Original Research

Music Training Enhances Auditory Inhibitory Control but Not Visual

Ying Liu[1],[†] ⓘ, Jing Ning[1],[†] ⓘ, Jiarui Ma[1] ⓘ, Jiejia Chen[2],* ⓘ

[1]College of Music, Southwest University, 400715 Chongqing, China
[2]School of Electronic and Information Engineering, Southwest University, 400715 Chongqing, China
*Correspondence: chenlele1993@email.swu.edu.cn (Jiejia Chen)
[†]These authors contributed equally.

Abstract

Background: Inhibitory control is an important component of cognitive processing that is influenced by multimodal information processing. Recent research has mainly focused on the influence of visual information on inhibitory control, paying less attention to the impact of auditory training, which limits the exploration of the mechanism and practical training of inhibitory control. The influence of music training on inhibitory control has received considerable attention in recent years. To explore the modality-specificity of inhibitory control, this study compares the behavioral and brain electrical activities of different music training experiences through visual and auditory inhibitory control tasks. **Methods**: This investigation utilized event-related potential (ERP) and time-frequency analysis methodologies to examine the behavioral and neural patterns of thirty participants with musical expertise alongside thirty individuals without such training while completing both visual and auditory Stroop tasks. Further analysis was conducted to examine the modality-specific effect of music training on inhibitory control. **Results**: The results showed no significant group differences in behavioral performance or traditional ERP components (N450 and sustained potential (SP)) in either modality. Time-frequency analysis revealed no significant differences in theta power in the visual modality. However, in the auditory modality, the music training group exhibited significantly lower beta power, suggesting that music training may more efficiently recruit neural resources when resolving auditory conflict. **Conclusions**: These findings indicate that inhibitory control exhibits modality specificity across both visual and auditory modalities. Specifically, music training mainly improves auditory conflict resolution, suggesting its impact on inhibitory control is limited to specific sensory modalities.

Keywords: music training; inhibitory control; conflict resolution; time-frequency analysis

1. Introduction

Inhibitory control is a core component of executive function and is defined as an individual's ability to suppress impulses or distracting information unrelated to a goal or task by regulating attention, behavior, thoughts, or emotions, thereby enabling goal-directed action [1]. This capacity not only underpins higher-order cognitive processes such as attentional control, conflict processing, and decision-making, it is also closely linked to a range of psychopathologies [2,3]. Consequently, enhancing inhibitory control has become a focal concern in both psychology and clinical medicine. In the current era of increasingly abundant online information, effectively improving inhibition control of redundant stimuli has become the focus of cognitive development. Previous research demonstrates enhanced inhibitory control via pharmacological interventions, cognitive training, and exercise programs [4–7]. Pharmacological interventions (e.g., methylphenidate, which is commonly used to improve attention deficit hyperactivity disorder) primarily enhance inhibitory control by modulating neurotransmitter levels; however, they often have side effects and may increase long-term treatment costs [8,9]. Cognitive training typically involves task-based repetitive practice, such as the Go/No-Go, Stroop, or working memory tasks, which strengthen executive func-

tions. However, the monotonous nature of such training often leads to participant fatigue, thereby reducing intervention effectiveness [1,10]. Exercise-based interventions have shown positive effects on inhibitory control; however, the magnitude of these effects varies considerably across populations and the evidence remains inconclusive [11,12]. Although these approaches have demonstrated some effectiveness in improving inhibitory control, they have notable limitations. Therefore, interventions that are both engaging and effective must be identified. Music training offers distinct and comprehensive advantages through rich cognitive stimulation and effective artistic engagement [13].

Music training, as a typical multisensory-channel and multilevel cognitive processing activity, constitutes a multi-dimensional stimulus involving motor, somatosensory, auditory, visual, and spatial processing, as well as localization and the integration of these diverse forms of processing [14,15], which has a particularly positive effect on inhibitory control ability [16]. According to the "Audiation" theory proposed by Edwin Gordon, the comprehension and processing of music depends on the brain's ability to encode, maintain, and retrieve rhythmic and pitch patterns, a process that requires sustained cognitive engagement. To process these auditory patterns accurately, individuals must continuously suppress irrelevant and distracting stim-

uli, thereby maintaining selective attention to the target auditory information [17]. Therefore, repeatedly activating auditory processing and regulatory mechanisms during music training may effectively enhance long-term inhibitory control capacity. This advantage is closely related to the operational demands of instrumental performance. Musicians need to continuously monitor and rapidly adjust their playing actions based on auditory and proprioceptive feedback while simultaneously suppressing irrelevant distractions, indicating that this highly refined motor process in music training fundamentally requires substantial involvement in inhibitory control ability [10,18].

Although music training relies heavily on auditory processing, most previous studies have focused on inhibitory control within the visual modality, paying limited attention to auditory channels. Even in relation to inhibitory control in music training [18], traditional research may neglect the inherent auditory-centered nature of music training. Some studies have found that professional musicians demonstrate significantly smaller interference effects in Stroop tasks than amateur musicians, reflecting their more efficient inhibitory control abilities [19]. The reduced interference observed in the professional group underscores the notion that music training, characterized by its highly refined integration of motor and cognitive demands, inherently necessitates the continuous engagement of inhibitory control mechanisms. However, assessing music training's impact on inhibitory control via only visual stimuli may fail to represent music's unique processing channels. In fact, the neural processing of inhibitory control exhibits significant modality-specific differences. Auditory conflict tasks are more closely aligned with the actual processing demands of musical training than visual tasks [20]. Previous research has demonstrated that musicians outperform non-musicians in auditory conflict tasks across behavioral measures, event-related potential (ERP) components, and time–frequency features, reflecting enhanced neural processing efficiency [21,22], which may be closely linked to the auditory-based nature of music training. Compared with visual tasks, auditory tasks can better simulate the real cognitive demands imposed by music training. Therefore, a unimodal visual perspective is insufficient to completely elucidate the impact of music training on the inhibitory control (IC), and it is necessary to investigate its neural processing characteristics in a multimodal context.

Based on this, the present study employed the Stroop paradigm, utilizing both visual and auditory channels, combined with ERP and time-frequency analyses, to compare differences in inhibitory control tasks between university students with long-term systematic music training (music training group) and non-music majors (control group) to investigate the differences in the neural processing of inhibitory control across different perceptual modalities. Furthermore, a particular focus should be on whether the effects of musical training on inhibitory control are modality-specific or modality-general. The Stroop task is a classi-cal method used to investigate the neural mechanisms of conflict processing. It consists of congruent (where semantic meaning and word color match) and incongruent (where semantic meaning and word color do not match) conditions [23]. Regarding ERPs, the Stroop task elicited prominent conflict-related N450/Ninc and late sustained potential (SP) components [24]. Specifically, N450 typically emerges approximately 300–500 ms after the presentation of visual stimuli, whereas Ninc is observed around 300–400 ms after auditory stimuli [25]. The SP component, appearing after N450/Ninc around 500–800 ms post-stimulus, is associated with the ability to resolve conflicts [26]. Regarding time-frequency, previous studies have shown that Stroop tasks can elicit significant theta- and beta-band activity. Theta power enhancement is widely considered an index of conflict monitoring and cognitive control and is closely associated with attentional regulation and error monitoring during tasks [27,28]. Beta-band activity, on the other hand, is closely related to conflict resolution or processing [29–31]. This study examined both frequency bands to explore the modality specificity of inhibitory control by comparing the behavioral and brain electrical activities of different music training experiences through visual and auditory inhibitory control tasks.

2. Materials and Methods

2.1 Participants

To minimize potential confounding factors, this study recruited 32 university students majoring in music and 32 university students from non-music majors. Each subject possessed unimpaired hearing, either normal or corrected vision, right-handed dominance, and no documented history of psychological conditions. We eliminated four participants from the analysis due to ERP artifact trials surpassing the 25% threshold, ultimately including 30 individuals in each experimental group (music training group: age 20.76 ± 1.19 years, eight males and 22 females; control group: age 21.16 ± 1.05 years, nine males and 21 females; details provided in **Supplementary Fig. 1**). Participants in the music training group were required to be proficient in playing at least one instrument (excluding those who had exclusively received vocal training or had studied music theory). All members had received training in playing an instrument for a minimum of eight years and were still actively practicing at the time of the study, with an average training duration of 12.53 ± 2.58 years and an average starting age of 6.33 ± 2.64 years. Participants in the control group were required to have received no musical training outside of standard school music courses. All participants completed a demographic questionnaire, including name, sex, and age, and completed the shortened Raven's Advanced Progressive Matrices assessment [32]. There were no significant differences between the two groups in age, $t(58) = -1.43$, $p = 0.15$, or intelligence scores, $t(58) = 0.99$, $p = 0.32$.

2.2 Experimental Materials and Procedure

The visual Stroop task materials consisted of Chinese characters "红" ("red") and "蓝" ("blue") displayed in either red or blue font, yielding a total of four stimulus combinations. Among them, two were congruent stimuli (e.g., the character "红" in red), and two were incongruent (e.g., the character "蓝" in blue). All stimuli were presented in a uniform font size against a gray background. Participants were instructed to ignore the semantic meaning of the characters and respond only to the font color. That is, 'press the "F" key if the color matched the word's meaning, and press the "J" key if it did not'. Following a practice session, participants undertook an assessment split into two segments, with a brief intermission between them.

The auditory Stroop task materials consisted of spoken words with the semantic content "big" or "small", each presented at two volume levels: 50 decibel (dB) (representing low volume) and 70 dB (representing high volume), forming congruent or incongruent combinations between loudness and meaning, for a total of four stimuli. Among these, two were congruent stimuli (e.g., the word "small" presented at 50 dB), and two were incongruent (e.g., the word "small" presented at 70 dB). All stimuli were presented through professional closed-back headphones (Sennheiser HD 280 Pro; frequency response 8–25,000 Hz; noise isolation up to 32 dB, Sennheiser electronic GmbH & Co. KG, Wedemark, Lower Saxony, Germany) to ensure consistent and high-quality audio. The experiment was conducted in a sound-attenuated electrophysiology laboratory, where background noise was maintained below 25 dB sound pressure level (SPL). To ensure that the subjective loudness differences across conditions were perceptually clear, all stimuli were pre-calibrated for SPL and verified using equal-loudness contours, according to ISO 226:2023 [33]. Participants were instructed to judge the loudness level of the voice while ignoring its semantic content. That is, 'press the "J" key for soft sounds (50 dB), and press the "F" key for loud sounds (70 dB)'. Before the formal task, participants listened to sample stimuli of different loudness levels to establish a reference for distinguishing between "soft" and "loud" volumes. Following a practice session, participants undertook an assessment split into two segments, with a brief intermission between them.

