ORIGINAL ARTICLE

WILEY **PSYCHOPHYSIOLOGY** SPR

Reward enhances cross-modal conflict control in object categorization: Electrophysiological evidence

Guanlan Kang[1] | Wenshuo Chang[1] | Lihui Wang[1] | Ping Wei[2] iD | Xiaolin Zhou[1,3,4,5,6] iD

[1]School of Psychological and Cognitive Sciences, Peking University, Beijing, China

[2]Beijing Key Laboratory of Learning and Cognition and School of Psychology, Capital Normal University, Beijing, China

[3]Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China

[4]Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, China

[5]PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China

[6]Institute of Psychological and Brain Sciences, Zhejiang Normal University, Zhejiang, China

Correspondence Xiaolin Zhou, Ph.D., School of Psychological and Cognitive Sciences, Peking University, 5 Yiheyuan Road, Beijing, 100871, China. Email: xz104@pku.edu.cn or Ping Wei, Ph.D, email: aweiping@gmail.com

Abstract

Cross-modal conflict arises when information from different sensory modalities are incompatible with each other. Such conflict may influence the processing of stimuli in the task-relevant modality and call for cognitive control to resolve this conflict. Here, we investigate how reward modulates cross-modal conflict control during object categorization. Participants categorized pictures as representing animate or inanimate objects while ignoring auditory stimuli. We manipulated the audiovisual congruency and performance-dependent reward (reward vs. no-reward). Behavioral results showed a significant cross-modal interference effect only in the no-reward condition, not in the reward condition. Neurally, we found that the frontocentral N2and theta band oscillations were larger in the incongruent condition than in the congruent condition, but only when there was no reward for performance. The converging behavioral and electrophysiological evidence demonstrates that reward enhances cognitive control in a cross-modal context and reduces cross-modal conflict.

KEYWORDS
cross-modal conflict, N2, reward, theta band oscillations

1 | INTRODUCTION

We live in a complex, multisensory environment. Sensory inputs from different modalities (e.g., vision, audition, touch) interact to influence human performance. For example, a noisy cell phone can be distracting to students studying in the library. The cell phone ringing may cause the students to miss words in the text and increase the amount of time needed to finish reading their current paragraphs. However, if there is an important exam on the next day, the noise may have little influence on the students. The motivation to pass the exam can act to enhance attention on reading and minimize the auditory distraction. One component of the motivation to study for the exam is reward expectation (e.g., achieving a high grade on the exam). The current study aimed to investigate how reward expectation influences cross-modal conflict control during object categorization.

The example given above illustrates how cross-modal conflict can arise when there is incongruence in information arriving from different modalities. The congruency between stimuli from target and nontarget modalities can be defined according to the stimuli's temporal, spatial, and/or

| **1 of 12**

PSYCHOPHYSIOLOGY SPR SOCIETY FOR PSYCHOPHYSIOLOGICAL RESEARCH

higher-level (such as semantics) characteristics. Previous studies have demonstrated cross-modal interference effects during object recognition (Diaconescu, Alain, & McIntosh, 2011; Molholm, Ritter, Javitt, & Foxe, 2004; Vogler & Titchener, 2011; Yuval-Greenberg & Deouell, 2009). For example, in Yuval-Greenberg and Deouell (2009), pairs of animal sounds and pictures were presented, and the sound and picture were associated with either the same animal (congruent) or different animal (incongruent). Participants were asked to recognize either auditory or visual objects in separate sessions. The authors found that incongruent trials led to slower reaction times than congruent trials. Studies using the ERP technique showed that incongruent audio-visual stimuli elicited larger N2 (220–380 ms) and N400 components than congruent stimuli (Molholm et al., 2004; Zimmer, Itthipanyanan, Grent-'t-Jong, & Woldorff, 2010). Neuroimaging studies showed that conflicting audiovisual stimuli elicited greater activity in dorsolateral prefrontal cortex (DLPFC) and anterior cingulate cortex (ACC), suggesting that DLPFC and ACC play important roles in cross-modal conflict monitoring and resolution (Weissman, Warner, & Woldorff, 2004, 2009).

Reward expectation enhances cognitive control and promotes human performance (Botvinick & Braver, 2015; Padmala & Pessoa, 2011; Soutschek, Stelzel, Paschke, Walter, & Schubert, 2015). However, this notion is mainly derived from studies in the visual domain, which limits the ability to generalize from visual contexts to cross-modal contexts. Meanwhile, a number of studies have suggested the existence of a supramodal mechanism of cognitive control (Haupt, Axmacher, Cohen, Elger, & Fell, 2009; Roberts & Hall, 2008; Spagna, Mackie, & Fan, 2015; Weissman et al., 2009). For example, in a study in which the Stroop effect was induced in both the visual and the auditory modalities, Roberts and Hall (2008) found that ACC, PFC, and the parietal lobe had consistent conflict-related (incongruent vs. congruent) activities across task modalities, suggesting a supramodal control mechanism in coping with conflict (see also Ye & Zhou, 2009). Extending these studies, we recently investigated the effect of reward on cognitive control in the cross-modal context and observed a facilitatory effect of reward on cross-modal conflict resolution (Kang, Wang, & Zhou, 2017). In Experiment 1 of this study, a cue indicating reward information of the current trial was presented in advance, followed by an auditory and a visual letter (*A* or *O*) presented simultaneously. Participants were instructed to discriminate either the auditory or the visual letter. Results showed a smaller interference effect (response times [RTs] for incongruent trials minus RTs for congruent trials) in the reward condition as compared with the no-reward condition, suggesting an enhancement of cognitive control under the reward expectation

context. However, this behavioral study speaks nothing to the neural substrates and dynamics underlying the observed effects; it remains unclear whether consistent reward modulations on both unimodal and cross-modal conflicts are supported by a common neural network (i.e., a supramodal mechanism) or by distinct neural substrates (i.e., modality-specific mechanisms).

To address this limitation and further investigate the electrophysiological dynamics of reward modulation on cross-modal conflict resolution, in the current study, we employed a cross-modal conflict task while recording scalp EEG signals during the task. We manipulated the audiovisual congruency by using pictures and sounds from the same object (animate/inanimate) or different objects. Participants were asked to categorize the pictures as animate or inanimate objects while ignoring the sounds. The performance-dependent reward was manipulated block-wise. With this manipulation, we were able to both replicate the previous findings that reward could reduce cross-modal conflict and reveal the electrophysiological evidence for this reward modulation.

Regarding scalp EEG, previous studies on conflict processing have shown that the frontocentrally distributed N2 (200–350 ms) is more negative on incongruent trials than on congruent trials, suggesting that N2 plays an important role in conflict detection (Larson, Clayson, & Clawson, 2014; van Veen & Carter, 2002; Yeung, Botvinick, & Cohen, 2004). Other studies have also shown conflict-related modulation on EEG activity in the theta (4–8 Hz) frequency (Cavanagh, Zambrano-Vazquez, & Allen, 2012; Cohen & Cavanagh, 2011; Cohen & Donner, 2013; Jiang, Zhang, & van Gaal, 2015). For example, Cohen and Donner (2013) used a Simon task to investigate the role of theta band activity in conflict processing. They observed a conflict-related modulation on the nonphase-locked midfrontal theta band activity; indeed, the theta band activity was able to predict performance in conflict resolution. These results suggest that theta band oscillations are related to conflict processing (Cohen & Donner, 2013). Studies have also suggested that N2 has a spectral signature in theta band (see Cavanagh & Frank, 2014, for a review). Taken together, for the current study, we predicted that reward would modulate conflict-related N2 amplitude and theta band oscillations.

Separately, previous studies using a trial-by-trial cue-target paradigm demonstrated that the cue to reward modulates the preparatory neural activity (alpha band oscillations) between the reward cue and the target (e.g., Sawaki, Luck, & Raymond, 2015; van den Berg, Krebs, Lorist, & Woldorff, 2014). In contrast, the present study manipulated reward presence in a blockwise manner. A secondary interest of the study was to examine whether reward would also affect the prestimulus neural activity in such a context.

