

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

Guanlan Kang<sup>1</sup> | Wenshuo Chang<sup>1</sup> | Lihui Wang<sup>1</sup> | Ping Wei<sup>2</sup>  | Xiaolin Zhou<sup>1,3,4,5,6</sup> 

1

## 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

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 audiovisual 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 (

(matal())TJ0.07 To1 (xts)27 ( )T2 0.05-6.50 (-1.3 -1.3 1.9 (w)"/(5rent(his s-10 (r)-c.9 (at[(t)disc Exper)-1iningueing wit)9 ( and  
he 18rl in toss te

## 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  $7^\circ \times 7^\circ$  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



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  $\pm 70 \mu\text{V}$  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 transforma-

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 ( $p < .001$ ).

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)  $\times$  2 (reward vs. no reward)  $\times$  3

statistic = 2719.28,  $p = .018$ ), suggesting that the frontal theta



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 (t)-1ve4) ERng tic (K1e4) \*[Pal2 (w)18 (ibuted )evi7y ]

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

- 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 Neurology*, *250*(1), 1–12. <https://doi.org/10.1007/s00415-012-2011-2>