Each trial followed the same sequence: first, a fixation point "+" was presented for 500 ms; this was followed by a blank screen lasting 200–500 ms. Then, a visual or auditory stimulus was presented for 300 ms, with a response window of 1000 ms. Finally, a blank screen was displayed for 1000–1500 ms as the inter-trial interval. The procedures were identical for the visual and auditory tasks.

For behavioral measures, the Stroop interference effect was used to assess participants' conflict control ability, which was determined by contrasting performance metrics between matching and non-matching test conditions. This measurement captures the contrast in both precision and speed when participants encounter congruent versus incongruent stimuli. Notably, a diminished Stroop interference effect reflects an enhanced aptitude for regulating conflicting information.

2.3 Data Acquisition and Analysis

Electroencephalogram (EEG) data were recorded using a 128-channel BioSemi ActiveTwo system (Active3, BioSemi, Heerlen, Netherlands). The EEG signals were sampled online at 1024 Hz, with all electrode impedances kept below 5 kΩ. A band-pass filter ranging from 0.1–100 Hz was applied during recording. Participants were instructed to minimize their head and body movements throughout the experiment to reduce electromyographic artifacts. The raw EEG data were stored in BioSemi Data Format (BDF) format and re-referenced offline to the averaged signal of the bilateral mastoids. Offline pre-processing was conducted using the EEGLAB toolbox (Active3, BioSemi, Heerlen, The Netherlands; https://www.biosemi.com/). After manually removing bad segments, the data were band-pass filtered from 0.1–30 Hz. Independent component analysis (ICA) was performed to correct for ocular artifacts, and signals surpassing ± 80 µV post-artifact filtering were omitted. ICA was subsequently performed to correct for ocular artifacts. Components corresponding to blinks and eye movements were identified based on their spatial distribution and temporal characteristics and then removed. Upon completing the ICA analysis, any experiments where the amplitudes at any electrode topped the ± 80 µV mark were ruled out for subsequent scrutiny. In the ERP analysis, the data was divided into 1000-millisecond windows starting from the stimulus presentation, while the 200-millisecond window prior to the stimulus served as the baseline for adjustment. ERP waveforms were averaged by experimental condition, and only correct trials were included. Finally, the grand-averaged ERPs for each condition were calculated for each participant.

Time-frequency amplitude and phase were calculated via a short-time Fourier transform using Discrete Fourier Transform (DFT). Specifically, EEG signals were decomposed in a time window spanning 800 ms before stimulus onset to 1000 ms after stimulus onset using a sliding window with a step size of 0.4 s. Frequency resolution was set to 1 Hz, covering the 3–30 Hz band. The resulting time-frequency power spectra were then baseline-corrected, with the baseline interval defined as –800 to –200 ms prior to stimulus onset. Baseline correction was performed using the subtraction method, whereby the power at each time point was adjusted by subtracting the mean power of the baseline interval, thereby eliminating nonspecific power shifts.

Based on previous studies, observations, and analyses of grand-average waveforms and time–frequency maps [25,34], the following time windows and electrode sites were selected for the ERP and time-frequency analyses. In

(a)

(b)
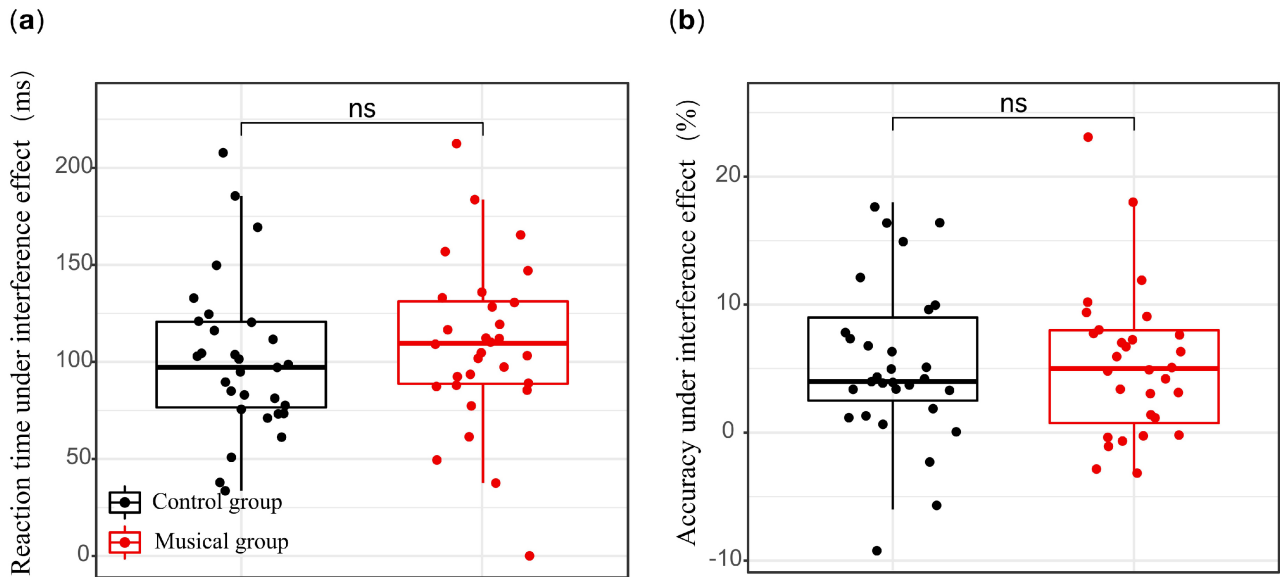
Visual Stroop task

Fig. 1. Behavioral performance of the music training group and control group in the visual modality. (a) Stroop interference effect in reaction time; (b) Stroop interference effect in accuracy. ns, not significant ($p > 0.05$).


(a)

(b)
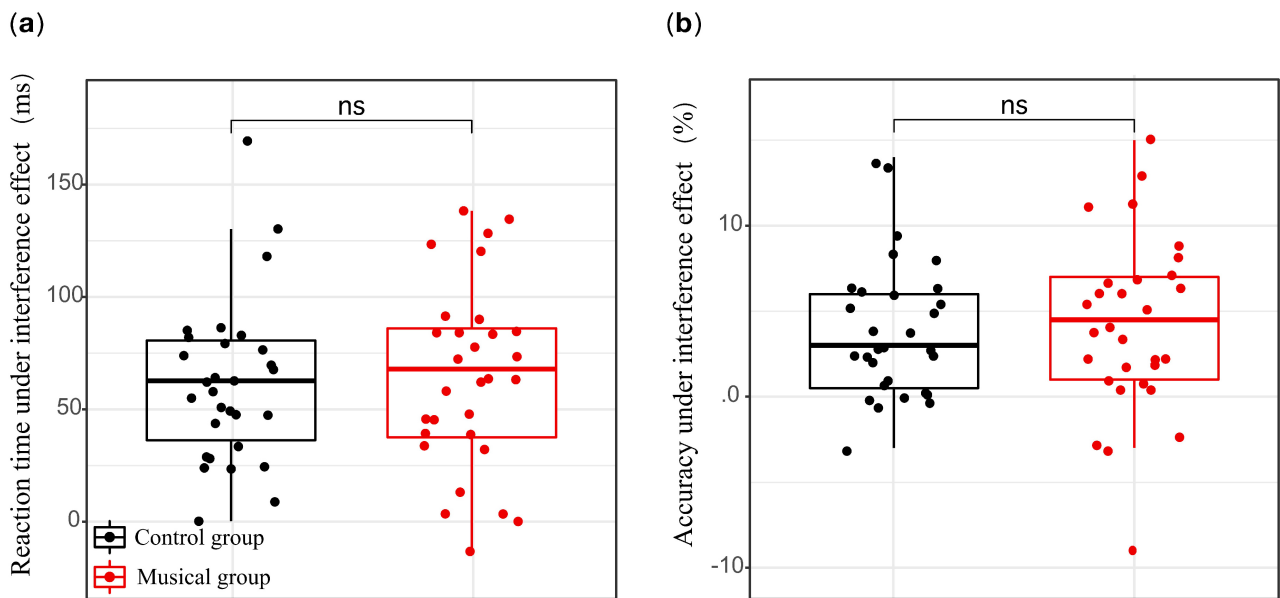
Auditory Stroop task

Fig. 2. Behavioral performance of the music training group and the control group in the auditory modality. (a) Stroop interference effect in reaction time; (b) Stroop interference effect in accuracy. ns, not significant ($p > 0.05$).

ERP analysis, the mean amplitudes of the N450 and late SP components were examined. In the visual modality, the N450 component's time window spanned 400–500 ms, with these electrode sites chosen: CZ, C1, C2, CPZ, CP1, CP2, Pz, P1, P2, FZ, F1, F2, FCz, FC1, and FC2. The SP component was analyzed in the 600–800 ms time window at the same electrode sites. In the auditory modality, the Ninc component was analyzed in the 300–400 ms time window at CPZ, P1, P3, Pz, P2, CP2, P4, and CP1. The SP component in the auditory modality was analyzed in the
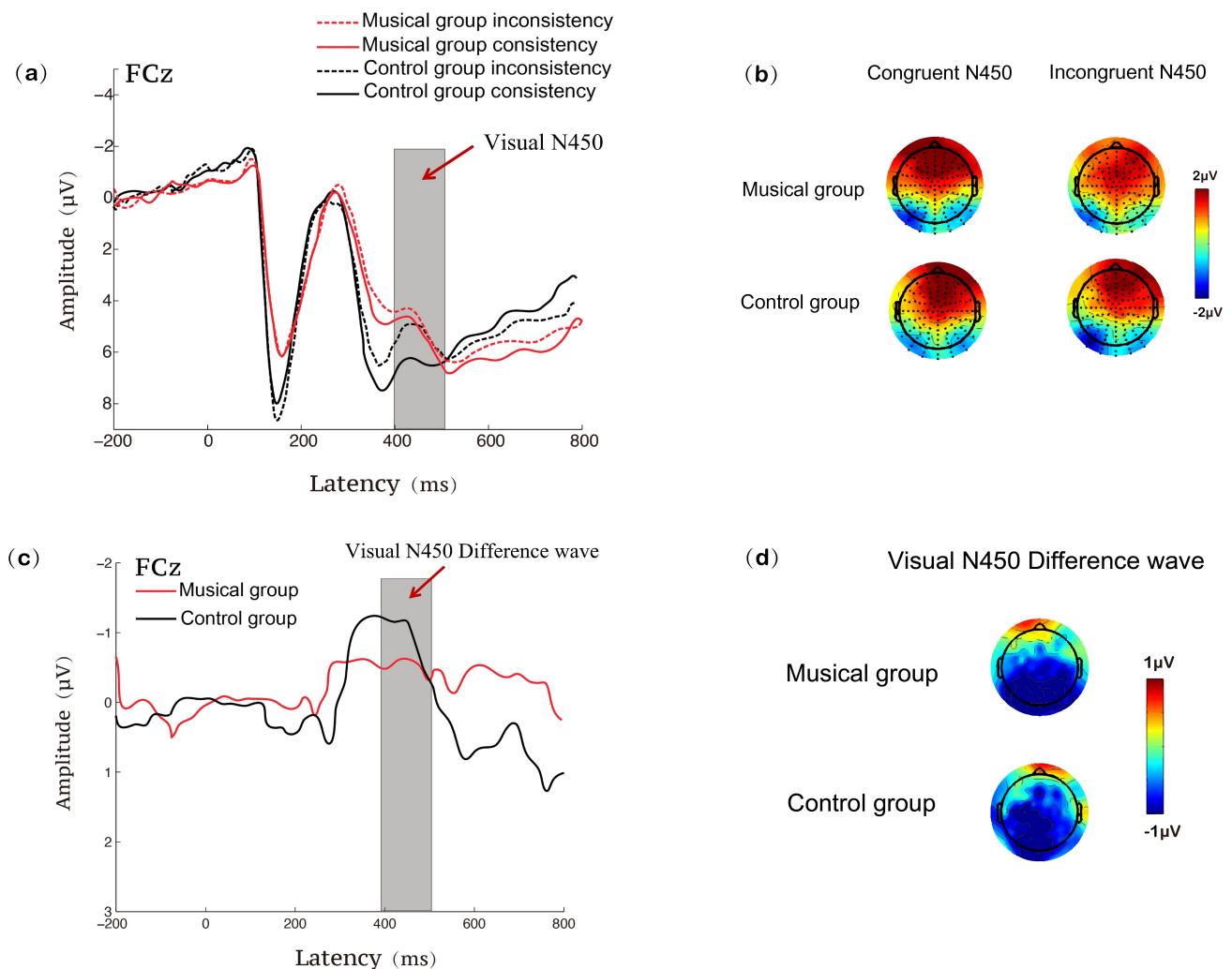
Fig. 3. ERP results at the FCz site during the visual Stroop task for the music training and control groups. (a) N450 waveforms under congruent and incongruent conditions; (b) N450 topographies under congruent and incongruent conditions; (c) N450 difference waveforms (incongruent minus congruent); (d) N450 difference topographies (incongruent minus congruent). ERP, event-related potential.