2 | METHOD

2.1 | Participants

Twenty-five graduate or undergraduate students (11 female, age range 18–26 years old) from universities in Beijing participated in this study. All participants were right-handed, had normal or corrected-to-normal vision, and had self-reported normal hearing. This study was carried out in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the School of Psychological and Cognitive Sciences, Peking University.

2.2 | Apparatus and materials

Both auditory and visual stimuli were used in the study. There were two types of stimuli (i.e., animate and inanimate) in the visual and auditory stimuli categories, respectively.

Visual stimuli were 40 black and white line drawings selected from Snodgrass and Vanderwart (1980) and were standardized on familiarity and complexity. There were 20 animate pictures and 20 inanimate pictures. All visual stimuli were presented at the center of a black background and were matched according to size. The size of the pictures was $7° \times 7°$ in visual angle. Participants were seated 57 cm from a CRT monitor.

Auditory stimuli were 40 complex sounds, which were semantically related to the corresponding visual stimuli. Twenty animate sounds and 20 inanimate sounds with a sampling rate of 44100 Hz (16-bit) were used in the study. All the auditory stimuli were normalized and delivered binaurally at an intensity level of about 60 dB. Auditory stimuli were presented via Etymotic ER-2 air ear phones (Etymotic Research, Elk Grove Village, IL). The duration of each sound was 500 ms.

2.3 | Design and procedure

Each trial began with a fixation sign at the center of the screen, lasting 500 ms (Figure 1). Then, an auditory/visual stimulus or an audiovisual stimulus pair was presented for 500 ms. The audiovisual stimulus pairs could semantically match audiovisual stimulus pairs (congruent; the picture was always paired with the sound from the same object, e.g., the picture of a bird was paired with the sound of a bird chirp), or nonmatch audiovisual stimulus pairs (incongruent; the picture was always paired with a sound from another category, e.g., the picture of a bird paired with the sound of bell). Participants were instructed to ignore auditory stimuli and they indicated whether the object in the picture was animate or inanimate by pressing a left or right response key with the index finger of the left or right hand, respectively.
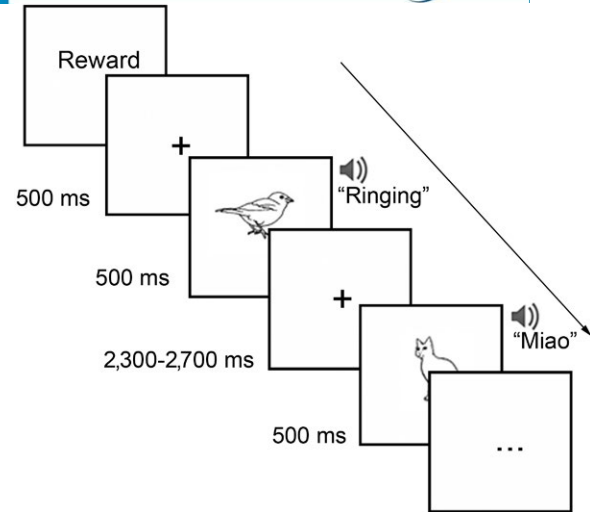
FIGURE 1 An exemplar block of the experimental task. At the beginning of each block, a cue was presented to instruct participants whether they would receive a performance-dependent reward in the current block. Each trial begins with a 500-ms fixation, followed by an audiovisual stimuli pair lasting 500 ms. Participants were asked to categorize pictures as representing animate or inanimate objects. The response within 1,500 ms postonset of the target would be recorded. The intertrial interval was set at 800–1,200 ms

The mapping between the two response keys and two object categories were counterbalanced across participants. Additionally, although we focused on how reward modulates the audiovisual conflict effect, we included visual-only and auditory-only trials to prevent participants from predicting the upcoming target. The auditory-only trials served as filler trials, and no response was required for these trials. The intertrial interval was 800–1,200 ms.

At the beginning of each experimental block, participants viewed a cue informing them whether they could obtain additional rewards in the current block. Specifically, a cue for "reward block" indicated that a reward would be given for each trial with a successful response, which was defined as a response that was both accurate and faster than a baseline RT; a cue for "no-reward block" indicated that no additional reward would be given in the current block. The baseline RT was dynamically adjusted after each block, depending on the performance of the individual participant. The baseline RT for the first block was set to the mean RT in practice trials. Then, the baseline RT was updated for the next block by calculating the mean of the baseline RT for the current block and the mean RT in the current block, adopting the following formula: baseline RT in block $n+1 =$ (baseline RT in block $n +$ mean RT in block n)/2 (see Soutschek et al., 2015).

Thus, the experiment had a 2 (Reward Type: reward vs. no-reward) $\times$ 3 (Stimulus Type: congruent vs. incongruent vs. visual-only) within-participant factorial design. The experiment comprised 360 trials, with 60 trials for each of the six experimental conditions. The 360 trials were divided into three reward

blocks and three no-reward blocks. In addition, there were 60 auditory-only filler trials, distributed in both the reward and no-reward blocks. All the trial types were equally distributed in each block and were presented in a pseudorandomized order.

Prior to the formal experiment, participants completed 30 practice trials. The procedure for the practice trials was the same as for the main experiment, except that no reward information was given. Participants were asked to respond as quickly and accurately as possible, and they received feedback (correct or incorrect) after the button press. The mean RT during the practice was calculated as the baseline RT for the first block of the formal experiment.

In the reward condition, participants won one coin per trial if the response met the threshold, and zero coin if not. No additional reward was delivered in the no-reward condition. Feedback denoting the total number of coins was presented every 10 trials. At the end of the experiment, the coins were exchanged for cash (i.e., 1 coin = 0.1 Chinese yuan), although the participants were not informed of the exchange rate before the experiment. Participants could earn up to 24 yuan of reward based on their performance, which was added to their basic payment of 55 yuan for participating in the study.

2.4 | Behavioral data analysis

Omissions and incorrect trials were excluded from the RT analysis. For each participant, trials with RT more than three standard deviations above or below the mean RT in each experimental condition were discarded as outliers (1.4% of all critical data points). We conducted a 2×3 repeated measures analysis of variance (ANOVA), with the first factor referring to reward type (reward vs. no-reward) and the second factor referring to stimulus type (congruent vs. incongruent vs. visual-only). Similar analysis was conducted on the mean error rates, which were calculated as the proportion of incorrect and missing trials in each condition. The mean RT and error rate in each experimental condition are shown in Figure 2.

In addition, a Bayesian ANOVA (BANOVA; Rouder, Morey, Speckman, & Province, 2012) was conducted on RTs with the software JASP (https://jasp-stats.org). This analysis compares the likelihood of linear models, which involves the main effects and/or interactions between variables while considering the nuisance variables (random effects; e.g., participants). The Bayes factor (BF_{10}) quantifies the ratio of the likelihood of a model with the alternative hypothesis (e.g., H_1: RT differences between reward conditions were reliable) and the likelihood of the default model with a null hypothesis (e.g., H_0: RT differences between reward conditions were not reliable). By convention, $BF_{10} < 3$ implies that the manipulation has no obvious effect on the dependent variable while $BF_{10} > 3$ implies that there is a reliable difference between conditions (Faivre, Mudrik, Schwartz, & Koch, 2014, Rouder et al., 2012). For example, $BF_{10} = 10$ indicates that the alternative hypothesis is 10 times as likely to be true compared with the null hypothesis.

