



OPEN

Electrophysiological evidence of different neural processing between visual and audiovisual inhibition of return

Xiaoyu Tang^{1✉}, Xueli Wang¹, Xing Peng^{2✉}, Qi Li³, Chi Zhang⁴, Aijun Wang^{5✉} & Ming Zhang^{5✉}

Inhibition of return (IOR) refers to the slower response to targets appearing on the same side as the cue (valid locations) than to targets appearing on the opposite side as the cue (invalid locations). Previous behaviour studies have found that the visual IOR is larger than the audiovisual IOR when focusing on both visual and auditory modalities. Utilising the high temporal resolution of the event-related potential (ERP) technique we explored the possible neural correlates with the behaviour IOR difference between visual and audiovisual targets. The behavioural results revealed that the visual IOR was larger than the audiovisual IOR. The ERP results showed that the visual IOR effect was generated from the P1 and N2 components, while the audiovisual IOR effect was derived only from the P3 component. Multisensory integration (MSI) of audiovisual targets occurred on the P1, N1 and P3 components, which may offset the reduced perceptual processing due to audiovisual IOR. The results of early and late differences in the neural processing of the visual IOR and audiovisual IOR imply that the two target types may have different inhibitory orientation mechanisms.

In a spatial cueing paradigm, the response to targets at the same side as the cue (also called the valid locations) is faster than that at the opposite side as the cue (invalid locations) when the cue and target stimulus onset asynchrony (SOA) is less than 300 ms. By contrast, when the SOA is more than 300 ms, the response to targets presented at valid locations is slower than that presenting at invalid locations. The latter phenomenon has been termed inhibition of return (IOR)¹, and is described as an orientation mechanism to improve the efficiency of visual search by preventing the reinspection of already searched locations^{2,3}.

IOR is observed within the visual⁴, auditory⁵, or tactile modalities⁶, suggesting that it may be a more general, than inhibitory process specific to unimodal search. Recently, researchers have extended the view to bimodal targets. Tang et al.⁷ investigated the IOR effect induced by visual, auditory and audiovisual targets. Their results showed a comparable IOR effect with audiovisual and visual targets when participants were asked to only focus on visual modality. However, they found that the magnitude of the IOR effect with bimodal audiovisual targets was smaller than that with unimodal visual targets when visual and auditory modalities were both attended (i.e., the divided attention condition). They also found that the simultaneous occurrence of visual and auditory targets was integrated (i.e., multisensory integration, MSI). The integrated bimodal stimuli can help improve signal salience⁸ and attract more attention^{9,10}. The IOR actually reflect that attention is blocked from returning a previously attended location¹¹ and facilitates visual exploration by reducing the perceptual salience of a previously explored location¹². Thus, Tang et al. (2019) hypothesised that the integration of audiovisual targets enhances perceptual salience and outcompetes the reduced perceptual processing because of IOR, so that the IOR effect of audiovisual targets is decreased. No direct neural evidence has proven this hypothesis, although previous studies may indirectly support it.

On the one hand, multisensory integration and IOR seem to have common neural correlates. Tanaka et al.¹³ demonstrated that the superior colliculus (SC) has multisensory neurons responsive to visual, auditory, and

¹School of Psychology, Liaoning Collaborative Innovation Center of Children and Adolescents Healthy Personality Assessment and Cultivation, Liaoning Normal University, Dalian 116029, China. ²Institute of Aviation Human Factors and Ergonomics, Civil Aviation Flight University of China, Guanghan 618307, China. ³School of Computer Science and Technology, Changchun University of Science and Technology, Changchun 130022, China. ⁴School of Biomedical Engineering, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China. ⁵Department of Psychology, Soochow University, Suzhou 215123, China. ✉email: tangyu-2006@163.com; fypx3688@163.com; psy_waj@126.com; psyzm@suda.edu.cn

tactile stimuli, but a suppressive interaction also exists in SC neurons. Macaluso and Driver¹⁴ showed that similar areas or even similar cells in subcortical areas or primary sensory cortices are responsible for both multisensory integration and crossmodal attention. Multisensory integration and attentional inhibition occur in similar brain areas, providing a possibility for spatial interplay between multisensory integration and IOR. On the other hand, multisensory integration can occur at multiple stages¹⁵. For example, early multisensory integration was found over the left parieto-occipital cortex (50-to-60 ms post-stimulus onset), sometimes followed by early positive modulation (80–100 ms) over occipital and temporal areas contralateral to audiovisual targets¹⁶. Multisensory integration also appeared at a time window of N2pc¹⁷ and even in later P3 stages^{18,19}. Notably, many studies have found IOR during the abovementioned time windows^{20–23}. The spatial and temporal overlaps of multisensory integration and IOR raise the possibility for their interaction at different levels. We hypothesised that at some time windows, particularly for some early ERP components, multisensory integration reduces IOR when these two effects appear simultaneously, or even eliminate IOR if MSI is sufficiently strong to compete with the decreased perceptual processing due to IOR.

Another more crucial question that must be considered is whether the time windows of IOR effects for unimodal visual targets and bimodal audiovisual targets are different due to the influence of multisensory integration or physical factors. Previous ERP studies have shown that the suppression of perceptual processing in extrastriate visual areas is the most common inhibition associated with the visual IOR effect—namely, a reduced amplitude of the earlier P1/N1 component at the valid locations compared with that at the invalid locations^{22,24–26}. IOR is also linked to the visual N2 component with a frontal/central^{27,28} or a posterior scalp distribution²⁹ involved in inhibitory executive processes—similarly, the amplitude is reduced at the valid locations^{30,31}. Additionally, a significant IOR effect was reported occasionally in association with a special P3 enhancements for valid locations compared with that for invalid locations^{6,23}, although this effect has not very often been reported^{19,24,32}. In different tasks, these ERP components may contribute to the visual IOR separately or collectively, and corresponding theories involve early perceptual inhibition^{25,33,34}, late motor inhibition^{35,36} or dual-component interpretation^{21,37–39}.