500–700 ms window at PO7, Pz, POZ, OZ, and PO8. In the time-frequency analysis, the average power in the theta (4–7 Hz) and beta (21–30 Hz) frequency bands was computed. Theta-band activity was analyzed at the following electrodes: FZ, F1, F2, FCz, FC1, FC2, CZ, C1, and C2. Beta activity was analyzed at CZ, C1, C2, FCz, FC1, and FC2.

2.4 Statistical Analysis

SPSS 27.0 (IBM Corporation, Armonk, NY, USA) was utilized for the statistical evaluations. For the behavioral data from the Stroop task (accuracy and reaction time) and the EEG data, including ERP components (mean amplitudes of N450 and SP) and time-frequency data (mean power in the theta and beta bands), separate two-way repeated-measures ANOVAs were performed with Condition (congruent, incongruent) as the within-subject factor and Group (music training, control) as the between-subject

factor. For the time–frequency analysis, power values were converted to dB units relative to the pre-stimulus baseline to represent changes in oscillatory activity. When the assumption of sphericity was violated, p-values were corrected using the Greenhouse-Geisser method.

3. Results

3.1 Behavioral Results

In the visual modality, ANOVA on accuracy revealed a significant main effect of condition, $F(1, 58) = 36.332$, $p < 0.001$, $\eta p^2 = 0.385$, with greater accuracy in the congruent condition than in the incongruent condition. The main effect of group was not significant, $F(1, 58) = 0.018$, $p = 0.893$, $\eta p^2 < 0.001$, indicating no discernible difference in accuracy between the music training and the control groups. The interaction between condition and group was also not significant, $F(1, 58) = 1.016$, $p = 0.318$, $\eta p^2 = 0.017$. An independent samples t-test further showed no significant
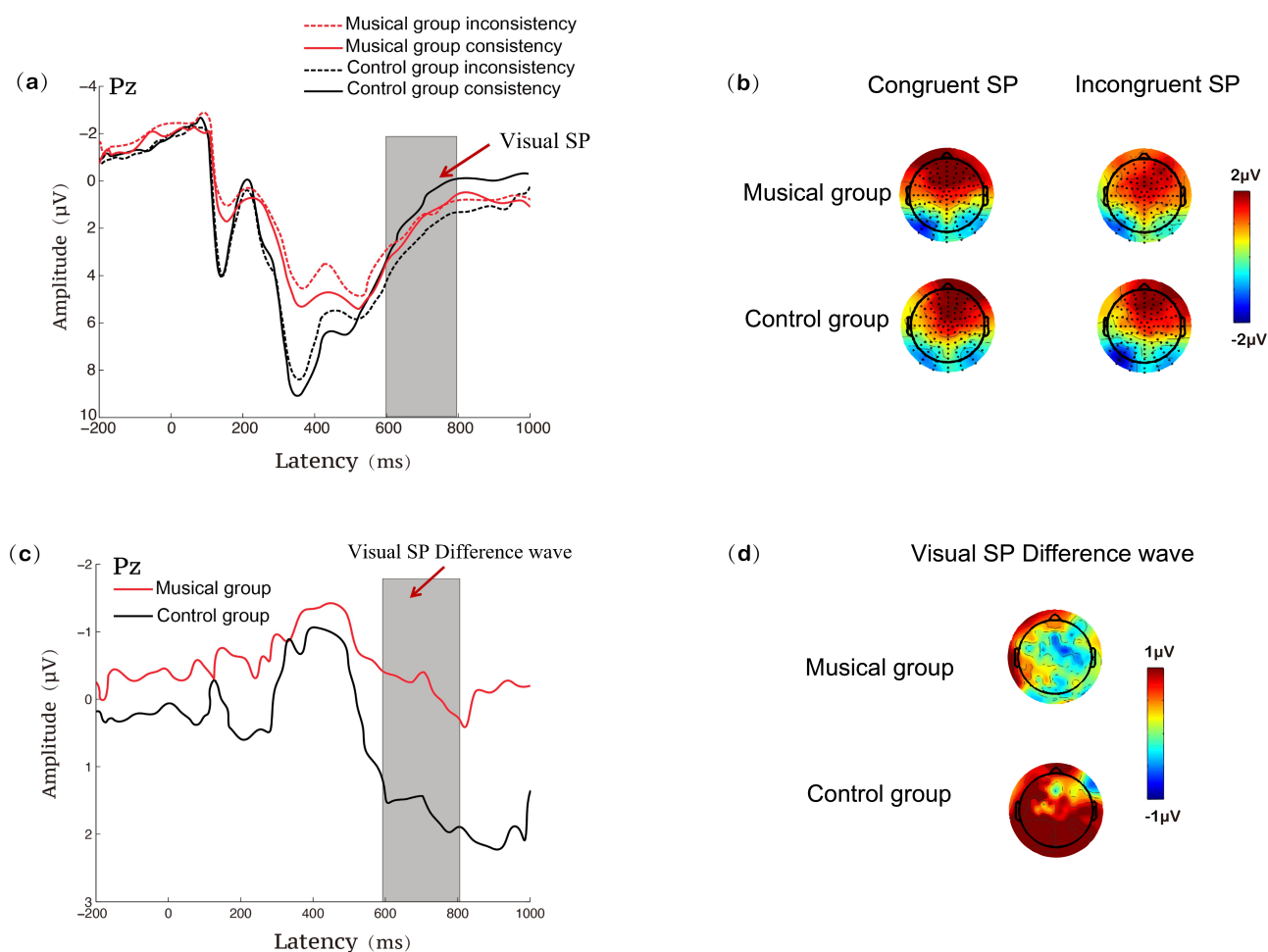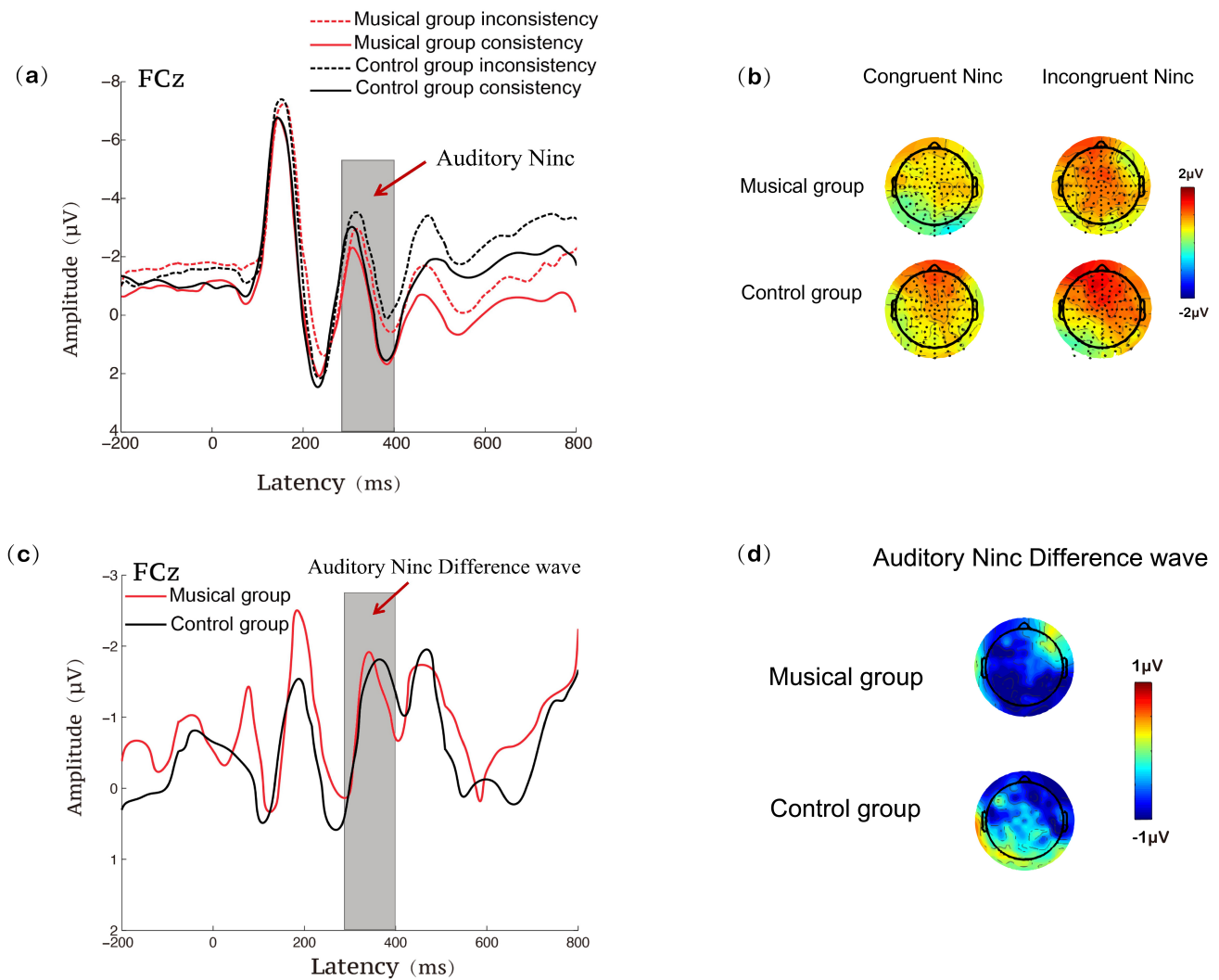
Fig. 4. ERP results at the Pz site during the visual Stroop task for the music training and control groups. (a) SP waveforms under congruent and incongruent conditions; (b) SP topographies under congruent and incongruent conditions; (c) SP difference waveforms (incongruent minus congruent); (d) SP difference topographies (incongruent minus congruent). SP, sustained potential.