2.5 | EEG recording and preprocessing

EEG was recorded from 64 Ag/AgCl electrodes mounted in an elastic cap (Easycap, Brain Products, Germany). Recordings were referenced online to the nose tip and rereferenced to linked mastoids offline (TP9, TP10). The electrooculogram (EOG) was recorded at two electrode sites; the vertical EOG was monitored from electrodes placed above the right eye and the horizontal EOG from electrodes situated at the outer canthus of the left eye. All electrode impedances were kept below 5 kΩ. The EEG and EOG recordings were amplified by BrainAmps (Brain Products, Germany) using a band-pass filter of 0.016 to 100 Hz, and digitized online at a sampling rate of 500 Hz.
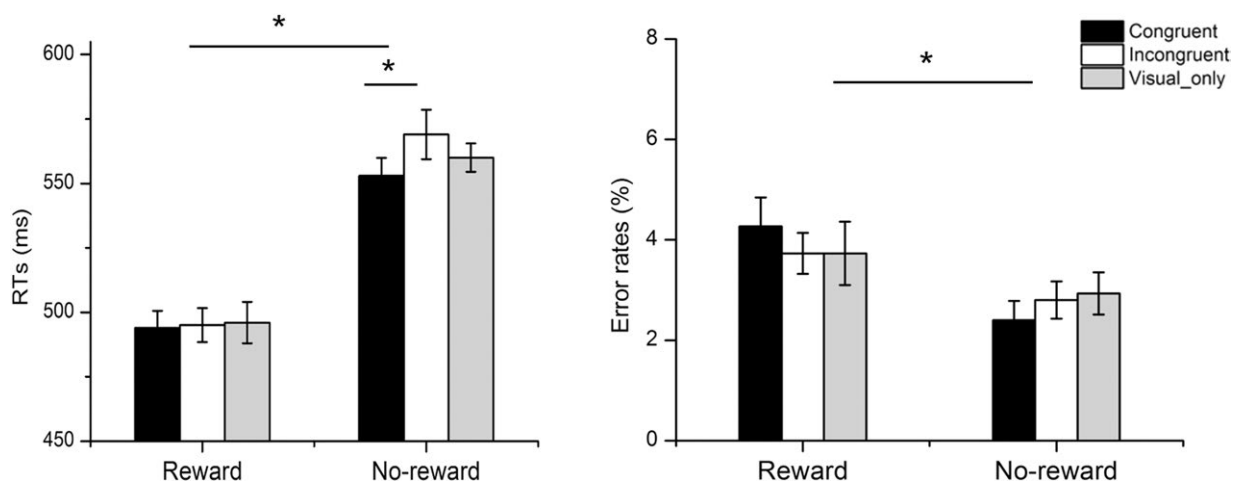
FIGURE 2 Mean reaction times (RTs; left) and error rates (%; right) with within-participant standard errors (Cousineau, 2005) as a function of the experimental conditions. The asterisk indicates a significant difference between conditions (*$p < .05$)

The EEGLAB toolbox (Delorme & Makeig, 2004) was used to preprocess the EEG data. The data were high-pass filtered offline above 0.5 Hz and low-pass filtered below 30 Hz. Ocular artifacts were corrected by a procedure based on independent component analysis (Jung et al., 2000).

2.6 | ERP analysis

For ERP analysis, epochs were extracted from 200 ms prestimulus to 800 ms poststimulus onset. The period from -200 ms to stimulus onset served as the prestimulus baseline. Trials with mean voltages of epochs exceeding $\pm 70\,\mu V$ and with omitted or incorrect responses were excluded. The remaining trials included 95.79% artifact-free trials in total (96.22% for the reward-congruent condition, 97.44% for the reward-incongruent condition, 97.37% for the reward-only-visual condition; 95.19% for the no-reward-congruent condition, 94.55% for the no-reward-incongruent condition, 93.97% for the no-reward-only-visual condition).

According to visual inspection of the potential conflict effect and previous findings on conflict processing, we calculated two cross-modal conflict effects: an early effect (N2, 240–320 ms), and a late effect (N400, 380–450 ms; see Figure 3). Given that the underlying cognitive functions of N2 were suggested to depend on its topographical distribution (Folstein & Van Petten, 2008), the scalp regions in units of electrode clusters were included in the statistical analysis. We selected 15 electrode positions from anterior to posterior areas (Fz, F3, F4, FCz, FC3, FC4, Cz, C3, C4, CPz, CP3, CP4, Pz, P3, P4), and divided them into five electrode clusters (frontal: Fz, F3, F4; frontocentral: FCz, FC3, FC4; central: Cz, C3, C4; parietocentral: CPz, CP3, CP4; parietal: Pz, P3, P4). Three-way repeated measures ANOVA was conducted on the mean amplitude of N2 and N400, respectively, with reward type (reward vs. no-reward), stimulus type (congruent vs. incongruent), and electrode cluster (frontal, frontocentral, central, parietocentral, parietal) as three within-participant factors.

The N2 amplitudes were further analyzed with BANOVA under each electrode cluster to provide stronger evidence for the presence or absence of an effect.

2.7 | Time-frequency analysis

Induced (nonphase-locked) EEG activity was computed by subtracting EEG activity of each individual trial from evoked (phase-locked) EEG activity (the average activity in each condition) for each participant. Time-frequency transformation was performed by convolving the induced activity with a complex Morlet wavelet with a Gaussian kernel of 4-cycle
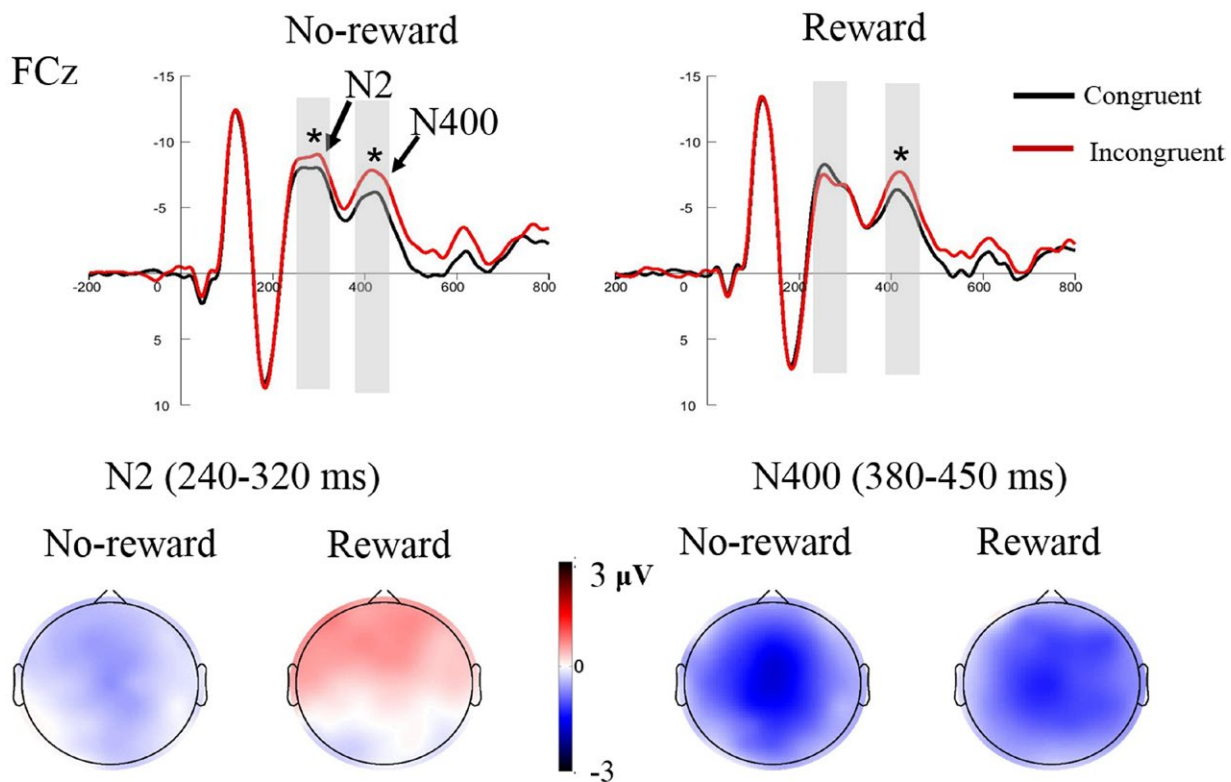
FIGURE 3 Top: ERP responses locked to the stimulus onset in the no-reward condition (left) and the reward condition (right) at FCz. The N2 and N400 are highlighted. In the no-reward condition, N2 and N400 were more negative for incongruent trials as compared with congruent trials (i.e., conflict effect). In the reward condition, only N400 showed a significant conflict effect (*$p < .05$). Bottom: Topography of the average of N2 conflict effects (incongruent vs. congruent; 240–320 ms; left) and N400 conflict effects (380–450 ms; right)

width using Fieldtrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). This procedure was applied to frequencies ranging from 2 to 30 Hz in steps of 1 Hz and time interval between -700 and 1,500 ms in steps of 10 ms. Event-related power was calculated as the percentage change in power relative to the baseline (i.e., -200 to 0 ms relative to stimulus onset).

