et al., 2007; Weinstein et al., 2006). The relatively intact receptive speech processing may reflect a compensatory process in patients, who invest more cognitive resources in receptive speech at the expense of other cognitive activities (Ross et al., 2007), such as speech production when facing pure visual stimuli. Therefore, when audiovisual stimuli occur simultaneously, patients' relatively intact receptive speech processing prompt them to rely more on auditory stimuli, resulting in auditory dominance.

The accuracy data in the RI condition failed to demonstrate the expected modality-specific pattern (i.e., higher accuracy during attended-auditory trials). This finding cannot be adequately explained by a speed-accuracy tradeoff, as maintaining comparable accuracy between attended visual and auditory conditions would theoretically require sacrificing RT in attended visual trials, which not observed in our data. Instead, we identified a novel visual dominance at the response conflict stage, which may reflect degraded auditory cortex activation in schizophrenia patients. Previous neuroimaging evidence has indicated that auditory dominance at the response conflict stage in healthy individuals is related to enhanced neural activity in the premotor cortex (Chen and Zhou, 2013). In Chen and Zhou's study, the authors speculated that the premotor cortex may be modulated by auditory signals, based on the anatomical connection between the auditory cortex and the premotor cortex (Morillon et al., 2019). Given the role of premotor cortex in response initiation, auditory stimuli naturally dominate visual stimuli at the response stage when there are conflicts (Chen and Zhou, 2013). However, in schizophrenia patients, the neural activation in their auditory cortex is not as strong as in healthy individuals (MacDonald et al., 2020; McKinney et al., 2019), which reduces the interference on the premotor cortex, thereby attenuating the auditory dominance and even leading to visual dominance.

Although the precise reason for this double dissociation of cross-modal conflicts at the semantic and response stages in schizophrenia relative to healthy individuals remains inconclusive at present, it is noteworthy that the current study was the first to distinguish these two conflict stages. Benefiting from the cross-modal conflict paradigm, we not only revealed the innate heterogeneity of cross-modal conflicts (Aine et al., 2017; Ettinger et al., 2018; Westerhausen et al., 2011), but also uncovered its potential clinical significance. We found that the cross-modal conflict pattern at the two stages, as a single diagnostics index, is able to distinguish schizophrenia from healthy individuals with a hit rate of 0.67 and a very minor false alarm rate (0.06). We believe that incorporating the cross-modal conflict task into the clinical diagnostic indicators may effectively reduce the diagnostic errors and associated costs. Future work should collect more data from multiple hospitals or medical centers, and verify the validity of the diagnostics in an enlarged cohort.

The potential explanation for the distinct cross-modal conflict pattern in schizophrenia patients, although not conclusive, converges on a critical point: early processing in the auditory cortex may be the candidate cause of this schizophrenia-like cross-modal conflict pattern. The current study further delved into this hypothesis in Experiments 2 and 3, in which we did not directly conduct neural or behavioral modulation on the patients but instead attempted to interfere with the normal function of healthy individuals to see whether they could behave like schizophrenia patients. The reasons for adopting such an indirect approach are twofold. First, a large number of brain regions and neural circuits in schizophrenia patients show abnormalities and interweave with each other (Hanlon et al., 2016; Sanfratello et al., 2018; Straube et al., 2013; Wroblewski et al., 2020), making it difficult to isolate the role of the auditory cortex. Second, previous studies have found that the initial disease course in first-episode schizophrenia patients is accompanied by lesions in the auditory cortex (Curtis et al., 2021; Donde et al., 2023), and our regulation of the auditory cortex in healthy individuals may provide a reference for modeling the initial disease course in patients.