group difference in the Stroop interference effect (incongruent minus congruent accuracy), $t(58) = 1.008$, $p = 0.318$, Cohen's $d = 0.260$, 95% CI [–0.016, 0.048]. ANOVA of reaction time revealed a significant main effect of condition, $F(1, 58) = 312.017$, $p < 0.001$, $\eta p^2 = 0.843$, with faster reaction times in the congruent condition than in the incongruent condition. Group differences were not statistically significant, $F(1, 58) = 0.213$, $p = 0.646$, $\eta p^2 = 0.004$, suggesting no significant difference in reaction times between the music training and control groups. The condition-group interaction was nonsignificant, $F(1, 58) = 0.011$, $p = 0.916$, $\eta p^2 < 0.001$. An independent samples t-test indicated no significant group difference in the Stroop interference effect (incongruent minus congruent reaction time), $t(58) = -0.106$, $p = 0.916$, Cohen's $d = -0.027$, 95% CI [–1.220, 11.483] (see Fig. 1a,b).

In the auditory modality, ANOVA on accuracy revealed a significant main effect of condition, $F(1, 58) = 48.906$, $p < 0.001$, $\eta p^2 = 0.457$, with higher accuracy in the congruent condition than in the incongruent condition. The main effect of group was not significant, $F(1, 58) =$

0.212, $p = 0.647$, $\eta p^2 = 0.004$, indicating no significant difference in accuracy between the music-trained group and control group. The interaction between group and condition was also not significant, $F(1, 58) = 0.594$, $p = 0.444$, $\eta p^2 = 0.010$. An independent samples t-test revealed no significant difference in the Stroop interference effect (accuracy in incongruent trials minus accuracy in congruent trials) between the music training and the control groups. $t(58) = -0.771$, $p = 0.444$, Cohen's $d = -0.199$, 95% CI [–0.032, 0.014]. For reaction time in the auditory modality, ANOVA revealed a significant main effect of condition, $F(1, 58) = 170.848$, $p < 0.001$, $\eta p^2 = 0.747$, with significantly markedly quicker reactions in the congruent condition than in the incongruent condition. No significant group effect was observed, $F(1, 58) = 0.226$, $p = 0.637$, $\eta p^2 = 0.004$, indicating no significant difference in reaction time between the music-trained group and control group. The condition-group interaction was nonsignificant, $F(1, 58) = 0.683$, $p = 0.412$, $\eta p^2 = 0.012$. An independent samples t-test revealed no significant difference in the Stroop interference effect (accuracy in incongruent trials minus accuracy

Fig. 5. ERP results at the FCz site during the auditory Stroop task for the music training and control groups. (a) Ninc waveforms under congruent and incongruent conditions; (b) Ninc topographies under congruent and incongruent conditions; (c) Ninc difference waveforms (incongruent minus congruent); (d) Ninc difference topographies (incongruent minus congruent).

in congruent trials) between the music training and the control groups, $t(58) = -0.826$, $p = 0.412$, Cohen's $d = -0.213$, 95% CI [-27.815, 11.559] (see Fig. 2a,b).

3.2 Event-Related Potential (ERP) Results

In the visual modality, the ANOVA results for the mean amplitude of the N450 showed a significant main effect of condition, $F(1, 58) = 6.635$, $p = 0.013$, $\eta p^2 = 0.103$, with the incongruent condition eliciting a more negative N450 amplitude than the congruent condition (see Fig. 3a,b). However, neither the main group effect ($F(1, 58) = 0.603$, $p = 0.441$, $\eta p^2 = 0.010$) nor the condition-group interaction ($F(1, 58) = 0.037$, $p = 0.848$, $\eta p^2 = 0.001$) reached significance. Independent-samples t-test results showed no significant difference in the N450 difference wave (incongruent minus congruent) between the musician group and the control group, $t(58) = 0.193$, $p = 0.848$, Cohen's $d = 0.050$, 95% CI [-1.226, 1.487] (see Fig. 3c,d). For the SP component, the ANOVA results revealed a significant main

effect of condition, $F(1, 58) = 7.101$, $p = 0.010$, $\eta p^2 = 0.109$, with the incongruent condition eliciting a larger SP amplitude than the congruent condition (see Fig. 4a,b). No significant group effect was observed, $F(1, 58) = 0.319$, $p = 0.575$, $\eta p^2 = 0.005$; however, the condition-group interaction was significant ($F(1, 58) = 6.914$, $p = 0.011$, $\eta p^2 = 0.107$). Independent-samples t-test results further showed that the SP difference wave (incongruent minus congruent) in the musician group was significantly smaller than in the control group, $t(58) = -2.630$, $p = 0.011$, 95% CI [-3.059, -0.414], Cohen's $d = -0.679$ (see Fig. 4c,d).

In the auditory modality, the ANOVA results for the mean amplitude of the Ninc revealed a significant main effect of condition, $F(1, 58) = 8.338$, $p = 0.005$, $\eta p^2 = 0.126$, with the incongruent condition eliciting a stronger negative Ninc response than congruent ones (see Fig. 5a,b). However, neither the main effect of group ($F(1, 58) = 0.001$, $p = 0.981$, $\eta p^2 < 0.001$) nor the condition × group interaction ($F(1, 58) = 0.194$, $p = 0.661$, $\eta p^2 = 0.003$) reached
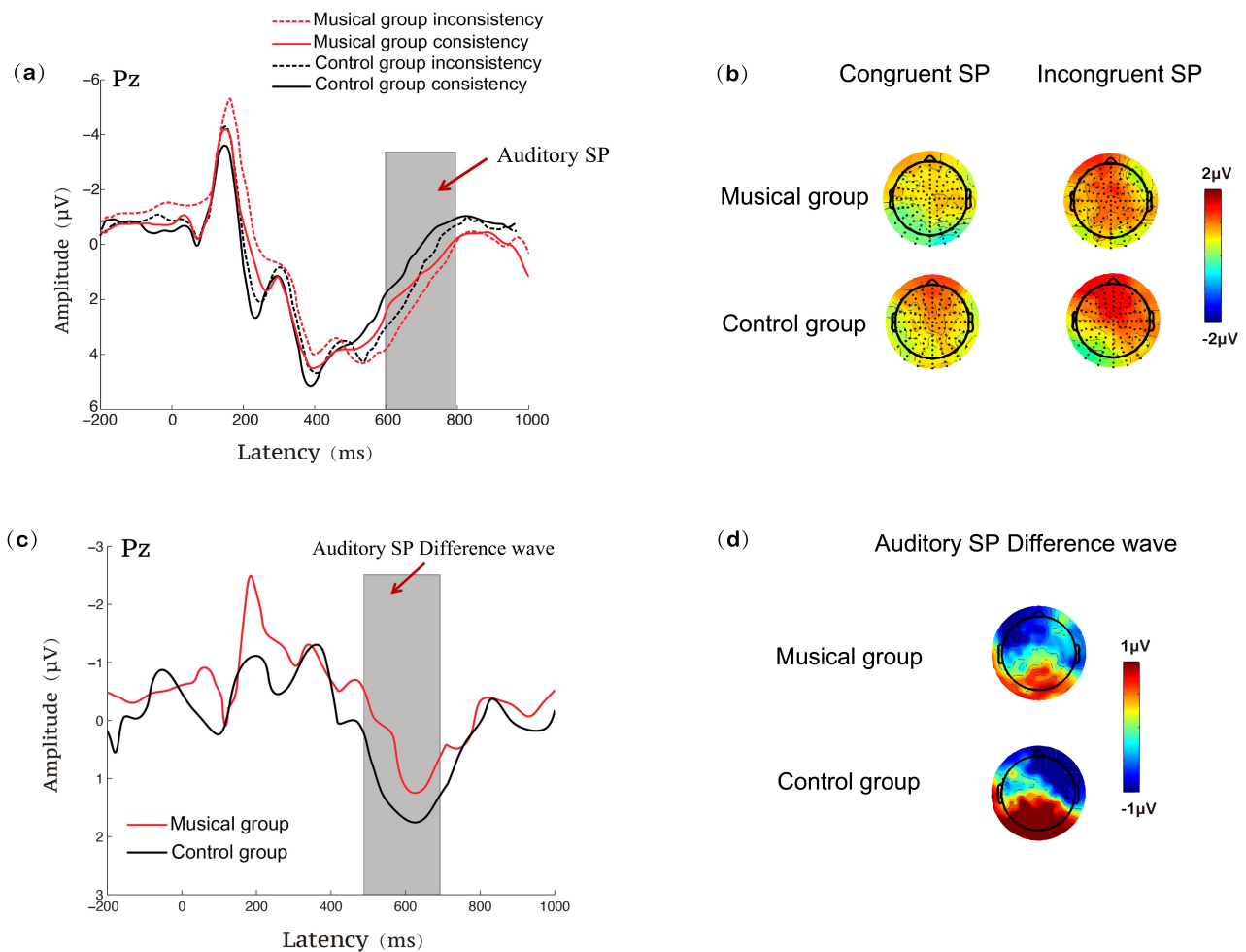
Fig. 6. ERP results at the Pz site during the auditory Stroop task for the music training and control groups. (a) SP waveforms under congruent and incongruent conditions; (b) SP topographies under congruent and incongruent conditions; (c) SP difference waveforms (incongruent minus congruent); (d) SP difference topographies (incongruent minus congruent).
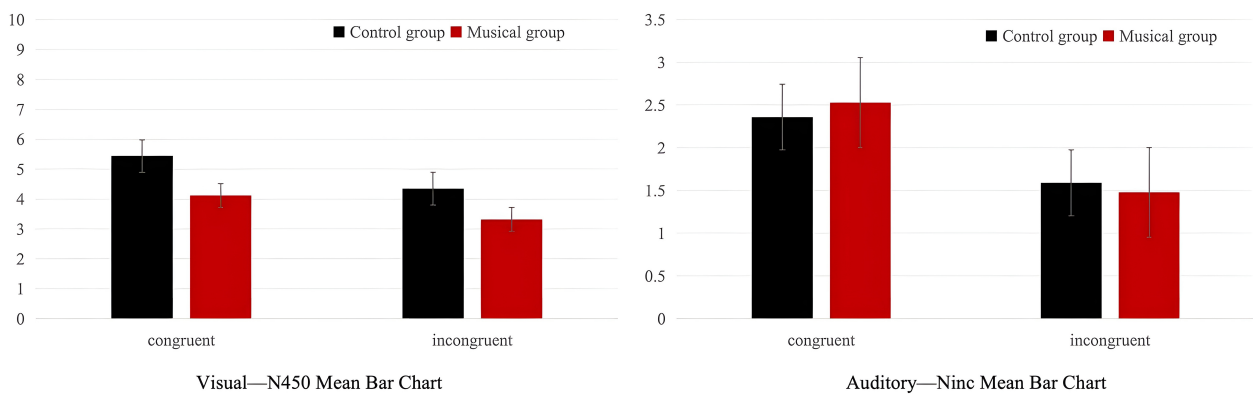
Fig. 7. Means and standard deviations of the N450 component in the musician and non-musician groups across visual and auditory modalities.