For the time frequency analysis, we focused on two components: frontocentral theta, which is a well-documented signature for cognitive control (Cavanagh & Frank, 2014; Cohen, 2014), and posterior alpha, which was found to be involved in expecting the upcoming visual stimuli and attentional preparation (Foxe & Snyder, 2011; van den Berg et al., 2014). Based on these hypotheses, the analysis on theta was focused on the frontocentral electrodes, while the analysis on alpha was focused on centroposterior electrodes.

For the statistical analysis of theta band activities, cluster-based permutation test (Maris & Oostenveld, 2007) was implemented on theta band activities across the frequency of 6–8 Hz during the time window of 200–600 ms poststimulus onset (Cohen & Donner, 2013; Hanslmayr et al., 2008; Jiang et al., 2015; Nigbur, Ivanova, & Stürmer, 2011) at the frontocentral channels (Fz, F1, F2, F3, F4, F5, F6, FCz, FC1, FC2, FC3, FC4, FC5, FC6, Cz, C1, C2, C3, C4, C5, C6). To investigate whether and how the theta oscillations driven by the cross-modal conflict was modulated by reward, t tests comparing incongruent with congruent trials were carried out for the reward and no-reward conditions separately. Dependent-sample t tests were conducted on theta oscillatory activities for each channel. Adjacent channels exceeding alpha level (0.05) were grouped into a cluster. The cluster-level statistic was calculated by taking the sum of the t values within the cluster. The number of random permutations using the Monte Carlo method was set to 5,000. Furthermore, we tested the interaction between congruency and reward by conducting a further cluster-based permutation t test comparing the differences between incongruent and congruent trials in the reward and no-reward conditions.

To examine the potential influence of reward information on prestimulus oscillatory activities (the reduced alpha band activities) reported in previous studies (Hughes, Mathan, & Yeung, 2013; Sawaki et al., 2015), the time-frequency analysis on alpha band was also applied to the time interval between -1,500 and 1,000 ms with the baseline time interval of -1,000 to -800 ms. Cluster-based permutation tests (Maris & Oostenveld, 2007) were conducted on alpha band activities across the frequency of 10–12 Hz during the time window of -800 to 0 ms prestimulus onset at the posterior channels (CPz, CP1, CP2, CP3, CP4, CP5, CP6, Pz, P1, P2, P3, P4, P5, P6, POz, PO3, PO4, PO7, PO8, Oz, O1, O2) to compare reward with no-reward trials. We chose this time window to ensure that there was a time interval long enough prior to the stimulus for the alpha band activity analysis while at the same time the categorization response to the previous trial was not included in this window.

3 | RESULTS

3.1 | Behavioral results

3.1.1 | Error rates

As shown in Figure 2 (right), ANOVA on error rates showed a main effect of reward type, $F(1, 24) = 7.538$, $p = .011$, $\eta_p^2 = .24$, suggesting that participants committed more errors in the reward condition than in the no-reward condition (3.9% vs. 2.8%). No other effect or interaction was found.

3.1.2 | RTs

As shown in Figure 2 (left), ANOVA on RTs showed a significant main effect of reward type, $F(1, 24) = 22.74$, $p < .001$, $\eta_p^2 = .49$, with shorter RTs in the reward condition than in the no-reward condition (495 vs. 560 ms). The main effect of stimulus type approached significance, $F(2, 48) = 2.86$, $p = .067$, $\eta_p^2 = .11$. However, the Bayes factor for this main effect ($BF_{10} = 0.084$) suggested that the null hypothesis was more likely to be true. Importantly, the interaction between reward and stimuli type was significant, $F(2, 48) = 4.089$, $p = .023$, $\eta_p^2 = .15$.

To explore the interaction, we conducted one-way ANOVA on RT with stimulus type as the within-participant factor in the reward and no-reward conditions, respectively. For the no-reward condition, there was a main effect of stimulus type, $F(2, 48) = 4.402$, $p = .018$, $\eta_p^2 = .16$. Pairwise comparisons revealed shorter RTs for the congruent condition than for the incongruent condition (553 vs. 569 ms; $p = .041$), and no RT differences in the visual-only condition as compared with in the congruent condition (560 vs. 553 ms; $p = .241$) or in the incongruent condition (560 vs. 569 ms; $p = .443$). For the reward condition, the ANOVA showed no significant main effect of stimulus type, $F(2, 48) = .19$, $p = .832$, $\eta_p^2 = .01$. These results suggested that the cross-modal conflict effect appeared in the no-reward condition, but not in the reward condition. Additionally, we analyzed the interaction from the other direction. The reaction times for reward trials were significantly faster than RTs for the no-reward trials in all the stimulus type conditions ($ps < .001$).

One may note that the overall reward effect (reward vs. no-reward) in terms of error rates showed a different pattern from the effect in terms of RT. Specifically, relative to the no-reward condition, participants committed more errors in the reward condition but responded much faster. This may indicate that participants weighted response speed more than accuracy in order to obtain reward. In other words, speed-accuracy trade-off that was induced by reward (Bijleveld, Custers, & Aarts, 2010). To test this hypothesis, we divided the trials into two RT bins according to the median RT in each experimental condition, and calculated the error rates in each RT

bin. A 2 (long vs. short RT bin) × 2 (reward vs. no-reward) × 3 (congruent vs. incongruent vs. visual-only) ANOVA on error rates showed main effects of RT bins, $F(1, 24) = 24.178$, $p < .001$, $\eta_p^2 = .502$, and reward, $F(1, 24) = 21.163$, $p < .001$, $\eta_p^2 = .469$. Moreover, the interaction between RT bins and reward was significant, $F(1, 24) = 8.737$, $p = .007$, $\eta_p^2 = .267$. No other main effect or interaction reached significance. The simple effect analysis showed that the error rate for reward trials was higher than for no-reward trials only in the short RT bin (6.1% vs. 3.4%, $t = 4.434$. $p < .001$), not in the long RT bin (1.5% vs. 1.1%, $t = .986$, $p = .334$). This pattern suggested that the reward-induced speed-accuracy trade-off occurred only in trials with fast responses.

To further investigate the potential influence of the speed-accuracy trade-off on the conflict effect, we conducted 2 (long vs. short RT bin) × 2 (reward vs. no-reward) × 3 (congruent vs. incongruent vs. visual-only) ANOVA on RTs as well. We hypothesized that if the reward-modulated conflict reduction observed here were simply a byproduct of the speed-accuracy trade-off caused by reward, the conflict reduction should have occurred only for the short RT bin. However, the ANOVA showed that there was no three-way interaction between RT bins, reward type, and stimuli type, $F(2, 48) = 1.588$, $p = .215$, $\eta_p^2 = .062$, suggesting that the reward modulation on the conflict effect occurred in both the short and long RT bins and the modulation could not simply be driven by the speed-accuracy trade-off.