Our results basically simulated the distinct cross-modal conflict pattern seen in patients, in healthy individuals when they received inhibitive tDCS (the cathodal group in Experiment 2) on their auditory cortex or saw the object name in vision instead of hearing it in audition (the visual-only group in Experiment 3). Individual data also demonstrated that the similarity between the cathodal group, the visual-only group and the patients (Fig. 3a and b). However, individuals in the patient group were less concentrated than those in the cathodal group and the visual-only group, which indicates larger heterogeneity among patients (Gröhn et al., 2022). All these results suggest that degraded activation of the auditory cortex may be the key factor that distinguishes schizophrenia from healthy individuals. Previous studies have found that individuals with a certain chromosomal syndrome have severe auditory impairments (Schneider et al., 2016) and are at a much higher risk of developing schizophrenia. The lifetime prevalence of these patients experiencing auditory hallucinations is also very high (Mancini et al., 2020). However, auditory hallucinations are closely related to enhanced activity of the auditory cortex (Perez-Rando et al., 2022), which contradicts the findings of the current study. This discrepancy may be because the degraded function of the auditory cortex, a neural correlate of auditory impairment, is an early symptom of schizophrenia. The cross-modal conflict pattern in patients found in the current study may provide a clue for in-depth understanding of early diagnostic markers or therapeutic targets for schizophrenia.

Notably, there are some limitations in the current study. Firstly, while some antipsychotic medications have been shown to improve behavioral performance in patients, often producing a trend toward normalization relative to healthy controls (Mehta et al., 2019; Meltzer, 2013), this effect cannot account for the completely reversed cross-modal conflict patterns we observed between groups. However, as all patients in our study were medicated, we cannot entirely exclude the possibility of medication-related confounding effects. Secondly, our study was a small-sample experiment and did not account for potential longitudinal changes in cross-modal conflict in schizophrenia. Besides, while we confirmed that the recruited patients had no visual or auditory hallucination symptoms, we did not determine the subtypes to which they belonged. Given the substantial clinical heterogeneity among patients with schizophrenia, our findings may not generalize uniformly across all subtypes. For instance, individuals with the paranoid subtype (characterized by prominent auditory hallucinations) might exhibit heightened cross-modal conflict due to interference from hallucinatory percepts. Conversely, those with the disorganized subtype, who demonstrate severe impairments in sensory integration, could show attenuated or even absent cross-modal conflict effects. Future large-scale cohort studies with extended follow-up periods and more comprehensive patient subtype representation may provide more definitive insights. Thirdly, healthy individuals and schizophrenia patients in our study were all within the adult age range, though the patient group exhibited a higher mean age and greater variability in age distribution. While existing literature provides no direct evidence for substantial age effects on cross-modal conflict patterns within the adult range (Li et al., 2019), future studies would benefit from stricter age-matching protocols to further mitigate this potential confound. Finally, the current study observed the schizophrenia-like pattern in healthy individuals using an indirect approach. It is necessary in the future to further clarify the role of the auditory cortex directly in the patient, using electrophysiological and imaging methods, to restore its function through interventions.

In conclusion, we reported a unique and distinct cross-modal conflict pattern in schizophrenia patients compared to healthy individuals, characterized by auditory dominance at the semantic conflict stage and visual dominance at the response conflict stage. This cross-modal conflict pattern may indicate the flexible response mechanisms that schizophrenia patients employ when processing external information, and is closely related to the insufficient activation of their auditory cortex.

CRediT authorship contribution statement

Heng Zhou: Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis. **Xiangyong Yuan:** Writing – review & editing, Writing – original draft, Visualization, Validation, Funding acquisition, Formal analysis. **Pei Xie:** Resources, Investigation, Funding acquisition. **Aijun Wang:** Validation, Software, Methodology, Funding acquisition, Conceptualization. **Yi Jiang:** Writing – review & editing, Visualization, Software, Funding acquisition, Conceptualization.

Role of funding source

This research was supported by the Ministry of Science and Technology of China (STI2030-Major Projects + 2021ZD0203804), the Major Program of Philosophy and Social Sciences in Jiangsu Province (2024SJZD137), Suzhou Science and Technology Development Plan (SKY2022113). Sichuan Science and Technology Program (2024NSFSC2099), Foundation of Sichuan Research Center of Applied Psychology of Chengdu Medical College (CSXL-24218). Youth Innovation Promotion Association of the Chinese Academy of Sciences.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank Dr. Yuhui Cheng for her helpful comments to the previous version of the manuscript.