significance. Independent-samples t-test results indicated no significant difference in the Ninc difference wave (incongruent minus congruent) between the musician group and the control group, $t(58) = -0.440$, $p = 0.661$, 95% CI [–1.537, 0.982], Cohen's $d = -0.114$ (see Fig. 5c,d). For

the SP component, the ANOVA findings revealed a notable primary impact of the condition, $F(1, 58) = 4.764$, $p = 0.033$, $\eta p^2 = 0.076$, with the incongruent condition eliciting greater SP amplitude compared to the consistent condition (see Fig. 6a,b). However, neither the main effect of group

($F(1, 58) = 0.814$, $p = 0.371$, $\eta p^2 = 0.014$) nor the condition × group interaction ($F(1, 58) = 1.053$, $p = 0.309$, $\eta p^2 = 0.018$) were significant. Independent-samples t-test results further revealed no significant difference in the SP difference wave (incongruent minus congruent) between the musician and control groups, $t(58) = -1.026$, $p = 0.309$, 95% CI [–2.124, 0.684], Cohen's $d = -0.265$ (see Fig. 6c,d).

The means and standard deviations of the amplitudes for the congruent and incongruent conditions across the visual and auditory tasks in both groups are presented in **Supplementary Table 1**.

3.3 Time-Frequency Results

In the visual modality, ANOVA on theta-band power within the 400–700 ms time window revealed no significant main effect of condition, $F(1, 58) = 0.196$, $p = 0.660$, $\eta p^2 = 0.003$, indicating no significant difference in theta power between congruent and incongruent conditions. The primary influence of participant type was also not significant, $F(1, 58) = 0.044$, $p = 0.835$, $\eta p^2 = 0.001$. Additionally, the interaction between participant type and condition was not significant, $F(1, 58) = 2.599$, $p = 0.112$, $\eta p^2 = 0.043$. Independent-samples t-tests further showed no significant differences between groups in the congruent condition ($t(58) = 0.944$, $p = 0.349$, 95% CI = [–0.016, 0.045], Cohen's $d = 0.244$) or in the incongruent condition ($t(58) = -0.584$, $p = 0.562$, 95% CI = [–0.040, 0.022], Cohen's $d = -0.151$).

Within the 400–700 ms time window, ANOVA on beta-band power showed no notable condition impact, $F(1, 58) = 0.264$, $p = 0.609$, $\eta p^2 = 0.005$, indicating no significant difference in beta power between congruent and incongruent conditions. The main effect of participant type was also not significant, $F(1, 58) = 0.610$, $p = 0.438$, $\eta p^2 = 0.010$, and the interaction between participant type and condition was not significant, $F(1, 58) = 0.015$, $p = 0.903$, $\eta p^2 < 0.001$. Independent-samples t-tests further revealed no notable variations between groups in the congruent condition ($t(58) = -0.794$, $p = 0.430$, 95% CI = [–0.002, 0.001], Cohen's $d = -0.205$) or in the incongruent condition ($t(58) = -0.524$, $p = 0.602$, 95% CI = [–0.002, 0.001], Cohen's $d = -0.135$).

In the auditory modality, ANOVA on theta-band power within the 400–700 ms time window revealed no significant main effect of condition, $F(1, 58) = 3.933$, $p = 0.052$, $\eta p^2 = 0.063$, indicating no significant difference in theta power between congruent and incongruent conditions. No significant effects were found for participant type ($F(1, 58) < 0.001$, $p = 0.984$, $\eta p^2 < 0.001$) or its interaction with condition ($F(1, 58) = 0.329$, $p = 0.568$, $\eta p^2 = 0.006$). Independent-samples t-tests further showed no significant group differences in the congruent condition ($t(58) = 0.418$, $p = 0.677$, 95% CI = [–0.014, 0.022], Cohen's $d = 0.108$) or in the incongruent condition ($t(58) = -0.326$, $p = 0.745$, 95% CI = [–0.025, 0.018], Cohen's $d = -0.084$).

In the auditory modality, ANOVA on high-beta band power within the 400–700 ms time window demonstrated no substantial primary effect of condition, $F(1, 58) = 2.917$, $p = 0.093$, indicating no significant difference in beta power between the congruent and incongruent conditions. The main effect of group was significant, $F(1, 58) = 3.999$, $p = 0.050$. Moreover, the interaction between group and condition was significant, $F(1, 58) = 6.103$, $p = 0.016$. Independent-sample t-tests revealed that the music training group exhibited significantly lower beta power than the control group under both congruent ($t(58) = -3.240$, $p = 0.002$) and incongruent conditions ($t(58) = 2.242$, $p = 0.033$; see Fig. 7, and Fig. 8).

Within the 400–700 ms time window, ANOVA on beta-band power demonstrated no substantial primary effect of condition, $F(1, 58) = 2.917$, $p = 0.093$, $\eta p^2 = 0.048$, indicating no significant difference in beta power between congruent and incongruent conditions. Participant type showed a marginally significant main effect, $F(1, 58) = 3.999$, $p = 0.050$, $\eta p^2 = 0.065$. Importantly, participant type and condition demonstrated a significant interaction, $F(1, 58) = 6.103$, $p = 0.016$, $\eta p^2 = 0.095$. Independent-samples t-tests showed that, in both congruent ($t(58) = -3.240$, $p = 0.002$, 95% CI = [–0.004, –0.001], Cohen's $d = -0.837$) and incongruent conditions ($t(58) = 2.242$, $p = 0.033$, 95% CI = [0.002, 0.049], Cohen's $d = 0.579$), beta power in the music group was significantly lower than in the control group (see Fig. 9 and Fig. 10).

The means and standard deviations of theta and beta power under the congruent and incongruent conditions in the two groups are presented in **Supplementary Table 2**.

3.4 Neural-Behavioral Correlations

To further assess whether the observed neural effects were functionally related to behavioral performance, we conducted correlation analyses between ERP/time-frequency measures and Stroop behavioral indices. Several significant associations emerged.

(a) The Stroop reaction time (RT) effect was positively correlated with the visual N450 difference wave ($r = 0.26$, $p = 0.048$). It showed an even stronger positive correlation with the late SP difference wave ($r = 0.38$, $p = 0.003$). These findings suggest that individuals with greater conflict-related ERP effects also exhibit greater behavioral Stroop interference.

(b) In the time-frequency domain, the Stroop RT effect was positively correlated with left-hemispheric high-β power (Condition 1; $r = 0.38$, $p = 0.003$), while the left SP difference wave was negatively correlated with the Stroop accuracy effect ($r = -0.30$, $p = 0.021$).

(c) Moreover, in the visual modality, frontal-central θ power correlated positively with accuracy ($r = 0.29$, $p = 0.027$), indicating a role of frontal θ oscillations in supporting task performance.
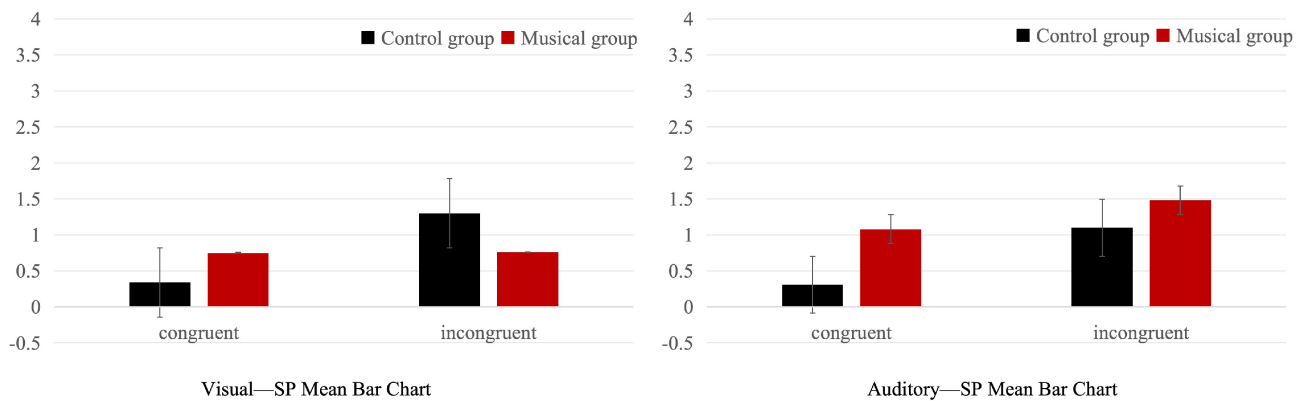
Fig. 8. Means and standard deviations of the SP component in the musician and non-musician groups across visual and auditory modalities.

Taken together, these results indicate that neural indices, including slow potentials and oscillatory activity, are not epiphenomenal but are directly related to behavioral measures of Stroop interference. Importantly, the correlations with the Stroop RT effect suggest that at least some of the group-related neural differences reflect variability in conflict-resolution efficiency, rather than purely motor-related processes (see **Supplementary Table 3**).