The neural activities of audiovisual stimuli are not equal to the neural activities of auditory or visual stimuli at some stages⁴⁰. P3 reflects the updating of working memory during the postdetection process⁴¹, it was viewed as an index of context updating⁴² with larger amplitudes elicited by an increased amount of engaged attentional resources and/or processing capacity⁴³. Audiovisual targets are novel stimuli compared with previous visual cues, which can attract more attention⁴⁴ or may lead to updated working memory after these target stimuli are detected. Visual targets have no modality novelty relative to visual cues, and the IOR is mainly related to the inhibition in early visual perception⁴⁵. The audiovisual targets might elicit a late IOR effect associated with the P3 component, and significant P1, and N1 inhibitory components should be found in visual targets. Because multisensory integration is more likely to occur in the early processing^{46–48} stages, it may cause the early P1 and N1 to be reduced or even disappear in audiovisual targets. Clarifying the problem of time windows may contribute to a better understanding of the behaviour difference between the auditory and visual IOR.

The primary goal of this study was to examine whether the enhanced perceptual salience induced by multisensory integration is the cause of the audiovisual IOR being smaller than the visual IOR. Considering that behavioural indicators (reaction time and accuracy) of the IOR effect usually reflect the sum of all information processing stages, the processing differences between visual IOR and audiovisual IOR cannot be unequivocally clarified from the internal mechanism. Therefore, we adopted a cue-target paradigm, using event-related potential (ERP) technique in the present study, to examine the neural correlates with the behaviour IOR differences between target types. In this paradigm, a thickened white box appears on the left or right side of the fixation as an exogenous cue that triggers exogenous spatial orientation. To facilitate the occurrence of an IOR, a central reorientation method is used, with the previous fixation extended to focus attention back to the centre. We manipulated the location validity (valid and invalid), and stimuli type (visual, auditory, and audiovisual stimuli) and asked participants to focus on both auditory and visual information simultaneously and then perform a target/non-target discrimination task (responding to the target, and withholding responding to the non-target). Finally, we compared the accuracy (ACC), mean reaction time (RT) or ERP amplitude of the target stimuli at the valid locations with those at the invalid locations. As explained earlier, we hypothesised that the enhanced perceptual salience induced by multisensory integration outcompetes the reduced perceptual processing partially because of IOR, leading to a decrease in the early IOR effect on audiovisual targets, that will be reflected in neural processes.

Method

Participants. We calculated the appropriate sample size based on the G*Power toolbox⁴⁹ and work of Van der Stoep⁵⁰. To achieve the recommended 95% statistical power at $\alpha = 0.05$, and an effect size of 0.70, the appropriate sample size was defined as at least 12 participants. Twenty undergraduate students (mean age = 20.33 years; SD = 2; 17 women, 3 men; 1 left-handed) were eventually recruited to participate in the current study. All the students had normal or corrected-to-normal vision and normal hearing capabilities, and had no neurological or psychiatric disorders. This study was approved by the Ethics Committee of Liaoning Normal University and was performed in accordance with the approved guidelines and the Declaration of Helsinki. All the participants provided written informed consent before participating.

Stimuli. The experiment was conducted in a dimly illuminated, sound-attenuated room. The participants were seated directly in front of a computer monitor, with their heads held steady by a chin rest. All visual (V) stimuli were presented on a black background display (with a screen resolution of 1024 × 768 black and a refresh rate of 100 Hz) 60 cm from the participant (see Fig. 1). The visual target was a checkerboard with two dots, and the visual non-target was a checkerboard without dots. Auditory (A) stimuli were presented via two speakers

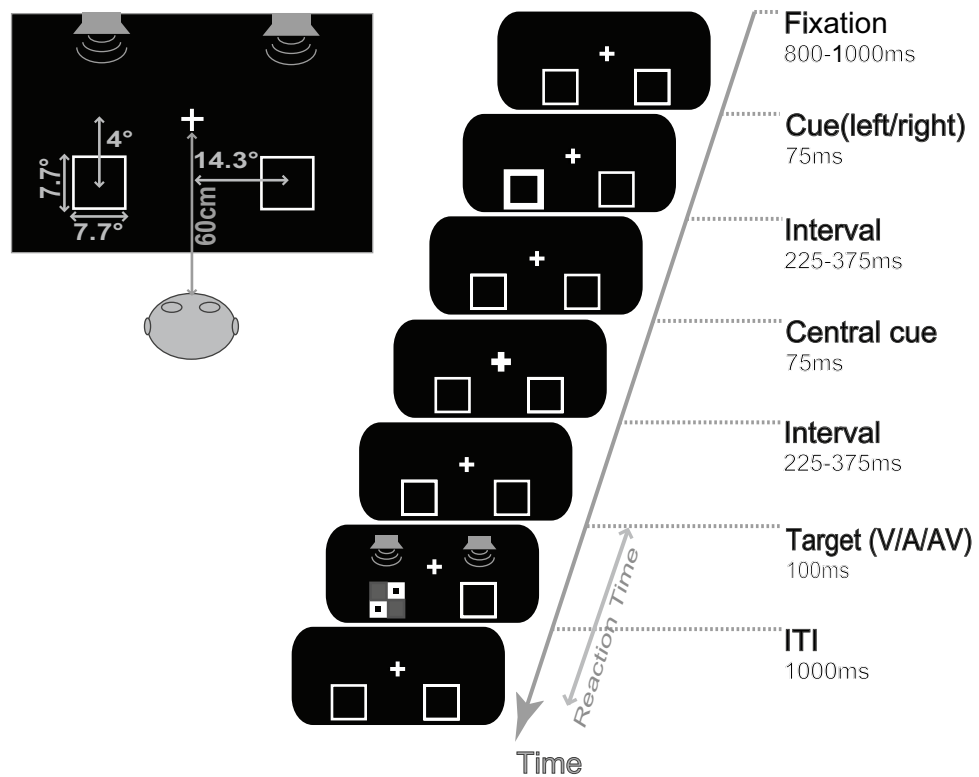


Figure 1. Illustration of the stimuli and experimental procedure. The size and position of the stimuli are shown in the left panel. The sequence of events and their duration (starting from the top) in the valid location of the visual (V) target condition are illustrated in the right panel.