Data availability

All primary data are publicly available (https://www.scidb.cn/en/anonymous/cnlldUV6).

References

Aine, C.J., Bockholt, H.J., Bustillo, J.R., Ca~nive, J.M., Caprihan, A., Gasparovic, C., Calhoun, V.D., 2017. Multimodal neuroimaging in schizophrenia: description and dissemination. Neuroinformatics 15 (4), 343–364. https://doi.org/10.1007/s12021-017-9338-9.

Brainard, D.H., 1997. The psychophysics toolbox. Spatial Vis 10 (4), 433–436. https://doi.org/10.1163/156856897X00357.

Callan, A., Callan, D., Ando, H., 2015. An fmri study of the ventriloquism effect. Cereb. Cortex 25 (11), 4248–4258. https://doi.org/10.1093/cercor/bhu306.

Chen, Q., Zhou, X.L., 2013. Vision dominates at the preresponse level and audition dominates at the response level in cross-modal interaction: behavioral and neural evidence. J. Neurosci. 33 (17), 7109–7121. https://doi.org/10.1523/JNEUROSCI.1985-12.2013.

Chen, L., Li, Q., Song, H., Gao, R.Q., Yang, J., Dong, W.T., Dang, W.M., 2020. Classification of schizophrenia using general linear model and support vector machine via Fnirs. Phys. Eng. Sci. Med. 43 (4), 1151–1160. https://doi.org/10.1007/s13246-020-00920-0.

Cloutier, M., Aigbogun, M.S., Guerin, A., Nitulescu, R., Ramanakumar, A.V., Kamat, S.A., DeLucia, M., Duffy, R., Legacy, S.N., Henderson, C., Francois, C., Wu, E., 2016. The economic burden of schizophrenia in the United States in 2013. J. Clin. Psychiatry 77 (6), 764–771. https://doi.org/10.4088/JCP.15m10278.

Curtis, M.T., Coffman, B.A., Salisbury, D.F., 2021. Pitch and duration mismatch negativity are associated with distinct auditory cortex and inferior frontal cortex volumes in the first-episode schizophrenia spectrum. Schizophr. Bull. Open 2 (1), sgab005. https://doi.org/10.1093/schizbullopen/sgab005.

de Gelder, B., Vroomen, J., Annen, L., Masthof, E., Hodiamont, P., 2003. Audio-visual integration in schizophrenia. Schizophr. Res. 59 (2–3), 211–218. https://doi.org/10.1016/S0920-9964(01)00344-9.

de Jong, J.J., Hodiamont, P.P., Van den Stock, J., de Gelr, B., 2009. Audiovisual emotion recognition in schizophrenia: reduced integration of facial and vocal affect. Schizophr. Res. 107 (2–3), 286–293. https://doi.org/10.1016/j.schres.2008.10.001.

Donde, C., Avissar, M., Weber, M.M., Javitt, D.C., 2019. A century of sensory processing dysfunction in schizophrenia. Eur. Psychiat. 59, 77–79. https://doi.org/10.1016/j.eurpsy.2019.04.006.

Donde, C., Brunelin, J., Haesebaert, F., 2020. Duration, pitch and intensity features reveal different magnitudes of tone-matching deficit in schizophrenia. Schizophr. Res. 215, 460–462. https://doi.org/10.1016/j.schres.2019.10.003.

Donde, C., Kantrowitz, J.T., Medalia, A., Saperstein, A.M., Balla, A., Sehatpour, P., Martinez, A., O'Connell, M.N., Javitt, D.C., 2023. Early auditory processing dysfunction in schizophrenia: mechanisms and implications. Neurosci. Biobehav. Rev. 148, 105098. https://doi.org/10.1016/j.neubiorev.2023.105098.

Donohue, S.E., Todisco, A.E., Woldorff, M.G., 2013. The rapid distraction of attentional resources toward the source of incongruent stimulus input during multisensory conflict. J. Cogn. Neurosci. 25 (4), 623–635. https://doi.org/10.1162/jocn_a_00336.