4. Discussion

To investigate the modality-specific versus modality-general effects of music training on inhibitory control, this study employed ERP and time-frequency analysis techniques to examine differences in inhibitory control between university-level music majors (music training group) and non-music majors (control group) during visual and auditory Stroop tasks. Previous studies have suggested that inhibitory control may exhibit both cross-modal and modality-specific characteristics, although the evidence remains inconsistent. Most research has focused on the visual modality, using N450 and SP to index conflict monitoring and resolution [24,25]. In contrast, the function of inhibitory control in the auditory modality has been largely neglected, despite auditory processing being a core component of musical training. This study compared a music and non-music group on Stroop tasks in both visual and auditory modalities to examine whether the effects of musical training on inhibitory control are modality-specific.

At the behavioral level, no significant differences were observed between the music training and control groups in terms of accuracy, reaction time, or Stroop interference effects across both the visual and auditory modalities. In the ERP results, a significant group difference was found only in the visual modality; the Stroop-related SP difference wave (incongruent minus congruent) was significantly smaller in the music training group than in the control group. This pattern may indicate that the music group allocated fewer cognitive resources to the visual modality

or employed more automated processing strategies to complete the task [35]. This does not preclude the possibility that the music group employed alternative information processing strategies in other brain regions or at different time windows, which warrants further investigation with additional electrode sites and temporal analyses in future studies. However, the time-frequency results revealed that in the auditory modality, the music training group exhibited significantly lower beta-band power under both congruent and incongruent conditions, suggesting more efficient neural resource allocation during auditory conflict resolution. Overall, the findings demonstrate that inhibitory control exhibits modality specificity across visual and auditory channels. Specifically, music training significantly enhanced individuals' conflict-resolution abilities in the auditory modality, indicating a modality-specific facilitative effect of music training on inhibitory control.

The SP component and beta-band activity are widely recognized as neural indices reflecting conflict resolution processing [24,36–39]. Larger SP difference amplitudes typically indicate that individuals recruit more cognitive resources to resolve conflict [24]. Conversely, reductions in beta-band power generally reflect greater neural regulatory efficiency during conflict resolution. Therefore, although the music training group exhibited smaller SP amplitudes in the visual modality, possibly suggesting a lack of advantage in that channel, the significant decrease in beta power observed in the auditory modality reflects more efficient neural processing [36,40]. Taken together, these results demonstrate that music training significantly enhances the neural regulatory capacity during conflict resolution in the auditory channel. However, no similar advantage was observed in the visual channel, indicating that the music training modality specifically facilitated inhibitory control. Previous studies have found that beta-band activity is associated with response inhibition and conflict resolution abilities, with lower beta power commonly linked to stronger conflict regulation capacity. For example, Bartoli *et al.* (2018) [41]
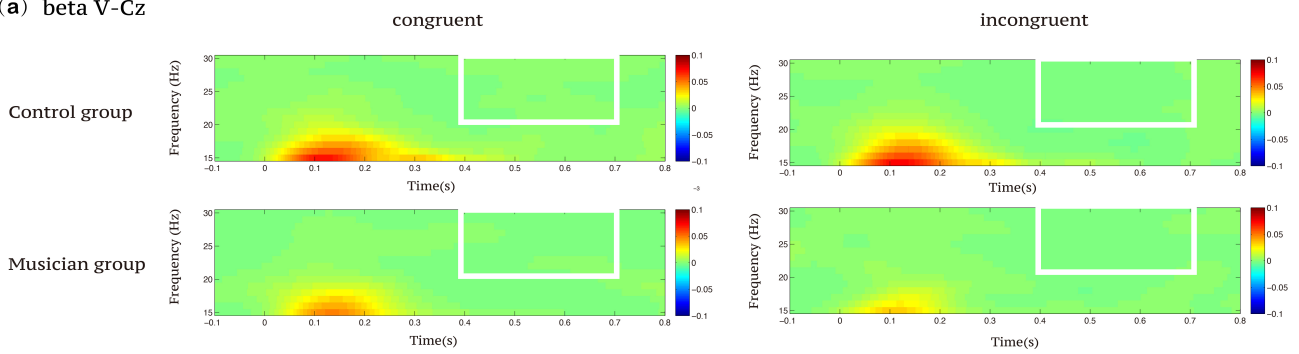
IMR Press

(**a**) theta V-FCz
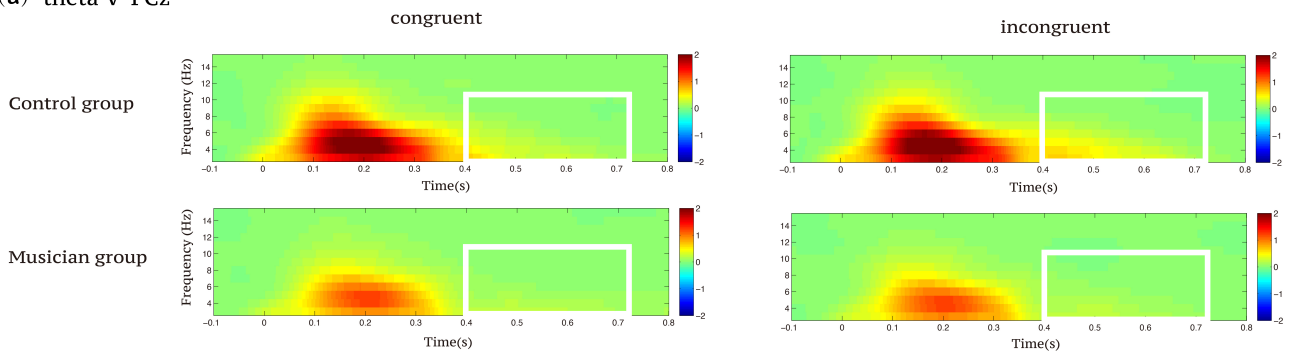
(**b**) theta L-FCz
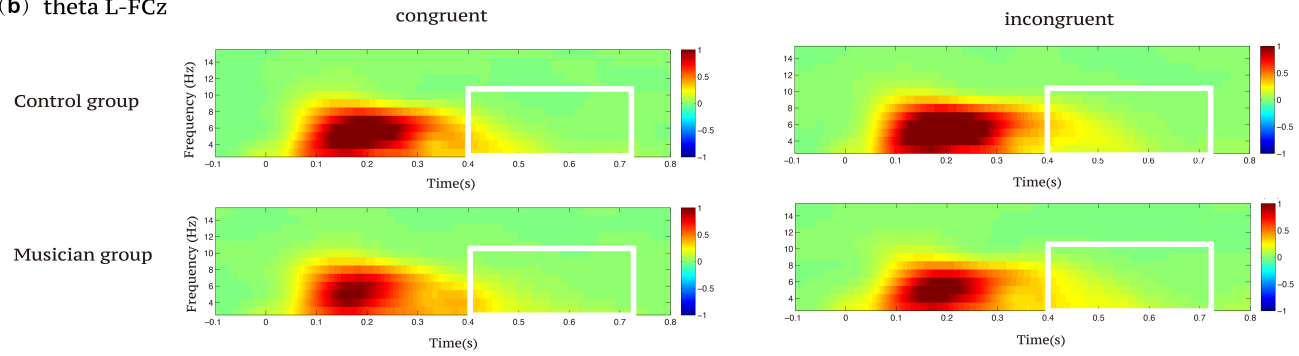
Fig. 9. Time–frequency representations of theta activity during the Stroop task in the musician and control groups across visual and auditory modalities. (a) Visual task-related theta-band activity at the FCz electrode site. (b) Auditory task-related theta-band activity at the FCz electrode site.
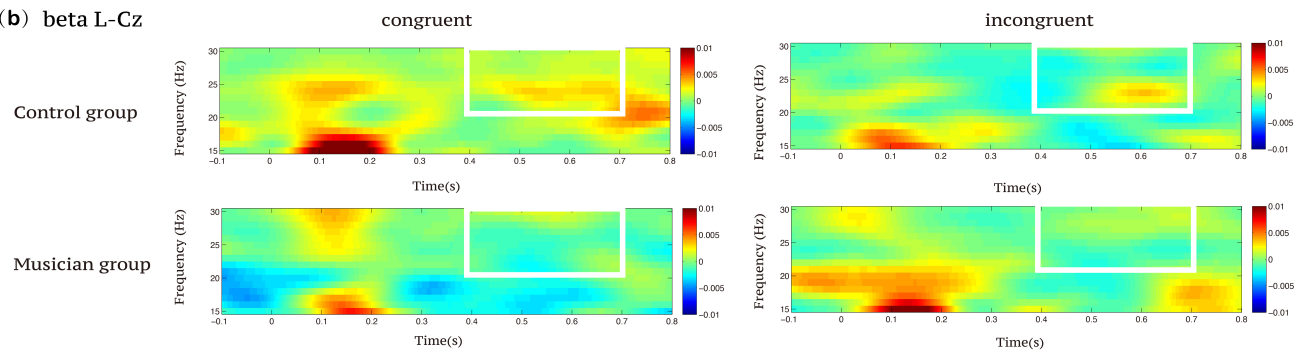
(**a**) beta V-Cz

(**b**) beta L-Cz

Fig. 10. Time–frequency representations of beta activity during the Stroop task in the musician and control groups across visual and auditory modalities. (a) Visual task-related beta-band activity at the Cz electrode site. (b) Auditory task-related beta-band activity at the Cz electrode site.

demonstrated a strong correlation between beta-band neural oscillations and conflict processing, with notably heightened beta-wave amplitudes detected in the dorsolateral prefrontal cortex during tasks involving cognitive incongruity. Similarly, Li *et al*. (2025) [42] demonstrated that beta band activity is tightly linked to the process of conflict resolution, showing significant changes during the later stages of the Stroop task (920–1040 ms), with lower beta power observed in incongruent trials. Consistent with these findings, the present study revealed that participants in the music training group exhibited lower beta band power, which may reflect more efficient conflict resolution. This suggests that music training enhances conflict resolution ability, potentially in a modality-specific manner, with more pronounced effects on auditory processing.