placed at both the left rear and right rear of the display. White noise was set as the auditory target (65 dB, rise/fall times of 10 ms) and another pure tone at 1600 Hz (65 dB, rise/fall times of 10 ms) was set as auditory non-target. Audiovisual (AV) stimulus comprised a combination of the simultaneous presentation of both auditory and visual target/non-target. For example, a checkerboard with two dots and white noise comprised an audiovisual target, a checkerboard without two dots and a pure tone comprised an audiovisual non-target. The fixation point ($0.05^\circ \times 0.05^\circ$ of the visual angle) was a small white cross (RGB, 255, 255, 255; 155.2 cd/m^2) flanked by two white square outline boxes ($7.7^\circ \times 7.7^\circ$), presented 4° below and 14.3° to the left and right of the fixation across.

Procedure and task. The experimental procedure is illustrated in Fig. 1. Each trial started with a fixation stimulus that appeared for 800–1000 ms. Subsequently, one of the white square boxes became thicker as a peripheral visual cue, which lasted for 75 ms. After a random first interval of 225–375 ms, the fixation cross was extended to $0.1^\circ \times 0.1^\circ$ to serve as a reorienting cue (lasting 75 ms) to bring focused attention back to the centre. Next, the second interval randomly appeared again for 225–375 ms. Thus, the SOAs were randomly assigned between 600 and 900 ms to help resist the temporal expectation of the appearance of the targets. Next, the A, V or AV targets/non-targets were randomly presented for 100 ms on the left/right side. Finally, the fixation stimulus appeared for 1000 ms to allow the participants to make reaction, after which the next trial started automatically. The participants were required to focus on both the visual and auditory modalities simultaneously. They performed a discrimination task in which they were requested to judge whether the stimulus was a target or a non-target. The participants were instructed to press "B" on the keyboard as quickly and accurately as possible when they detected an auditory, a visual, or an audiovisual target, but withhold their response when a non-target stimulus appeared in the left or right location. During the experiment, the participants were instructed to reduce their actions of swallowing, frowning, and blinking and to, keep their bodies from moving. Matters requiring attention were only target stimuli trials that were used in further analyses.

Design. Two factors of location validity and target type were manipulated in the behaviour experiment. Location validity involved two levels: valid and invalid. Half of the trials in each block involved valid locations, and half involved invalid locations. There were three target types: A targets, V targets and AV targets. Before the formal experiment, each participant needed to complete 60 practice trials first to ensure that the instructions were clear and that the task was performed correctly. The formal experiment comprised 10 blocks of 1200 trials: 960 target trials (80%) and 240 non-target trials (20%). There were 6 experimental conditions and 200 trials contained for each; 160 were target trials, and 40 were non-target trials. All the conditions contained an equal number of A, V, and AV stimuli trials. Within each block, the participants were asked to respond to the

targets, and their EEG data were recorded simultaneously. Feedback about the number of correct responses was provided at the end of each block, and rest periods were provided between blocks. The order in which the participants completed the target condition blocks was randomised and counterbalanced across the participants.

Behavioural data recording and analysis. E-prime software (1.0 version) was used to present the stimuli and record the responses. Trials with RTs faster than 100 ms (anticipations) or slower than 1000 ms (misses) were counted as errors. These error trials and trials with incorrect responses were excluded from analysis. Thus, 2% of the target trials were removed from further analysis. The mean reaction time (RT) and accuracy (ACC) were entered into a 3 (target type: A, V, AV) \times 2 (location validity: valid and invalid) repeated measures ANOVA. Paired *t* test (two-tailed) was used to compare the IOR effect (valid vs. invalid) in different target modalities. All statistical significance levels were set at 0.05. The effect size of Cohen's *d* or partial eta-squared (η_p^2) was calculated for mean comparisons or ANOVA, respectively.

ERP data recording and analysis. An EEG system (Brain Products, Brain Vision Recorder 2.0) was used to record EEG signals using 32 electrodes mounted on an electrode cap as specified by the International 10–20 system, with the reference electrode on the left earlobe and a ground electrode on the medial frontal aspect. A horizontal electrooculogram (EOG) was recorded from the outer canthi of the left eye, and a vertical EOG was recorded from an electrode placed 1.5 cm above and below the left eye. The raw signals were digitised with a sample frequency of 500 Hz, and all impedances were kept below 5 k Ω .

Offline, an EEG system (Brain Products, Brain Vision Analyzer 2.0) was used to analyse ERP signals. The data were digitally filtered with a high-pass filter of 0.1 Hz and a low-pass filter of 30 Hz (24 dB/octave)⁵¹. The data were segmented into 1000 ms epochs from –200 to 800 ms. Artefact rejection was performed using a semi-automated procedure⁵² to remove epochs that contained eye movements and blinks from EEG. Also, trials with excessive artefacts ($\pm 80 \mu\text{V}$) were excluded from further analysis. Finally, the data from each electrode were averaged and a grand average ERP was computed across all participants for each experimental condition.