Ettinger, U., Aichert, D.S., Wöstmann, N., Dehning, S., Riedel, M., Kumari, V., 2018. Response inhibition and interference control: effects of schizophrenia, genetic risk, and schizotypy. J. Neuropsychol. 12 (3), 484–510. https://doi.org/10.1111/jnp.12126.

Ferri, F., Venskus, A., Fotia, F., Cooke, J., Romei, V., 2018. Higher proneness to multisensory illusions is driven by reduced temporal sensitivity in people with high schizotypal traits. Conscious. Cogn. 65, 263–270. https://doi.org/10.1016/j.concog.2018.09.006.

Fung, G.M., Mangasarian, O.L., 2005. Multicategory proximal support vector machine classifiers. Mach. Learn. 59 (1–2), 77–97. https://doi.org/10.1007/s10994-005-0463-6.

Gröhn, G., Norgren, E., Eriksson, L., 2022. A systematic review of the neural correlates of multisensory integration in schizophrenia. Schizophr. Res-Cogn. 27, 100219. https://doi.org/10.1016/j.scog.2021.100219.

Haist, F., Song, A.W., Wild, K., Faber, T.L., Popp, C.A., Morris, R.D., 2001. Linking sight and sound: fMRI evidence of primary auditory cortex activation during visual word recognition. Brain Lang. 76, 340–350. https://doi.org/10.1006/brln.2000.2433.

Hanlon, F.M., Shaff, N.A., Dodd, A.B., Ling, J.M., Bustillo, J.R., Abbott, C.C., Stromberg, S.F., Abrams, S., Lin, D.S., Mayer, A.R., 2016. Hemodynamic response function abnormalities in schizophrenia during a multisensory detection task. Hum. Brain Mapp. 37 (2), 745–755. https://doi.org/10.1002/hbm.23063.

Haß, K., Sinke, C., Reese, T., Roy, M., Wiswede, D., Dillo, W., Szycik, G.R., Oranje, B., 2017. Enlarged temporal integration window in schizophrenia indicated by the double-flash illusion. Cogn. Neuropsychiatry 22 (2), 145–158. https://doi.org/10.1080/13546805.2017.1287693.

Hirst, R.J., Cragg, L., Allen, H.A., 2018. Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect. Neurosci. Biobehav. Rev. 94, 286–301. https://doi.org/10.1016/j.neubiorev.2018.07.012.

Hutmacher, F., 2019. Why is there so much more research on vision than on any other sensory modality? Front. Psychol. 10, 2246. https://doi.org/10.3389/fpsyg.2019.02246.

Koenig, T., van Swam, C., Dierks, T., Hubl, D., 2012. Is gamma band EEG synchronization reduced during auditory driving in schizophrenia patients with