Musicians' advantages in conflict resolution may stem from continuous training in conflict monitoring and behavioral regulation abilities during long-term music practice [16,43]. First, music training is a cognitively demanding, integrative activity in which individuals must continuously monitor multiple dimensions such as rhythm, pitch, and fingering during performance. They need to rapidly adjust their actions based on auditory feedback and bodily sensations, a process that requires ongoing error detection and conflict monitoring [44,45]. Especially during instrumental performance, and more so in ensemble or improvisational contexts, performers must maintain high levels of attention and focus on complex tasks while actively suppressing interference from both internal sources (e.g., mind-wandering and irrelevant thoughts) and external distractions (e.g., environmental noise and phone notifications). Additionally, they must coordinate their timing and intonation with others [10], placing high demands on cognitive control systems. Therefore, long-term music training may enhance individuals' selective processing of goal-relevant information during conflict tasks and improve their ability to inhibit irrelevant distractions [46]. Secondly, the inhibitory control advantages exhibited by the music training group may also arise from long-term training-induced plasticity in brain regions related to inhibitory control, such as the dorsolateral prefrontal cortex, anterior cingulate cortex, supplementary and pre-supplementary motor areas (SMA/pre-SMA), insula, and inferior frontal gyrus [47]. These regions are core neural substrates of executive function and conflict resolution [48,49]. Previous studies indicated that musical training intensity correlates with inhibitory control-related brain activity. For example, James *et al*. (2014) [50] used functional magnetic resonance imaging (fMRI) and found a significant positive correlation between musical training experience and gray matter density in the middle orbitofrontal gyrus of the right hemisphere and the left inferior frontal gyrus. Fauvel *et al*. (2014) [51] reported that a music training group showed significantly enhanced functional connectivity between the right inferior frontal gyrus and the anterior cingulate cortex compared to the control group. These brain regions are thought to play key roles in

inhibitory control processes [52–54]. Additionally, Gaser and Schlaug (2003) [55] found that amateur musicians exhibit moderate increases in gray matter volume compared with non-musicians and professional musicians. Other evidence also indicates structural and functional differences in the prefrontal cortex between musicians and non-musicians [51].

Music training shows a more pronounced advantage in the auditory modality than in the visual modality, a phenomenon supported by research across multiple dimensions, including sensory channel processing advantages and neural plasticity. For example, Sharma *et al*. (2019) [56] found that musicians outperformed controls under incongruent stimulus conditions in an auditory Stroop task, indicating that music training significantly enhanced inhibitory control efficiency in the auditory system during conflict resolution. This suggests that the facilitative effect of music training on inhibitory control is not universally applicable across all sensory modalities but is more explicitly manifested in the auditory channel, closely related to its processing pathways. Furthermore, music training can continuously shape individuals' multilevel auditory processing pathways from the auditory brainstem to the cortex. Studies have shown that early musical experience promotes the development of key regions, such as the left temporal auditory cortex, with musicians generally exhibiting greater cortical thickness and enhanced auditory perceptual abilities [57]. Additional research has demonstrated that musicians exhibit superior auditory perception in noisy environments, with their auditory brainstem responses showing more precise neural encoding of speech signals than non-musicians [58].

It is worth noting that this study employed a cross-sectional design comparing a music training group with a control group, which limits the ability to establish a causal relationship between music training and inhibitory control. Additionally, it cannot entirely rule out potential individual differences between the two groups, such as personality traits and baseline cognitive abilities [59]. Therefore, this study controlled for extraneous variables that may affect inhibitory control, including age and intelligence level, as rigorously as possible. This strict control of variables may partially explain why no significant group differences were found in visual modality. This result differs from those of previous studies that have reported differences in visual conflict processing. We speculate that such discrepancies may stem from the varying degrees of control over confounding variables in different research designs. Therefore, to interpret the nonsignificant group differences observed in the visual modality, we considered the effect sizes and their 95% confidence intervals. The intervals indicated that any differences were minimal and likely of little practical significance, suggesting that the findings more accurately reflected the limited impact of musical training on visual conflict processing than the complete absence of an effect. In contrast, time-frequency analysis in the auditory

modality revealed a significant group difference in beta-band power, with the music training group exhibiting a distinct neural advantage. The observed modulation of beta activity, which may reflect enhanced efficiency in resolving auditory conflicts among musically trained individuals, suggests that the facilitative effects of music training on inhibitory control are primarily modality-specific and most prominent in the auditory domain. The identification of this modality-specific pattern offers novel evidence of the domain-constrained benefits of musical training and provides a promising direction for future research aimed at uncovering the underlying neural mechanisms.

5. Conclusions

Our study aimed to investigate the modality-specific versus modality-general effects of music training on inhibitory control by examining processing differences across perceptual modalities. To this end, we employed ERP and time–frequency analysis techniques to compare the neural processing characteristics and behavioral performance of two groups during both visual and auditory Stroop tasks. The results revealed that significant group differences emerged only in the beta-band power during the auditory modality: the music-trained participants demonstrated reduced beta activity compared to controls under both congruent and incongruent circumstances. These findings suggest that inhibitory control is modality-specific. Specifically, music training primarily enhanced conflict resolution abilities in the auditory modality, indicating a facilitative effect of music training on inhibitory control within auditory processing contexts.

Availability of Data and Materials

The original contributions presented in this study are included in this article/supplementary material, further inquiries can be directed to the corresponding authors.

Author Contributions

JN, JJC, JRM: the acquisition, analysis, and interpretation of data. YL: design of the work, funding acquisition, project administration, supervision. JN, JJC: writing–review & editing. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

These experiments were approved by the Ethics Committee of the School of Electronics and Information Engineering at Southwest University (Ethics Approval No. 18A08998) and complied with local regulations and institutional requirements. All participants provided written informed consent before their involvement in the study. The study was carried out in accordance with the guidelines of the Declaration of Helsinki.

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/JIN45358.

References

[1] Diamond A. Executive functions. Annual Review of Psychology. 2013; 64: 135–168. https://doi.org/10.1146/annurev-psych-113011-143750.

[2] Pandey AK, Kamarajan C, Tang Y, Chorlian DB, Roopesh BN, Manz N, *et al*. Neurocognitive deficits in male alcoholics: an ERP/sLORETA analysis of the N2 component in an equal probability Go/NoGo task. Biological Psychology. 2012; 89: 170–182. https://doi.org/10.1016/j.biopsycho.2011.10.009.

[3] Tiryaki İ, Kadak KS, Karakulak EZ, Korkmaz OE, Sever O, Esin İS. Effects of exercise on response inhibition performance in adolescent males with attention deficit hyperactivity disorder: A fNIRS Study. Psychiatry Research. Neuroimaging. 2025; 348: 111963. https://doi.org/10.1016/j.pscychresns.2025.111963.

[4] Dhir S, Teo WP, Chamberlain SR, Tyler K, Yücel M, Segrave RA. The Effects of Combined Physical and Cognitive Training on Inhibitory Control: A Systematic Review and Meta-Analysis. Neuroscience and Biobehavioral Reviews. 2021; 128: 735–748. https://doi.org/10.1016/j.neubiorev.2021.07.008.

[5] Green CS, Bavelier D. Exercising your brain: a review of human brain plasticity and training-induced learning. Psychology and Aging. 2008; 23: 692–701. https://doi.org/10.1037/a0014345.

[6] Li CSR, Morgan PT, Matuskey D, Abdelghany O, Luo X, Chang JLK, *et al*. Biological markers of the effects of intravenous methylphenidate on improving inhibitory control in cocaine-dependent patients. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107: 14455–14459. https://doi.org/10.1073/pnas.1002467107.

[7] Cheng X, Qian L, Fan Y, Tang Q, Wu H. The Effect of Equine-Assisted Activities in Children Aged 7–8 Years Inhibitory Control: An fNIRS Study. Journal of integrative neuroscience. 2023; 22: 89. https://doi.org/10.31083/j.jin2204089.

[8] Brackenridge R, McKenzie K, Murray GC, Quigley A. An examination of the effects of stimulant medication on response inhibition: a comparison between children with and without attention deficit hyperactivity disorder. Research in Developmen-

tal Disabilities. 2011; 32: 2797–2804. https://doi.org/10.1016/j.ridd.2011.05.027.

[9] Wang A, Yang H, Yang Y, Yang J, Yang X, Wen Q, *et al*. Neural markers of methylphenidate response in children with attention deficit hyperactivity disorder and the impact on executive function. Frontiers in Psychiatry. 2025; 16: 1475889. https://doi.org/10.3389/fpsyt.2025.1475889.

[10] Okada BM, Slevc LR. Music training: Contributions to executive function. In Novick JM, Bunting MF, Dougherty MR, Engle RW (eds.) Cognitive and working memory training: Perspectives from psychology, neuroscience, and human development (pp. 206–226). Oxford University Press: New York, NY, USA. 2020. https://doi.org/10.1093/oso/9780199974467.003.0011.

[11] Ludyga S, Gerber M, Brand S, Holsboer-Trachsler E, Pühse U. Acute effects of moderate aerobic exercise on specific aspects of executive function in different age and fitness groups: A meta-analysis. Psychophysiology. 2016; 53: 1611–1626. https://doi.org/10.1111/psyp.12736.

[12] Berger M, Bruggisser F, Schwarz A, Leuenberger R, Hanke M, Gerber M, *et al*. Acute effect of aerobic exercise on inhibitory control and task-related heart rate variability in preadolescent children. Physiology & Behavior. 2025; 292: 114820. https://doi.org/10.1016/j.physbeh.2025.114820.

[13] Zhou M, Wei C, Xie X, Wang Z, Chen L, Zhang X. Effect of Music Therapy on Cognitive Function Among Older Adults: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. The Gerontologist. 2025; 65: gnaf106. https://doi.org/10.1093/geront/gnaf106.

[14] Koshimori Y, Thaut MH. New Perspectives on Music in Rehabilitation of Executive and Attention Functions. Frontiers in Neuroscience. 2019; 13: 1245. https://doi.org/10.3389/fnins.2019.01245.

[15] Wang L. Music Aptitude, Training, and Cognitive Transfer: A Mini-Review. Frontiers in Psychology. 2022; 13: 903920. https://doi.org/10.3389/fpsyg.2022.903920.

[16] Velasquez MA, Winston JL, Sur S, Yurgil K, Upman AE, Wroblewski SR, *et al*. Music training is related to late ERP modulation and enhanced performance during Simon task but not Stroop task. Frontiers in Human Neuroscience. 2024; 18: 1384179. https://doi.org/10.3389/fnhum.2024.1384179.

[17] Brink E. A look at Gordon, Edwin E. theories + music-education. Bulletin of the Council for Research in Music Education. 1983: 1–13.

[18] Slevc LR, Davey NS, Buschkuehl M, Jaeggi SM. Tuning the mind: Exploring the connections between musical ability and executive functions. Cognition. 2016; 152: 199–211. https://doi.org/10.1016/j.cognition.2016.03.017.

[19] Travis F, Harung HS, Lagrosen Y. Moral development, executive functioning, peak experiences and brain patterns in professional and amateur classical musicians: interpreted in light of a Unified Theory of Performance. Consciousness and Cognition. 2011; 20: 1256–1264. https://doi.org/10.1016/j.concog.2011.03.020.

[20] Yu B, Wang X, Ma L, Li L, Li H. The Complex Pre-Execution Stage of Auditory Cognitive Control: ERPs Evidence from Stroop Tasks. PLoS One. 2015; 10: e0137649. https://doi.org/10.1371/journal.pone.0137649.