3.2 | ERP results

3.2.1 | N2 (240–320 ms)

As shown in Figure 3, a 2 (Reward Type: reward vs. no-reward) × 2 (Stimulus Type: congruent vs. incongruent) × 5 (Electrode Cluster: frontal, frontocentral, central, parietocentral, parietal) repeated measures ANOVAs showed a main effect of reward type, $F(1, 24) = 7.983$, $p = .009$, $\eta_p^2 = .250$, with more negative N2 in the no-reward condition than in the reward condition (-2.57 vs. -1.71 μV), and a main effect of electrode cluster, $F(4, 96) = 96.419$, $p < .001$, $\eta_p^2 = .801$. The main effect of stimulus type was not significant, $F(1, 24) = 1.109$, $p = .303$, $\eta_p^2 = .044$). The interaction between reward type and stimulus type was marginally significant, $F(1, 24) = 4.085$, $p = .055$, $\eta_p^2 = .145$. Moreover, the three-way interaction was significant, $F(4, 96) = 4.380$, $p = .003$, $\eta_p^2 = .154$.

To investigate the three-way interaction, we conducted 2 (reward vs. no-reward) × 2 (congruent vs. incongruent) ANOVAs on N2 amplitudes in each electrode cluster, respectively. The two-way interaction was significant in frontal, $F(1, 24) = 7.738$, $p = .01$, $\eta_p^2 = .244$, and frontocentral, $F(1, 24) = 6.482$, $p = .018$, $\eta_p^2 = .213$, clusters, and marginally significant in the central cluster, $F(1, 24) = 3.344$, $p = .08$, with more negativ = .122, while it was absent in parietocentral,

$F(1, 24) = 0.978$, $p = .333$, with more negativ = .039, and parietal, $F(1, 24) = .090$, $p = .766$, with more negativ = .004, clusters. The Bayes factor for the model involving both main effects and interaction decreased from frontal to parietal clusters (frontal: $BF_{10} = 35.485$; frontocentral: $BF_{10} = 32.766$; central: $BF_{10} = 7.001$; parietocentral: $BF_{10} = 3.289$; parietal: $BF_{10} = 0.487$), suggesting that the null hypothesis could be true in the parietal region but was unlikely to be true in the frontal, frontocentral, and central regions. Planned t tests showed that N2 was more negative for the no-reward trials than for the reward trials only in the incongruent condition ($ps < .005$; frontal: -6.38 vs. -4.87 μV; frontocentral: -6.75 vs. -5.22 μV; central: -5.24 vs. -3.98 μV), but not in the congruent condition. These results indicated that reward affected the processing of incongruent audiovisual stimuli, but not the congruent audiovisual stimuli. To analyze the interaction in the other direction, we calculated the N2 amplitude difference between the incongruent and congruent trials (i.e., N2 conflict effect). Planned t tests showed a reduced N2 conflict effect in the reward condition as compared with the no-reward condition in frontal (-.687 vs. .421 μV, $t(24) = 2.785$, $p = .010$) and frontocentral (-.668 vs. .305 μV, $t(24) = 2.547$, $p = .018$) clusters, while the effect was marginally significant in the central cluster (-.552 vs. .072 μV, $t(24) = 1.830$, $p = .080$). These results suggested that the reduced N2 conflict by reward manifested mainly in the frontocentral regions.

3.2.2 | N400 (380–450 ms)

The ANOVA on N400 amplitude showed a main effect of stimulus type, $F(1, 24) = 21.906$, $p < .001$, with more negativ = .477, with more negative N400 for incongruent trials than for congruent trials (-1.912 vs. -.887 μV). The main effect of electrode cluster was significant, $F(4, 96) = 161.834$, $p < .001$, with more negativ = .871. Pairwise comparisons showed that the N400 decreased from frontal to parietal region (frontal: -6.000 μV; frontocentral: -5.546 μV; central: -3.150 μV; parietocentral: 1.433 μV; parietal: 6.264 μV). No other effect or interaction reached significance.

3.3 | Time-frequency analysis

3.3.1 | Theta band oscillations

The spectrogram and topographical distribution of theta band activities are shown in Figure 4. On the spectrogram, black lines encircle regions of significance clusters, corrected for multiple comparisons using cluster-based statistics. The cluster-based permutation-dependent t tests showed no significant clusters in the reward condition, whereas Fz, F1, F2, F4, F6, FCz, FC1, FC2, FC3, FC4, FC6, Cz, C1, C2, C3, C4, and C6 were grouped as a significant positive cluster (6–8 Hz, 200–600 ms) in the no-reward condition (cluster

PSYCHOPHYSIOLOGY SPR
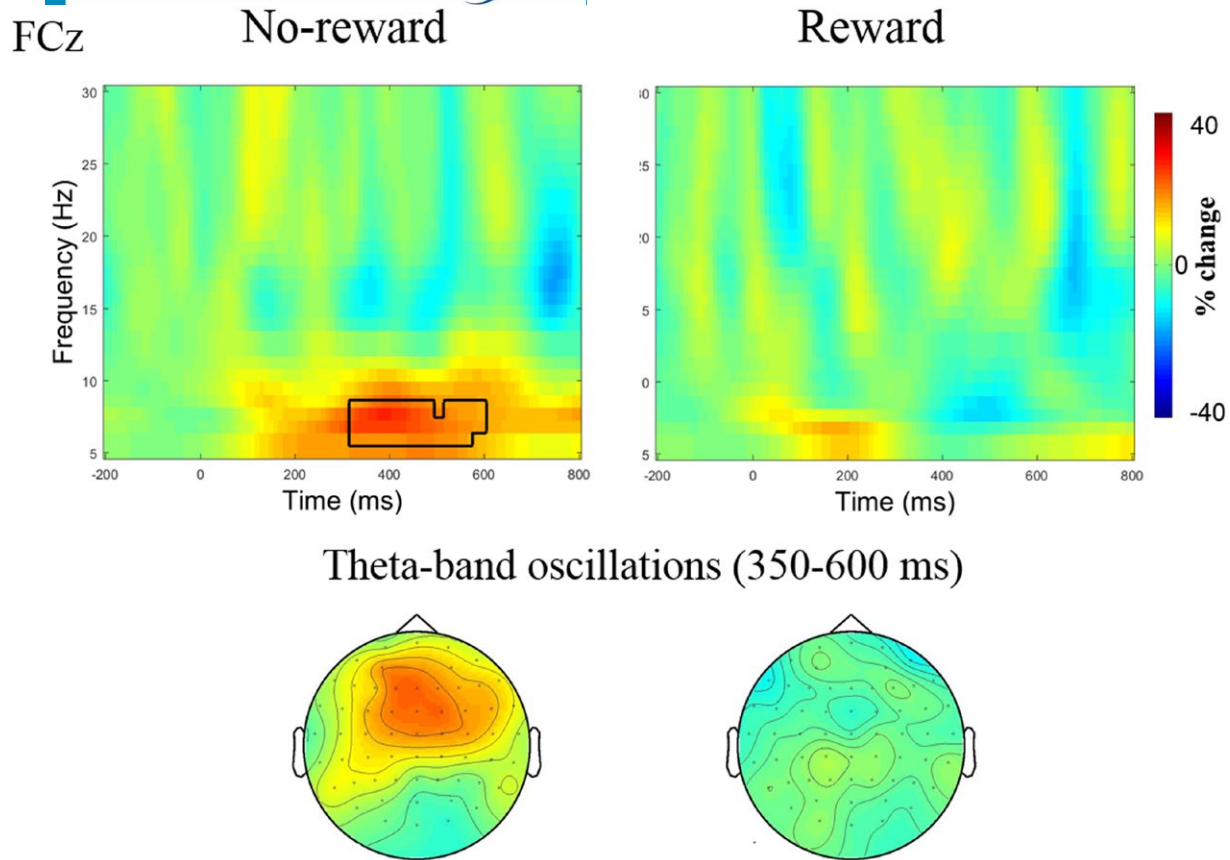
FIGURE 4 The difference in power change percentage of oscillatory activities relative to the prestimulus baseline (-200–0 ms) between incongruent and congruent audiovisual pairs in the no-reward condition (left) and the reward condition (right). Top: Oscillatory activities starting from 0 to 800 ms after stimulus onset at FCz. Black lines encircle regions of clusters showing significant differences between incongruent and congruent conditions (i.e., conflict effect), corrected for multiple comparisons using cluster-based statistics. There was a significant conflict effect in the no-reward condition, but not in the reward condition. Bottom: Topography of the average of theta oscillations (350–600 ms)