auditory verbal hallucinations? Schizophr. Res. 141 (2–3), 266–270. https://doi.org/10.1016/j.schres.2012.07.016.
Li, Z., Gu, R., Qi, M., Cen, J., Zhang, S., Gu, J., Zeng, X., Chen, Q., 2019. Loss of vision dominance at the preresponse level in tinnitus patients: preliminary behavioral evidence. Front. Neurosci. 13, 482. https://doi.org/10.3389/fnins.2019.00482.
MacDonald, M.L., Garver, M., Newman, J., Sun, Z., Kannarkat, J., Salisbury, R., Glausier, J., Ding, Y., Lewis, D.A., Yates, N., Sweet, R.A., 2020. Synaptic proteome alterations in the primary auditory cortex of individuals with schizophrenia. JAMA Psychiatry 77 (1), 86–95. https://doi.org/10.1001/jamapsychiatry.2019.2974.
Mancini, V., Zöller, D., Schneider, M., Schaer, M., 2020. Abnormal development and dysconnectivity of distinct thalamic nuclei in patients with 22q11.2 deletion syndrome experiencing auditory hallucinations. Biol. Psychiat.-Cogn. Neurosci. Neuroimag. 5 (9), 875–890. https://doi.org/10.1016/j.bpsc.2020.04.015.
McKinney, B.C., MacDonald, M.L., Newman, J.T., Shelton, M.A., DeGiosio, R.A., Kelly, R. M., Fish, K.N., Sampson, A.R., Lewis, D.A., Sweet, R.A., 2019. Density of small dendritic spines and microtubule-associated-protein-2 immunoreactivity in the primary auditory cortex of subjects with schizophrenia. Neuropsychopharmacology 44 (6), 1055–1061. https://doi.org/10.1038/s41386-019-0350-7.
Mehta, N.D., Won, M.J., Babin, S.L., Patel, S.S., Wassef, A.A., Chuang, A.Z., Sereno, A.B., 2019. Differential benefits of olanzapine on executive function in schizophrenia patients: Preliminary findings. Hum. Psychopharmacol.-Clin. Exp. 35, 1. https://doi.org/10.1002/hup.2718.
Meltzer, H.Y., 2013. Update on typical and atypical antipsychotic drugs. Annu. Rev. Med. 64, 393–406. https://doi.org/10.1146/annurev-med-050911-161504.
Morillon, B., Arnal, L.H., Schroeder, C.E., Keitel, A., 2019. Prominence of delta oscillatory rhythms in the motor cortex and their relevance for auditory and speech perception. Neurosci. Biobehav. Rev. 107, 136–142. https://doi.org/10.1016/j.neubiorev.2019.09.012.
Moyer, C.E., Delevich, K.M., Fish, K.N., Asafu-Adjei, J.K., Sampson, A.R., Dorph-Petersen, K.A., Lewis, D.A., Sweet, R.A., 2012. Reduced glutamate decarboxylase 65 protein within primary auditory cortex inhibitory boutons in schizophrenia. Biol. Psychiatry 72 (9), 734–743. https://doi.org/10.1016/j.biopsych.2012.04.010.
Pearl, D., Yodashkin-Porat, D., Katz, N., Valevski, A., Aizenberg, D., Sigler, M., Weizman, A., Kikinzon, L., 2009. Differences in audiovisual integration, as measured by McGurk phenomenon, among adult and adolescent patients with schizophrenia and age-matched healthy control groups. Compr. Psychiatry 50 (2), 186–192. https://doi.org/10.1016/j.comppsych.2008.06.004.
Perez-Rando, M., Elvira, U.K.A., Garcia-Marti, G., Gadea, M., Aguilar, E.J., Escarti, M.J., Ahullo-Fuster, M.A., Grasa, E., Corripio, I., Sanjuan, J., Nacher, J., 2022. Alterations in the volume of thalamic nuclei in patients with schizophrenia and persistent auditory hallucinations. NeuroImage-Clin. 35, 103070. https://doi.org/10.1016/j.nicl.2022.103070.
Ross, L.A., Saint-Amour, D., Leavitt, V.M., Molholm, S., Javitt, D.C., Foxe, J.J., 2007. Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. Schizophr. Res. 97 (1–3), 173–183. https://doi.org/10.1016/j.schres.2007.08.008.
Sanfratello, L., Aine, C., Stephen, J., 2018. Neuroimaging investigations of dorsal stream processing and effects of stimulus synchrony in schizophrenia. Psychiatry Res. Neuroimaging 278, 56–64. https://doi.org/10.1016/j.pscychresns.2018.05.005.
Sass, K., Heim, S., Sachs, O., Straube, B., Schneider, F., Habel, U., Kircher, T., 2013. Neural correlates of semantic associations in patients with schizophrenia. Eur. Arch. Psych. Clin. Neurosci. 264 (2), 143–154. https://doi.org/10.1007/s00406-013-0425-0.