According to previous studies, the present ERP components were divided by the time windows in which they occurred: P1, 90–140 ms; N1, 150–190 ms; N2, 220–290 ms; P3, 350–400 ms. We selected electrodes P3, P4, P7, P8, CP5, CP6, F3, F4, F7, F8, FC5, FC6, and Cz⁵³ in the lateral and frontal, central, and occipital areas for statistical analyses of P1, N1, N2, and P3. In each time window, the mean amplitudes were analysed using repeated measures ANOVA with factors of location validity and electrode.

Results

Behavioural results. Accuracy rate analysis showed that only the main effect of the target type was significant [$F(2, 38) = 7.328$; $p < 0.01$; $\eta_p^2 = 0.278$]. The response to the AV targets (97%) was more accurate than those to the A targets (95%) and V targets (95%). No other significant main effects or interactions were found. For the mean correct RTs, ANOVA revealed a significant main effect of the target type [$F(2, 38) = 29.194$; $p < 0.001$; $\eta_p^2 = 0.606$], and the RTs to AV targets (408 ms) were significantly shorter than those to V (468 ms) and A targets (453 ms). The main effect of the location validity was significant [$F(1, 19) = 45.747$, $p < 0.001$, $\eta_p^2 = 0.707$], reflecting faster RTs at the invalid locations (437 ms) than at the valid locations (448 ms). Crucially, the interaction between the target type and location validity was significant [$F(2, 38) = 18.828$; $p < 0.001$; $\eta_p^2 = 0.498$]. The difference in RTs between valid and invalid locations was significant for the V targets [$t(19) = 10.23$; $p < 0.01$; $d = 0.360$] and the AV targets [$t(19) = 7.04$; $p < 0.001$; $d = 0.164$], but not for the A targets [$t(19) = 0.49$; $p = 0.628$; $d = 0.028$]. Further analysis found that the IOR effect of the V targets (21 ms) was larger than that of the AV targets (10 ms) [$t(19) = -5.75$; $p < 0.001$; $d = 1.362$].

ERP results. Consistent with the behavioural results of previous studies, no auditory IOR was found. We still observed that the visual and audiovisual IOR effects were different in magnitude. The reasons for the differences are not clear at present. Thus, we provide a neural process comparison across the A targets, V targets and AV targets. The mean amplitudes were subjected to a 2 (location validity: valid and invalid) \times 13 (electrode) two factor repeated measures ANOVA.

Event-related potential of the auditory IOR effect: (A valid vs. A invalid). *P1 component (90–140 ms).* Figure 2a,c shows the P1 waveforms and scalp topographies of the A targets, respectively. The analysis yielded a significant main effect for the electrode [$F(12, 228) = 11.480$; $p < 0.001$; $\eta_p^2 = 0.377$]. Not surprisingly, neither the main effect of the location validity [$F(1, 19) = 0.062$; $p = 0.806$; $\eta_p^2 = 0.003$], nor the interaction between the location validity and electrode [$F(12, 228) = 0.437$; $p = 0.947$; $\eta_p^2 = 0.022$] were significant. The maximum difference between the valid and invalid locations was found in the P7 electrode (P7v-i = $-0.273 \mu\text{V}$; $p = 0.476$; v-i means valid minus invalid).

N1 component (150–190 ms). Figure 2b,d shows the N1 waveforms and scalp topographies of the A targets. The main effect of the electrodes [$F(12, 228) = 11.380$; $p < 0.001$; $\eta_p^2 = 0.431$] was significant. Similarly, neither the main effect of the location validity [$F(1, 19) = 0.065$; $p = 0.802$; $\eta_p^2 = 0.003$], nor the interaction between the location validity and electrode [$F(12, 228) = 0.378$; $p = 0.970$; $\eta_p^2 = 0.019$] was significant. The maximum difference between the valid and invalid locations was found in the P7 electrode (P7v-i = $-0.288 \mu\text{V}$; $p = 0.331$).

N2 component (220–290 ms). Figure 2e,g shows the waveforms and scalp topographies of the N2 components elicited by the A targets. The main effect of the electrode [$F(12, 228) = 3.399$; $p < 0.001$; $\eta_p^2 = 0.152$] was observed,