statistic $= 2719.28$, $p = .018$), suggesting that the frontal theta oscillation activities were stronger for the audiovisual incongruent trials than for the congruent trials in the no-reward condition (142.98 vs. 122.75% in power change). Further analysis for the interaction showed that Fz, F1, F2, F3, F5, FCz, FC1, FC2, FC3, FC5, Cz, C1, C2, C3, C4, and C6 were grouped as a significant negative cluster with frequency ranging from 6 to 8 Hz and time interval ranging from 350 to 600 ms after stimulus onset (cluster statistic $= -1479.68$, $p = .049$), suggesting that the difference between incongruent and congruent trials was smaller in the reward condition than in the no-reward condition (-7.1 vs. 18.61% in power change).

3.3.2 | Prestimulus alpha band oscillations

The cluster-based permutation-dependent t tests showed no significant clusters ($p > .1$), suggesting that reward expectation did not affect the prestimulus neural activity under the blockwise reward manipulation.

4 | DISCUSSION

The aim of the present study was to investigate the electrophysiological dynamics underlying the effect of reward on cross-modal conflict processing in object categorization. Behaviorally, we found that reward not only improved behavioral performance in general, but also reduced the cross-modal conflict effect when information from visual and auditory modalities were incompatible. This finding is consistent with the notion that reward enhances cognitive control and facilitates conflict resolution (Botvinick & Braver, 2015; Kang et al., 2017; Padmala & Pessoa, 2011; Soutschek et al., 2015). By combining ERP and time-frequency analyses, we showed that incongruent audiovisual stimulus pairs evoked larger N2, N400, and frontocentral theta band oscillation activities, as compared with congruent audiovisual pairs. Importantly, we found reward modulation on the conflict-related N2 and theta band effects, but not on the N400 effect. These results suggest that reward can enhance cognitive control in a cross-modal context and reduce cross-modal conflict in an early stage of cognitive processing.

The ERP analysis demonstrated a larger frontocentral N2 conflict effect in the no-reward condition as compared with the reward condition. Previous studies have shown that the frontocentrally distributed N2 peaking approximately 200–400 ms after stimulus onset is related to early conflict processing, response inhibition, or error monitoring (Donohue, Appelbaum, McKay, & Woldorff, 2016; Larson et al., 2014; Nigbur et al., 2011; van Veen & Carter, 2002; Yeung et al., 2004). For example, van Veen and Carter (2002) investigated the timing of ACC activity during conflict and error detection by using a flanker task. They found a frontocentral N2 component, which was more negative in the response incongruent condition as compared with the congruent condition. They also showed that ACC generated the N2, implicating conflict detection. By using a color flanker paradigm with negative and neutral words, Kanske and Kotz (2010) showed that the N2 component peaking at around 230 ms was larger for incongruent trials than for congruent trials and this N2 conflict effect was enhanced for negative words, as compared with neutral words. The authors suggested that N2 may reflect the amount of resources recruited for conflict detection. These results are consistent with previous EEG findings, suggesting that reward as a motivator modulates the early stage of processing in cognitive control (Kiss, Driver, & Eimer, 2009; Sawaki et al., 2015; Wei, Wang, & Ji, 2015). Sawaki et al. (2015), for example, examined the reward effect on attentional selection in visual search. For the target array, the N2pc was larger for the low-reward trials than for the reward trials, suggesting that reward increased efficiency of attentional orienting and reduced the need for focused attention.

It should be noted that the frontocentral N2 and the posterior N2 showed different patterns of reward modulation in the current study. Specifically, the frontocentral N2 showed an interaction between reward and congruency, indicating that this frontocentral N2 is closely related to conflict resolution and that this resolution process can be modulated by reward. By contrast, the interaction between reward and congruency was not observed for the posterior N2. Instead, relative to no-reward trials, the posterior N2 showed an overall decreased activity in reward trials, regardless of the congruency in the current trial. The differential patterns of frontocentral N2 and posterior N2 are consistent with the notion that different subcomponents of N2 are related to different cognitive functions (Folstein & Van Petten, 2008). Folstein and Van Petten (2008) proposed that the N2 component could be divided into three subcomponents: a frontocentral component indicating perceptual template mismatch, a second frontocentral component related to cognitive control, and a posterior N2 related to visual attention. Indeed, Suwazono, Machado, and Knight (2000) showed that the posterior N2 was larger for the target stimuli (presented at 20% of trials) than the standard stimuli, suggesting that the posterior N2 is related to attentional processing of the target stimuli. Given these arguments and the

different patterns of the frontocentral N2 and the posterior N2 in the current study, it is likely that, while the frontocentral N2 reflects the process of conflict control, the posterior N2 reflects an enhanced efficiency of attentional processing of the upcoming visual stimulus that can lead to reward. Thus, our results demonstrate that the differential cognitive functions indexed by different subcomponents of N2 can manifest in the cross-modal context.

The ERP analysis also showed a conflict-modulation on N400, with larger N400 amplitudes for incongruent trials than for congruent trials. A large number of previous studies have shown that N400 is implicated in semantic processing of objects in linguistic (Kutas & Federmeier, 2000) and pictorial (Ganis & Kutas, 2003) contexts. Previous studies also showed that N400 is related to cross-modal semantic matching (Molholm et al., 2004; Schneider, Debener, Oostenveld, & Engel, 2008; Sinke et al., 2014). For example, Molholm et al. (2004) investigated audiovisual object recognition processes by using images and vocalizations of animals. Consistent with the current study, the authors observed that incongruent audiovisual stimulus pairs elicited more negative ERP responses around 400 ms postonset as compared with congruent audiovisual stimulus pairs. This cross-modal conflict effect distributed over the centroparietal sites. Importantly, however, the current N400 semantic conflict effect was not affected by reward manipulation, in contrast with the reward modulation on the frontocentral N2 conflict effect. This contrast might be due to the nature or difficulty of the current task: categorization (animate vs. inanimate) does not need, and occurs earlier than, elaborated semantic processing (Grill-Spector & Kanwisher, 2005; Liu, Harris, & Kanwisher, 2002), which is indexed by N400. It is possible that reward modulation in conflict processing occurs very early. Nevertheless, it would be interesting to investigate whether reward modulation on the ERP components of conflict processing (N2 vs. N400) is critically dependent on the particular task type or task difficulty (e.g., categorization vs. identification).

Our time-frequency analysis revealed that the frontocentral oscillatory theta band power increased with conflict in the no-reward condition, but not in the reward condition. Previous studies suggested that theta band oscillations govern cognitive control processes, including conflict control, response inhibition, working memory, etc. (Cavanagh & Frank, 2014; Cohen, 2014; Cohen & Cavanagh, 2011; Cohen & Donner, 2013; Hanslmayr et al., 2008; Jiang et al., 2015; Nigbur et al., 2011; Raghavachari et al., 2001; Töllner et al., 2017; Yamanaka & Yamamoto, 2010). Nigbur et al. (2011), for example, investigated whether theta activity in the mediofrontal cortex indexes increased cognitive control demands in a Simon task, a flanker task, and a go/no-go task. They observed a significant conflict-modulation on theta activity (FCz, 200–300 ms) across all tasks; the sources of

PSYCHOPHYSIOLOGY SPR SOCIETY FOR PSYCHOPHYSIOLOGICAL RESEARCH

theta activity distributed along the rostral cingulate zone and premotor areas. Similar results were reported by Cohen and Donner (2013), which showed in a Simon task that the midfrontal (FCz) theta power was stronger in the high conflict condition than in the low conflict condition in the 200–600 ms interval after stimulus onset. This conflict modulation on theta band activity is also present in the Stroop task (Hanslmayr et al., 2008). Taken together, these findings suggest that theta band oscillations support conflict detection and resolution processes, and the enhanced theta band oscillations are related to increased recruitment of cognitive control (Hanslmayr et al., 2008; Nigbur et al., 2011).