Schneider, M., Armando, M., Pontillo, M., Vicari, S., Debbane, M., Schultze-Lutter, F., Eliez, S., 2016. Ultra high risk status and transition to psychosis in 22q11.2 deletion syndrome. World Psychiatry 15 (3), 259–265. https://doi.org/10.1002/wps.20347.
Snodgrass, J.G., Vanderwart, M., 1980. A standardized set of 260 picture: norms for name agreement, image agreement, familiarity, and visual complexity. J. Exp. Psychol. Learn. Mem. Cogn. 6 (2), 174–215. https://doi.org/10.1037/0278-7393.6.2.174.
Stein, B.E., Stanford, T.R., 2008. Multisensory integration: current issues from the perspective of the single neuron. Nat. Rev. Neurosci. 9 (4), 255–266. https://doi.org/10.1038/nrn2331.
Straube, B., Green, A., Sass, K., Kirner-Veselinovic, A., Kircher, T., 2013. Neural integration of speech and gesture in schizophrenia: evidence for differential processing of metaphoric gestures. Hum. Brain Mapp. 34 (7), 1696–1712. https://doi.org/10.1002/hbm.22015.
Surguladze, S., Rossell, S., Rabe-Hesketh, S., David, A.S., 2002. Cross-modal semantic priming in schizophrenia. J. Int. Neuropsychol. Soc. 8 (7), 884–892. https://doi.org/10.1017/S1355617702870023.
Tang, X.Y., Wu, J.L., Shen, Y., 2016. The interactions of multisensory integration with endogenous and exogenous attention. Neurosci. Biobehav. Rev. 61, 208–224. https://doi.org/10.1016/j.neubiorev.2015.11.002.
Vanes, L.D., White, T.P., Wigton, R.L., Joyce, D., Collier, T., Shergill, S.S., 2016. Reduced susceptibility to the sound-induced flash fusion illusion in schizophrenia. Psychiatry Res. 245, 58–65. https://doi.org/10.1016/j.psychres.2016.08.016.
Wallace, M.T., Stevenson, R.A., 2014. The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. Neuropsychologia 64, 105–123. https://doi.org/10.1016/j.neuropsychologia.2014.08.005.
Wallace, M.T., Woynaroski, T.G., Stevenson, R.A., 2020. Multisensory integration as a window into orderly and disrupted cognition and communication. Annu. Rev. Psychol. 71, 193–219. https://doi.org/10.1146/annurev-psych-010419-051112.
Weinstein, S., Werker, J.F., Vouloumanos, A., Woodward, T.S., Ngan, E.T., 2006. Do you hear what I hear? Neural correlates of thought disorder during listening to speech in schizophrenia. Schizophr. Res. 86 (1–3), 130–137. https://doi.org/10.1016/j.schres.2006.05.011.
Westerhausen, R., Kompus, K., Hugdahl, K., 2011. Impaired cognitive inhibition in schizophrenia: A meta-analysis of the Stroop interference effect. Schizophr. Res. 133, 172–181. https://doi.org/10.1016/j.schres.2011.08.025.
White, T.P., Wigton, R.L., Joyce, D.W., Bobin, T., Ferragamo, C., Wasim, N., Lisk, S., Shergill, S.S., 2014. Eluding the illusion? Schizophrenia, dopamine and the McGurk effect. Front. Hum. Neurosci. 8, 565. https://doi.org/10.3389/fnhum.2014.00565.
Wroblewski, A., He, Y., Straube, 2020. Dynamic causal modelling suggests impaired effective connectivity in patients with schizophrenia spectrum disorders during gesture-speech integration. Schizophr. Res. 216, 175–183. https://doi.org/10.1016/j.schres.2019.12.005.
Xu, H.H., Yang, G.C., Wu, H.Y., Xiao, J., Li, Q., Liu, X., 2024. Distinct mechanisms underlying cross-modal semantic conflict and response conflict processing. Cereb. Cortex 34 (2), 1–10. https://doi.org/10.1093/cercor/bhad539.
Zhou, W., Jiang, Y., He, S., Chen, D., 2010. Olfaction modulates visual perception in binocular rivalry. Curr. Biol. 20 (15), 1356–1358. https://doi.org/10.1016/j.cub.2010.05.059.
Zhuo, C.J., Tian, H.J., Fang, T., Li, R.L., Li, Y.C., Kong, L.G., Cai, Z.Y., Zheng, L.D., Lin, X. D., Chen, C., 2020. Neural mechanisms underlying visual and auditory processing impairments in schizophrenia: insight into the etiology and implications for tailoring preventive and therapeutic interventions. Am. J. Transl. Res. 12 (12), 7657–7669.