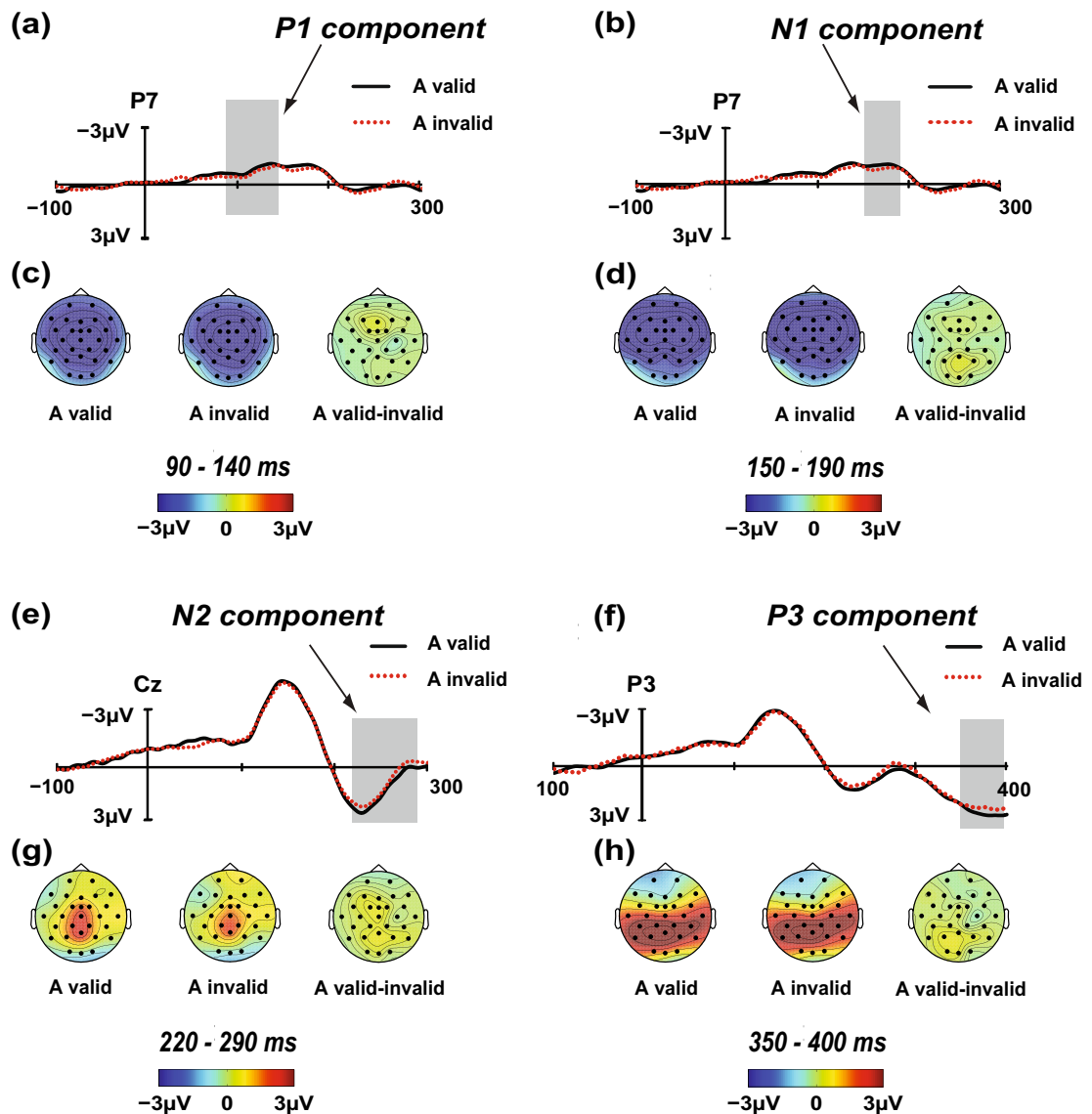


Figure 2. Direct comparisons of the auditory IOR effect for the time windows 90–140 ms, 150–190 ms, 220–290 ms, and 350–400 ms. The contrasts between the valid and invalid locations were analysed using *t* test (sig. two-tailed; 95% confidence interval, CI), and the maximum difference between the valid and invalid locations is shown from P7 electrode, P7 electrode, Cz electrode and P3 electrode, for the P1, N1, N2 and P3 components, respectively. The grand average amplitudes elicited by the A targets presented in either the valid (black solid) or invalid (red dotted) locations are shown in panels (a), (b), (e), and (f). The scalp topographies of the P1, N1, N2 and P3 components in the valid and invalid waveforms are shown in panels (c), (d), (g), and (h), respectively. Valid-invalid scalp topographies represent the IOR effect.

but the main effect of the location validity [$F(1, 19) = 0.602$; $p = 0.447$; $\eta_p^2 = 0.031$] was not observed. No interaction was found between the location validity and electrode [$F(12, 228) = 0.697$; $p = 0.754$; $\eta_p^2 = 0.035$]. The maximum difference between the valid and invalid locations was found in the Cz electrode (Cz v-i = 0.521 μV ; $p = 0.253$).

P3 component (350–400 ms). The P3 waveforms and scalp topographies of the A targets are included in Fig. 2f,h. The analysis yielded a significant main effect for the electrode [$F(12, 228) = 17.469$; $p < 0.001$; $\eta_p^2 = 0.479$]. We did not find a significant effect of the location validity [$F(1, 19) = 0.205$; $p = 0.656$; $\eta_p^2 = 0.011$] and a significant interaction between the location validity and electrode [$F(12, 228) = 0.631$; $p = 0.815$; $\eta_p^2 = 0.032$]. The maximum difference between the valid and invalid locations was found in the P3 electrode (P3v-i = 0.488 μV ; $p = 0.247$).

Event-related potential of the visual IOR effect: (V valid vs. V invalid). **P1 component (90–140 ms).** Figure 3a,c shows the P1 waveforms and scalp topographies of the V targets, respectively. There was a main effect of the electrode [$F(12, 228) = 11.865$; $p < 0.001$; $\eta_p^2 = 0.384$] but no significant main effect of