Extending these findings, we observed conflict modulation on theta band oscillations in a cross-modal object categorization task. These conflict-related theta band oscillations were over frontocentral electrode sites during the 350–600 ms interval posttarget onset in the no-reward condition. These results indicate that the theta band oscillations are sensitive to control demands, and the enhanced theta band activities may indicate increased activation of the control system in resolving cross-modal conflict. Importantly, the conflict-modulated theta oscillation effect was not present in the reward condition. One of the important parts of control processes is the realization of control (Cavanagh & Frank, 2014); the enhanced theta band activities following stimulus presentation may fulfill this role by indicating a need for increased control (Cavanagh et al., 2012). The reduced conflict effect in terms of theta oscillations in the reward condition demonstrates the reward-enhanced cognitive control in the face of cross-modal conflicts.

The theta band oscillation showed almost the same pattern as the frontocentral N2. This similarity may indicate a functional connection between the frontal theta oscillation and the frontocentral N2. It has been suggested that N2 has a spectral signature in theta band (4–8 Hz) (for a review, see Cavanagh & Frank, 2014), and there might be functional overlap between theta power and N2 in cognitive control (Nigbur et al., 2011). Our results support this proposal.

As the results suggest that reward modulates poststimulus N2 and theta band oscillation conflict effects, but not prestimulus oscillatory alpha activity, one potential mechanism is that the participants were better at filtering out the information in the task-irrelevant sensory modality and resolving conflict in a reactive manner in the reward condition. The dual mechanisms of control (Braver, 2012) suggest that there are two distinct manners for cognitive control: proactive control and reactive control. In proactive control, the goal-directed information is early selected and maintained; in reactive control, attentional control is recruited in a just-in-time manner (Braver, 2012). Previous studies using a trial-by-trial reward manipulation (e.g., Sawaki et al. 2015; van den Berg et al., 2014) have shown that reward modulates the neural activity (alpha band oscillations) postcue onset but before target onset, suggesting an enhanced preparatory control for the subsequent task. In the current study, we did not observe reward modulation on prestimulus neural activity but only conflict effects on N2 and theta band oscillations posttarget onset, suggesting that the impact of reward on control is implemented only after a conflict is detected. These results imply that the blockwise reward manipulation may lead to an increased tendency of control in a reactive manner.

It is well established in the visual domain that reward can facilitate conflict resolution, and this reward modulation is associated with certain neural signatures such as N2, theta oscillation, and the BOLD activity in ACC (Krebs, Boehler, Appelbaum, & Woldorff, 2013; Padmala & Pessoa, 2011; Soutschek et al., 2015). Behaviorally, our results showed a similar reduced conflict caused by reward in an audiovisual context, extending the reward modulation into the cross-modal domain. At the neural level, our results showed that the reward modulation on conflict processing is also related to the frontocentral N2 and theta oscillation. Taking together the current results and the mounting evidence from previous studies in the visual domain, we suggest a supramodal role of reward in modulating cognitive control. This suggestion asks for further evidence based on direct comparisons between unimodal and cross-modal contexts in a single experiment.

To conclude, the present study investigated how reward modulates cross-modal conflict control in object categorization. Behaviorally, reward improved behavioral performance and reduced the cross-modal conflict effect in response times. Neurally, reward reduced the conflict effect on N2 and theta band oscillations. More detailed analyses showed that reward enhanced cross-modal conflict control mainly by facilitating the processing of distractors in the task-irrelevant sensory modality. Thus, reward can enhance cognitive control in a cross-modal context and reduce cross-modal conflict.

ACKNOWLEDGMENTS

ORCID

Ping Wei (iD) http://orcid.org/0000-0002-3781-3844
Xiaolin Zhou (iD) http://orcid.org/0000-0001-7363-4360

REFERENCES

Bijleveld, E., Custers, R., & Aarts, H. (2010). Unconscious reward cues increase invested effort, but do not change speed–accuracy

tradeoffs. *Cognition*, *115*(2), 330–335. https://doi.org/10.1016/j.cognition.2009.12.012

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: From behavior to neural mechanism. *Annual Review of Psychology*, *66*(1), 83–113. https://doi.org/10.1146/annurev-psych-010814-015044

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106–113. https://doi.org/10.1016/j.tics.2011.12.010

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology*, *49*(2), 220–238. https://doi.org/10.1111/j.1469-8986.2011.01293.x

Cohen, M. X. (2014). A neural microcircuit for cognitive conflict detection and signaling. *Trends in Neurosciences*, *37*(9), 480–490. https://doi.org/10.1016/j.tins.2014.06.004

Cohen, M. X., & Cavanagh, J. F. (2011). Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. *Frontiers in Psychology*, *2*(30), 1–12. https://doi.org/10.3389/fpsyg.2011.00030

Cohen, M. X., & Donner, T. H. (2013). Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior. *Journal of Neurophysiology*, *110*(12), 2752–2763. https://doi.org/10.1152/jn.00479.2013

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45. https://doi.org/10.3758/s13423-012-0230-1

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Diaconescu, A. O., Alain, C., & McIntosh, A. R. (2011). The co-occurrence of multisensory facilitation and cross-modal conflict in the human brain. *Journal of Neurophysiology*, *106*(6), 2896–2909. https://doi.org/10.1152/jn.00303.2011

Donohue, S. E., Appelbaum, L. G., McKay, C. C., & Woldorff, M. G. (2016). The neural dynamics of stimulus and response conflict processing as a function of response complexity and task demands. *Neuropsychologia*, *84*, 14–28. https://doi.org/10.1016/j.neuropsychologia.2016.01.035

Faivre, N., Mudrik, L., Schwartz, N., & Koch, C. (2014). Multisensory integration in complete unawareness: Evidence from audiovisual congruency priming. *Psychological Science*, *25*(11), 2006–2016. https://doi.org/10.1177/0956797614547916

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, *45*(1), 152–170. https://doi.org/10.1111/j.1469-8986.2007.00602.x

Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology*, *2*, 154. https://doi.org/10.3389/fpsyg.2011.00154

Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Cognitive Brain Research*, *16*(2), 123–144. https://doi.org/10.1016/S0926-6410(02)00244-6

Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science*, *16*(2), 152–160. https://doi.org/10.1111/j.0956-7976.2005.00796.x

Hanslmayr, S., Pastötter, B., Bäuml, K. H., Gruber, S., Wimber, M., & Klimesch, W. (2008). The electrophysiological dynamics of interference during the Stroop task. *Journal of Cognitive Neuroscience*, *20*(2), 215–225. https://doi.org/10.1162/jocn.2008.20020

Haupt, S., Axmacher, N., Cohen, M. X., Elger, C. E., & Fell, J. (2009). Activation of the caudal anterior cingulate cortex due to task-related interference in an auditory Stroop paradigm. *Human Brain Mapping*, *30*(9), 3043–3056. https://doi.org/10.1002/hbm.20731

Hughes, G., Mathan, S., & Yeung, N. (2013). EEG indices of reward motivation and target detectability in a rapid visual detection task. *NeuroImage*, *64*(1), 590–600. https://doi.org/10.1016/j.neuroimage.2012.09.003