[21] Strait DL, Kraus N. Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. Frontiers in Psychology. 2011; 2: 113. https://doi.org/10.3389/fpsyg.2011.00113.

[22] Sharma VV, Thaut M, Russo FA, Alain C. Neural Dynamics of Inhibitory Control in Musicians with Absolute Pitch: Theta Synchrony as an Oscillatory Signature of Information Conflict. Cerebral Cortex Communications. 2021; 2: tgab043. https://doi.org/10.1093/texcom/tgab043.

[23] Chung RS, Cavaleri J, Sundaram S, Gilbert ZD, Del Campo-Vera RM, Leonor A, *et al*. Understanding the human conflict processing network: A review of the literature on direct neural recordings during performance of a modified stroop task. Neuroscience Research. 2024; 206: 1–19. https://doi.org/10.1016/j.neures.2024.03.006.

[24] Larson MJ, Clayson PE, Clawson A. Making sense of all the conflict: a theoretical review and critique of conflict-related ERPs. International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology. 2014; 93: 283–297. https://doi.org/10.1016/j.ijpsycho.2014.06.007.

[25] Buzzell GA, Roberts DM, Baldwin CL, McDonald CG. An electrophysiological correlate of conflict processing in an auditory spatial Stroop task: the effect of individual differences in navigational style. International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology. 2013; 90: 265–271. https://doi.org/10.1016/j.ijpsycho.2013.08.008.

[26] van Mourik R, Sergeant JA, Heslenfeld D, Konig C, Oosterlaan J. Auditory conflict processing in ADHD. Journal of Child Psychology and Psychiatry, and Allied Disciplines. 2011; 52: 265–274. https://doi.org/10.1111/j.1469-7610.2010.02339.x.

[27] Giller F, Bensmann W, Mückschel M, Stock AK, Beste C. Evidence for a causal role of superior frontal cortex theta oscillations during the processing of joint subliminal and conscious conflicts. Cortex; a Journal Devoted to the Study of the Nervous System and Behavior. 2020; 132: 15–28. https://doi.org/10.1016/j.cortex.2020.08.003.

[28] Eisma J, Rawls E, Long S, Mach R, Lamm C. Frontal midline theta differentiates separate cognitive control strategies while still generalizing the need for cognitive control. Scientific Reports. 2021; 11: 14641. https://doi.org/10.1038/s41598-021-94162-z.

[29] Oehrn CR, Hanslmayr S, Fell J, Deuker L, Kremers NA, Do Lam AT, *et al*. Neural communication patterns underlying conflict detection, resolution, and adaptation. The Journal of Neuroscience: the Official Journal of the Society for Neuroscience. 2014; 34: 10438–10452. https://doi.org/10.1523/JNEUROSCI.3099-13.2014.

[30] Shenhav A, Botvinick MM, Cohen JD. The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron. 2013; 79: 217–240. https://doi.org/10.1016/j.neuron.2013.07.007.

[31] Wang K, Li Q, Zheng Y, Wang H, Liu X. Temporal and spectral profiles of stimulus-stimulus and stimulus-response conflict processing. NeuroImage. 2014; 89: 280–288. https://doi.org/10.1016/j.neuroimage.2013.11.045.

[32] Arthur Jr W, Day DV. Development of a Short form for the Raven Advanced Progressive Matrices Test. Educational and Psychological Measurement. 1994; 54: 394–403. https://doi.org/10.1177/0013164494054002013.

[33] Suzuki Y, Takeshima H, Kurakata K. Revision of ISO 226" Normal Equal-Loudness-Level Contours" from 2003 to 2023 edition: The background and results. Acoustical Science and Technology. 2024; 45: 1–8. https://doi.org/10.1250/ast.e23.66.

[34] Hao J, Pang Y, Liu Y, Jing Y, Li J, Mi R, *et al*. The Relationship between Formal Music Training and Conflict Control: An ERP Study. Brain Sciences. 2023; 13: 723. https://doi.org/10.3390/brainsci13050723.

[35] Jiang Y, Zheng M. EEG microstates are associated with music training experience. Frontiers in Human Neuroscience. 2024; 18: 1434110. https://doi.org/10.3389/fnhum.2024.1434110.

[36] Tang H, Yu HY, Chou CC, Crone NE, Madsen JR, Anderson WS, *et al*. Cascade of neural processing orchestrates cognitive control in human frontal cortex. eLife. 2016; 5: e12352. https://doi.org/10.7554/eLife.12352.

[37] Khan AU, Hoy CW, Anderson KL, Piai V, King-Stephens D, Laxer KD, *et al*. Neural dynamics of proactive and reactive cog-

IMR Press

nitive control in medial and lateral prefrontal cortex. iScience. 2025; 28: 113375. https://doi.org/10.1016/j.isci.2025.113375.

[38] Larson MJ, Kaufman DAS, Perlstein WM. Conflict adaptation and cognitive control adjustments following traumatic brain injury. Journal of the International Neuropsychological Society: JINS. 2009; 15: 927–937. https://doi.org/10.1017/S1355617709990701.

[39] Zhao J, Liang WK, Juan CH, Wang L, Wang S, Zhu Z. Dissociated stimulus and response conflict effect in the Stroop task: evidence from evoked brain potentials and brain oscillations. Biological Psychology. 2015; 104: 130–138. https://doi.org/10.1016/j.biopsycho.2014.12.001.

[40] Griffiths BJ, Mayhew SD, Mullinger KJ, Jorge J, Charest I, Wimber M, *et al*. Alpha/beta power decreases track the fidelity of stimulus-specific information. eLife. 2019; 8: e49562. https://doi.org/10.7554/eLife.49562.

[41] Bartoli E, Conner CR, Kadipasaoglu CM, Yellapantula S, Rollo MJ, Carter CS, *et al*. Temporal Dynamics of Human Frontal and Cingulate Neural Activity During Conflict and Cognitive Control. Cerebral Cortex (New York, N.Y.: 1991). 2018; 28: 3842–3856. https://doi.org/10.1093/cercor/bhx245.

[42] Li Y, Lin Y, Li Q, Chen Y, Li Z, Chen A. Temporal dynamics analysis reveals that concurrent working memory load eliminates the Stroop effect through disrupting stimulus-response mapping. eLife. 2025; 13: RP100918. https://doi.org/10.7554/eLife.100918.

[43] Wang X, Yamashita M, Guo X, Stiernman L, Kakihara M, Abe N, *et al*. Never too late to start musical instrument training: Effects on working memory and subcortical preservation in healthy older adults across 4 years. Imaging Neuroscience (Cambridge, Mass.). 2025; 3: IMAG.a.48. https://doi.org/10.1162/IMAG.a.48.

[44] Zuk J, Benjamin C, Kenyon A, Gaab N. Correction: Behavioral and Neural Correlates of Executive Functioning in Musicians and Non-Musicians. PloS One. 2018; 13: e0191394. https://doi.org/10.1371/journal.pone.0191394.

[45] Hennessy SL, Sachs ME, Ilari B, Habibi A. Effects of Music Training on Inhibitory Control and Associated Neural Networks in School-Aged Children: A Longitudinal Study. Frontiers in Neuroscience. 2019; 13: 1080. https://doi.org/10.3389/fnins.2019.01080.

[46] Wang X, Zeng T. Musical Expertise Reshapes Cross-Domain Semantic Integration: ERP Evidence from Language and Music Processing. Brain Sciences. 2025; 15: 401. https://doi.org/10.3390/brainsci15040401.

[47] Niendam TA, Laird AR, Ray KL, Dean YM, Glahn DC, Carter CS. Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. Cognitive, Affective & Behavioral Neuroscience. 2012; 12: 241–268. https://doi.org/10.3758/s13415-011-0083-5.

[48] Sachs M, Kaplan J, Der Sarkissian A, Habibi A. Increased engagement of the cognitive control network associated with music training in children during an fMRI Stroop task. PloS One. 2017; 12: e0187254. https://doi.org/10.1371/journal.pone.0187254.

[49] Zhang Z, Peng P, Eickhoff SB, Lin X, Zhang D, Wang Y. Neural substrates of the executive function construct, age-related changes, and task materials in adolescents and adults: ALE meta-analyses of 408 fMRI studies. Developmental Science. 2021; 24: e13111. https://doi.org/10.1111/desc.13111.

[50] James CE, Oechslin MS, Van De Ville D, Hauert CA, Descloux C, Lazeyras F. Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks. Brain Structure & Function. 2014; 219: 353–366. https://doi.org/10.1007/s00429-013-0504-z.

[51] Fauvel B, Groussard M, Chételat G, Fouquet M, Landeau B, Eustache F, *et al*. Morphological brain plasticity induced by musical expertise is accompanied by modulation of functional connectivity at rest. NeuroImage. 2014; 90: 179–188. https://doi.org/10.1016/j.neuroimage.2013.12.065.

[52] Badzakova-Trajkov G, Barnett KJ, Waldie KE, Kirk IJ. An ERP investigation of the Stroop task: the role of the cingulate in attentional allocation and conflict resolution. Brain Research. 2009; 1253: 139–148. https://doi.org/10.1016/j.brainres.2008.11.069.

[53] Simmonds DJ, Pekar JJ, Mostofsky SH. Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. Neuropsychologia. 2008; 46: 224–232. https://doi.org/10.1016/j.neuropsychologia.2007.07.015.

[54] Carter CS, van Veen V. Anterior cingulate cortex and conflict detection: an update of theory and data. Cognitive, Affective & Behavioral Neuroscience. 2007; 7: 367–379. https://doi.org/10.3758/cabn.7.4.367.

[55] Gaser C, Schlaug G. Brain structures differ between musicians and non-musicians. The Journal of Neuroscience: the Official Journal of the Society for Neuroscience. 2003; 23: 9240–9245. https://doi.org/10.1523/JNEUROSCI.23-27-09240.2003.

[56] Sharma VV, Thaut M, Russo F, Alain C. Absolute Pitch and Musical Expertise Modulate Neuro-Electric and Behavioral Responses in an Auditory Stroop Paradigm. Frontiers in Neuroscience. 2019; 13: 932. https://doi.org/10.3389/fnins.2019.00932.

[57] Shahin A, Bosnyak DJ, Trainor LJ, Roberts LE. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. The Journal of Neuroscience: the Official Journal of the Society for Neuroscience. 2003; 23: 5545–5552. https://doi.org/10.1523/JNEUROSCI.23-13-05545.2003.

[58] Musacchia G, Sams M, Skoe E, Kraus N. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104: 15894–15898. https://doi.org/10.1073/pnas.0701498104.

[59] Corrigall KA, Schellenberg EG, Misura NM. Music training, cognition, and personality. Frontiers in Psychology. 2013; 4: 222. https://doi.org/10.3389/fpsyg.2013.00222.