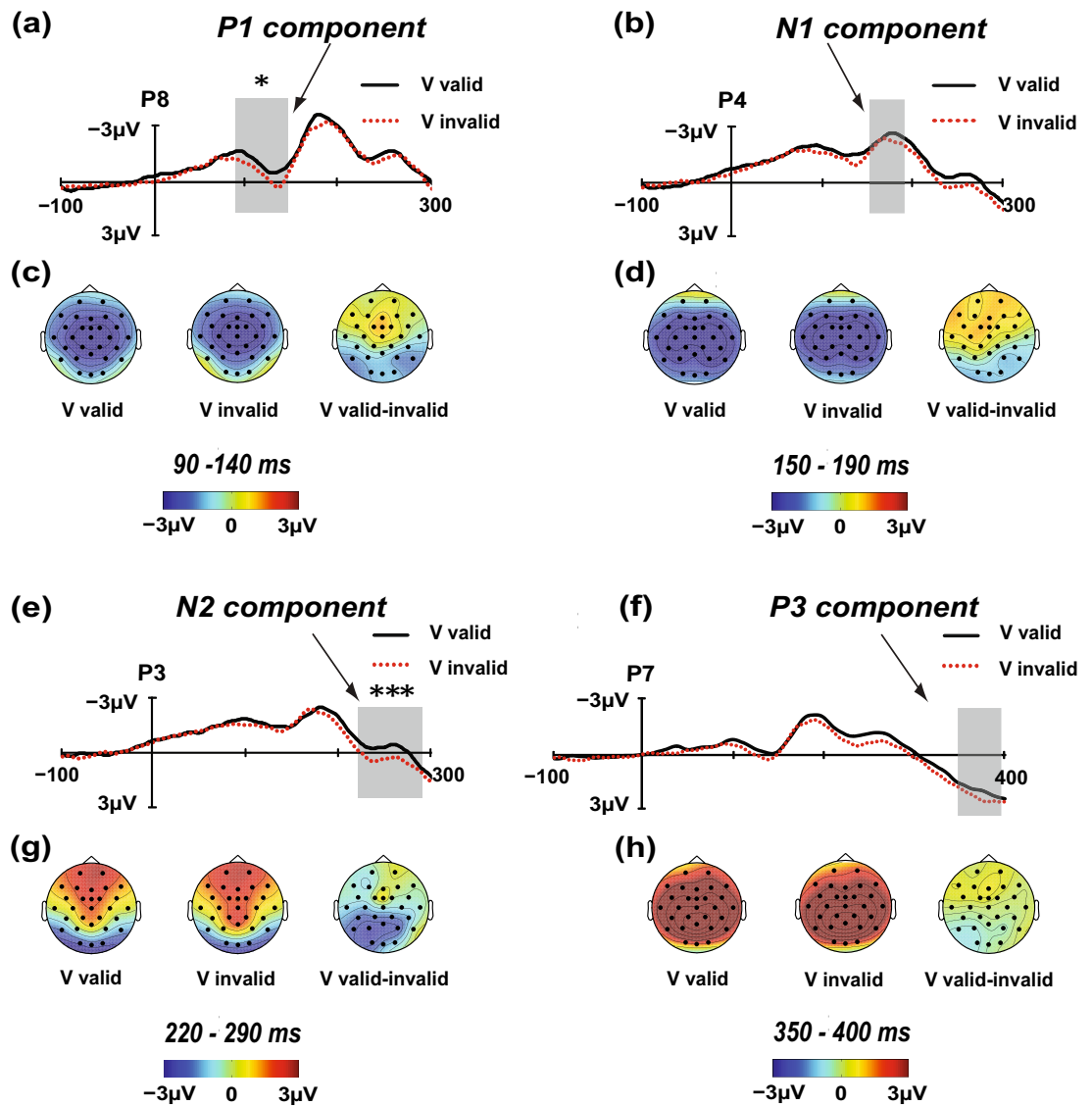


Figure 3. Direct comparisons of the visual IOR effect for the time windows 90–140 ms, 150–190 ms, 220–290 ms, and 350–400 ms. The contrasts between the valid and invalid locations were analysed using *t* test (sig. two-tailed; 95% confidence interval, CI), and the maximum difference between the valid and invalid locations is shown from P4 electrode, and P7 electrode, for the N1 and P3 components, respectively. The difference between the valid and invalid locations for the P1 and N2 components reaches significance in the P8 electrode and P3 electrode, respectively (* $p < 0.05$; *** $p < 0.001$). The grand average amplitudes elicited by the V targets presented in either the valid (black solid) or invalid (red dotted) locations are shown from panels (a), (b), (e), and (f). The scalp topographies of the P1, N1, N2 and P3 components in the valid and invalid waveforms are shown in panels (c), (d), (g), and (h), respectively. Valid-invalid scalp topographies represent the IOR effect.

the location validity [$F(1, 19) = 0.302$; $p = 0.589$; $\eta_p^2 = 0.016$]. Importantly, the interaction between the location validity and electrode was significant [$F(12, 228) = 2.414$; $p < 0.01$; $\eta_p^2 = 0.113$]. Simple effect analysis showed that significant or marginally significant differences between the valid and invalid locations occurred at electrodes P4, P7, and P8 ($P4v-i = -0.507 \mu V$, $p = 0.06$; $P7v-i = -0.506 \mu V$, $p < 0.01$; $P8v-i = -0.538 \mu V$, $p < 0.05$), and a significant maximum difference between the valid and invalid locations was found in the P8 electrode ($-0.538 \mu V$).

N1 component (150–190 ms). Figure 3b,d shows the N1 waveforms and scalp topographies of the V targets. The main effect of the location validity [$F(1, 19) = 0.076$; $p = 0.786$; $\eta_p^2 = 0.004$] was not significant. However, the main effect of the electrode [$F(12, 228) = 3.723$; $p < 0.001$; $\eta_p^2 = 0.164$] was significant. The location validity also interacted with electrode [$F(12, 228) = 2.029$; $p < 0.05$; $\eta_p^2 = 0.096$]. Further simple effect analysis revealed no significant differences between the valid and invalid locations at any of the 13 electrodes. Only a maximum difference between the valid and invalid locations was found in the P4 electrode ($P4v-i = -0.386 \mu V$; $p = 0.194$).

N2 component (220–290 ms). The N2 waveforms and scalp topographies of the V targets are illustrated in Fig. 3e,g. ANOVA revealed a significant main effect of the electrode [$F(12, 228) = 13.421$; $p < 0.001$; $\eta_p^2 = 0.414$]. ANOVA also revealed significant interaction between the location validity and electrode [$F(12, 228) = 1.805$; $p < 0.05$; $\eta_p^2 = 0.087$]. No main effect of the location validity was obtained [$F(1, 19) = 1.965$; $p = 0.177$; $\eta_p^2 = 0.094$]. Simple effect analysis revealed significant amplitude differences between the valid and invalid locations at electrodes P3, P4, P7, and CP5 (P3v-i = $-0.880 \mu\text{V}$, $p < 0.001$; P4v-i = $-0.730 \mu\text{V}$, $p < 0.05$; P7v-i = $-0.692 \mu\text{V}$, $p < 0.01$; CP5v-i = $-0.599 \mu\text{V}$, $p < 0.01$). The significant maximum difference was found in the P3 electrode ($-0.880 \mu\text{V}$) (Supplementary Information).