Jiang, J., Zhang, Q., & van Gaal, S. (2015). Conflict awareness dissociates theta-band neural dynamics of the medial frontal and lateral frontal cortex during trial-by-trial cognitive control. *NeuroImage*, *116*, 102–111. https://doi.org/10.1016/j.neuroimage.2015.04.062

Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, *111*(10), 1745–1758. https://doi.org/10.1016/S1388-2457(00)00386-2

Kang, G., Wang, L., & Zhou, X. (2017). Reward interacts with modality shift to reduce cross-modal conflict. *Journal of Vision*, *17*(1), 1–14. https://doi.org/10.1167/17.1.19

Kanske, P., & Kotz, S. A. (2010). Modulation of early conflict processing: N200 responses to emotional words in a flanker task. *Neuropsychologia*, *48*(12), 3661–3664. https://doi.org/10.1016/j.neuropsychologia.2010.07.021

Kiss, M., Driver, J., & Eimer, M. (2009). Reward priority of visual target singletons modulates event-related potential signatures of attentional selection. *Psychological Science*, *20*(2), 245–251. https://doi.org/10.1111/j.1467-9280.2009.02281.x

Krebs, R. M., Boehler, C. N., Appelbaum, L. G., & Woldorff, M. G. (2013). Reward associations reduce behavioral interference by changing the temporal dynamics of conflict processing. *PLOS One*, *8*(1), e53894. https://doi.org/10.1371/journal.pone.0053894

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6

Larson, M. J., Clayson, P. E., & Clawson, A. (2014). Making sense of all the conflict: A theoretical review and critique of conflict-related ERPs. *International Journal of Psychophysiology*, *93*(3), 283–297. https://doi.org/10.1016/j.ijpsycho.2014.06.007

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience*, *5*(9), 910–916. https://doi.org/10.1038/nn909

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Molholm, S., Ritter, W., Javitt, D. C., & Foxe, J. J. (2004). Multisensory visual-auditory object recognition in humans: A high-density electrical mapping study. *Cerebral Cortex*, *14*(4), 452–465. https://doi.org/10.1093/cercor/bhh007

Nigbur, R., Ivanova, G., & Stürmer, B. (2011). Theta power as a marker for cognitive interference. *Clinical Neurophysiology*, *122*(11), 2185–2194. https://doi.org/10.1016/j.clinph.2011.03.030

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 1–9. https://doi.org/10.1155/2011/156869

Padmala, S., & Pessoa, L. (2011). Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *Journal of Cognitive Neuroscience*, *23*(11), 3419–3432. https://doi.org/10.1162/jocn_a_00011

Raghavachari, S., Kahana, M. J., Rizzuto, D. S., Caplan, J. B., Kirschen, M. P., Bourgeois, B., … Lisman, J. E. (2001). Gating of human theta oscillations by a working memory task. *Journal of Neuroscience*, *21*(9), 3175–3183. https://doi.org/10.1523/JNEUROSCI.21-09-03175.2001

Roberts, K. L., & Hall, D. A. (2008). Examining a supramodal network for conflict processing: A systematic review and novel functional magnetic resonance imaging data for related visual and auditory stroop tasks. *Journal of Cognitive Neuroscience*, *20*(6), 1063–1078. https://doi.org/10.1162/jocn.2008.20074

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Sawaki, R., Luck, S. J., & Raymond, J. E. (2015). How attention changes in response to incentives. *Journal of Cognitive Neuroscience*, *27*(11), 2229–2239. https://doi.org/10.1162/jocn_a_00847

Schneider, T. R., Debener, S., Oostenveld, R., & Engel, A. K. (2008). Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *NeuroImage*, *42*(3), 1244–1254. https://doi.org/10.1016/j.neuroimage.2008.05.033

Sinke, C., Neufeld, J., Wiswede, D., Emrich, H. M., Bleich, S., Münte, T. F., & Szycik, G. R. (2014). N1 enhancement in synesthesia during visual and audio-visual perception in semantic cross-modal conflict situations: An ERP study. *Frontiers in Human Neuroscience*, *8*, 21. https://doi.org/10.3389/fnhum.2014.00021

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174–215. http://dx.doi.org/10.1037/0278-7393.6.2.174

Soutschek, A., Stelzel, C., Paschke, L., Walter, H., & Schubert, T. (2015). Dissociable effects of motivation and expectancy on conflict processing: An fMRI study. *Journal of Cognitive Neuroscience*, *27*(2), 409–423. https://doi.org/10.1162/jocn_a_00712

Spagna, A., Mackie, M. A., & Fan, J. (2015). Supramodal executive control of attention. *Frontiers in Psychology*, *6*, 65. https://doi.org/10.3389/fpsyg.2015.00065

Suwazono, S., Machado, L., & Knight, R. T. (2000). Predictive value of novel stimuli modifies visual event-related potentials and behavior. *Clinical Neurophysiology*, *111*(1), 29–39. https://doi.org/10.1016/S1388-2457(99)00186-8

Töllner, T., Wang, Y., Makeig, S., Müller, H. J., Jung, T.-P., & Gramann, K. (2017). Two independent frontal midline theta oscillations during conflict detection and adaptation in a Simon-type manual reaching task. *Journal of Neuroscience*, *37*(9), 2504–2515. https://doi.org/10.1523/JNEUROSCI.1752-16.2017

van den Berg, B., Krebs, R. M., Lorist, M. M., & Woldorff, M. G. (2014). Utilization of reward-prospect enhances preparatory attention and reduces stimulus conflict. *Cognitive, Affective & Behavioral Neuroscience*, *14*(2), 561–577. https://doi.org/10.3758/s13415-014-0281-z

van Veen, V., & Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology and Behavior*, *77*, 477–482. https://doi.org/10.1016/S0031-9384(02)00930-7

Vogler, J. N., & Titchener, K. (2011). Cross-modal conflicts in object recognition: Determining the influence of object category. *Experimental Brain Research*, *214*(4), 597–605. https://doi.org/10.1007/s00221-011-2858-x

Wei, P., Wang, D., & Ji, L. (2015). Reward expectation regulates brain responses to task-relevant and task-irrelevant emotional words: ERP evidence. *Social Cognitive and Affective Neuroscience*, *11*(2), 191–203. https://doi.org/10.1093/scan/nsv097

Weissman, D. H., Warner, L. M., & Woldorff, M. G. (2004). The neural mechanisms for minimizing cross-modal distraction. *Journal of Neuroscience*, *24*(48), 10941–10949. https://doi.org/10.1523/JNEUROSCI.3669-04.2004

Weissman, D. H., Warner, L. M., & Woldorff, M. G. (2009). Momentary reductions of attention permit greater processing of irrelevant stimuli. *NeuroImage*, *48*(3), 609–615. https://doi.org/10.1016/j.neuroimage.2009.06.081

Yamanaka, K., & Yamamoto, Y. (2010). Single-trial EEG power and phase dynamics associated with voluntary response inhibition. *Journal of Cognitive Neuroscience*, *22*(4), 714–727. https://doi.org/10.1162/jocn.2009.21258

Ye, Z., & Zhou, X. (2009). Conflict control during sentence comprehension: fMRI evidence. *NeuroImage*, *48*, 280–290. https://doi.org/10.1016/j.neuroimage.2009.06.032

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*(4), 931–959. https://doi.org/10.1037/0033-295X.111.4.939

Yuval-Greenberg, S., & Deouell, L. Y. (2009). The dog's meow: Asymmetrical interaction in cross-modal object recognition. *Experimental Brain Research*, *193*(4), 603–614. https://doi.org/10.1007/s00221-008-1664-6

Zimmer, U., Itthipanyanan, S., Grent-'t-Jong, T., & Woldorff, M. G. (2010). The electrophysiological time course of the interaction of stimulus conflict and the multisensory spread of attention. *European Journal of Neuroscience*, *31*(10), 1744–1754. https://doi.org/10.1111/j.1460-9568.2010.07229.x