P3 component (350–400 ms). The P3 waveforms and scalp topographies of the V targets are displayed in Fig. 3f,h. No main effect of the location validity was observed [$F(1, 19) = 0.001$; $p = 0.979$; $\eta_p^2 = 0.000$]. A significant main effect of the electrode was found [$F(12, 228) = 8.615$; $p < 0.001$; $\eta_p^2 = 0.312$], and a significant interaction between the location validity and electrode was also observed [$F(12, 228) = 2.599$; $p < 0.01$; $\eta_p^2 = 0.120$]. However, simple effect analysis revealed no significant differences between the valid and invalid locations at any of the 13 electrodes. The maximum difference was found in the P7 electrode (P7v-i = $-0.567 \mu\text{V}$; $p = 0.072$).

Event-related potential of the audiovisual IOR effect: (AV valid vs. AV invalid). **P1 component (90–140 ms).** Figure 4a,c shows the P1 waveforms and scalp topographies of the AV targets, respectively. We observed a significant main effect of the electrode [$F(12, 228) = 20.112$; $p < 0.001$; $\eta_p^2 = 0.514$]. The main effect of the location validity [$F(1, 19) = 0.321$; $p = 0.577$; $\eta_p^2 = 0.017$], and interaction between the location validity and electrode were not significant [$F(12, 228) = 0.787$; $p = 0.663$; $\eta_p^2 = 0.040$]. The maximum difference between the valid and invalid locations was found in the F7 electrode (F7v-i = $-0.393 \mu\text{V}$; $p = 0.110$).

N1 component (150–190 ms). Figure 4b,d shows the waveforms and scalp topographies of the N1 components elicited by the AV targets. The main effect of the electrode [$F(12, 228) = 10.393$; $p < 0.001$; $\eta_p^2 = 0.354$] was significant. The main effect of the location validity [$F(1, 19) = 1.141$; $p = 0.299$; $\eta_p^2 = 0.057$] and interaction between the location validity and electrode [$F(12, 228) = 0.480$; $p = 0.925$; $\eta_p^2 = 0.025$] were not significant. The maximum difference between the valid and invalid locations was found in the Cz electrode (Cz v-i = $0.506 \mu\text{V}$; $p = 0.255$).

N2 component (220–290 ms). As shown in Fig. 4e,g, the analysis of the N2 component showed a significant main effect of the electrode [$F(12, 228) = 8.373$; $p < 0.001$; $\eta_p^2 = 0.306$], while no significant main effect of the location validity [$F(1, 19) = 0.814$; $p = 0.378$; $\eta_p^2 = 0.041$] or a significant interaction between the location validity and electrode [$F(12, 228) = 1.113$; $p = 0.351$; $\eta_p^2 = 0.055$] was found. The maximum difference between the valid and invalid locations was found in the F4 electrode (F4v-i = $-0.529 \mu\text{V}$; $p = 0.260$).

P3 component (350–400 ms). The P3 waveforms and scalp topographies of the AV targets are illustrated in Fig. 4f,h. We observed a significant main effect of the location validity [$F(1, 19) = 19.338$; $p < 0.001$; $\eta_p^2 = 0.504$], and a significant main effect of the electrode [$F(12, 228) = 14.631$; $p < 0.001$; $\eta_p^2 = 0.435$]. We also found a significant interaction between the location validity and electrode [$F(12, 228) = 4.381$; $p < 0.001$; $\eta_p^2 = 0.187$]. As in a previous study, a subsequent simple analysis revealed a larger P3 amplitude in the valid location for the AV targets, the involved electrodes were CP5, P3, P7, P4, P8, CP6, Cz, FC6, and F8 (CP5v-i = $1.032 \mu\text{V}$, $p < 0.01$; P3v-i = $1.272 \mu\text{V}$, $p < 0.01$; P7v-i = $0.680 \mu\text{V}$, $p < 0.05$; P4v-i = $1.892 \mu\text{V}$, $p < 0.001$; P8v-i = $1.406 \mu\text{V}$, $p < 0.001$; CP6v-i = $1.472 \mu\text{V}$, $p < 0.001$; Cz v-i = $1.222 \mu\text{V}$, $p < 0.01$; FC6v-i = $0.739 \mu\text{V}$, $p < 0.05$; F8v-i = $0.643 \mu\text{V}$, $p < 0.05$). A significant maximum difference was found in the P4 electrode ($1.892 \mu\text{V}$).

MSI effect (AV vs. A + V). According to the above results, the difference in the time course between the visual IOR and audiovisual IOR mainly exists for the P1, N2 and P3 components. The visual IOR was generated from the first two stages, while the audiovisual IOR appeared in only one late processing stage. In this study, we employed a bimodal divided attention task, which may be a key factor. Multisensory integration could occur when targets from both the visual and auditory modalities are presented simultaneously. According to a previous study, multisensory integration played an essential role in partially offsetting the audiovisual IOR effect. This effect was demonstrated by comparing the ERP data of the AV stimuli with the sum of the ERP to the corresponding A and V stimuli, and showing differences between the AV and A + V waveforms⁵⁴. To fully elucidate the reasons for the time course differences between the two IOR effects, the multisensory integration effect was analysed at the time windows of P1 (90–120 ms), N1 (140–180 ms), N2 (220–290 ms) and P3 (350–400 ms) in the corresponding analysis window of the IOR. Figure 5a shows the mean ERP waveforms of 13 electrodes to A, V, A + V and AV targets. To better understand the brain regions where integration occurs, Fig. 5b shows the mean ERP waveforms of 5 cortical regions to AV and A + V. The corresponding scalp topographies of A, V, A + V, AV and AV – (A + V) are shown in Fig. 5c. An overall ANOVA contained the following within-subject factors: target type (2: AV or A + V) and cortical region (5 levels: left frontal F3/F7/FC5, right frontal F4/F8/FC6, central Cz, left posterior P3/P7/CP5 and right posterior P4/P8/CP6).

P1 component (90–120 ms). The test for the P1 component indicated that the main effects of the target type [$F(1, 19) = 34.819$; $p < 0.001$; $\eta_p^2 = 0.647$] and cortical region [$F(4, 76) = 9.512$; $p < 0.001$; $\eta_p^2 = 0.334$] were significant. The P1 amplitude elicited by the AV targets ($-2.535 \mu\text{V}$) was larger than the sum ($-4.495 \mu\text{V}$) of the P1 amplitudes elicited by the A targets and V targets. A statistically significant interaction [$F(4, 76) = 7.767$;

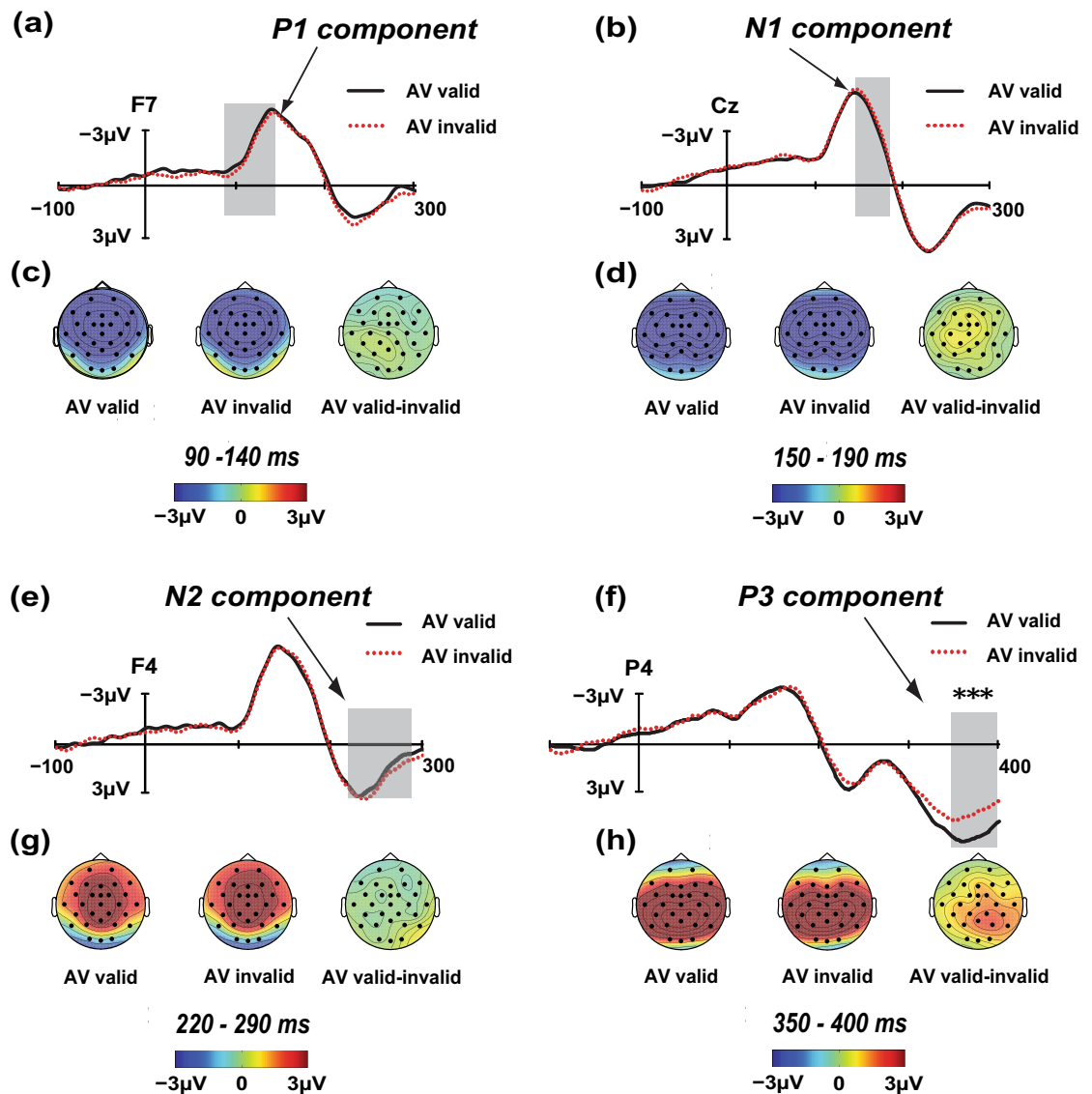


Figure 4. Direct comparisons of the audiovisual IOR effect for the time windows 90–140 ms, 150–190 ms, 220–290 ms, and 350–400 ms. The contrasts between the valid and invalid locations were analysed using *t* test (sig. two-tailed; 95% confidence interval, CI), and the maximum difference between the valid and invalid locations is shown from F7 electrode, Cz electrode, and F4 electrode, for the P1, N1 and N2 components, respectively. The difference between the valid and invalid locations for the P3 component reaches significance in the P4 electrode (***p* < 0.001). The grand average amplitudes elicited by the AV targets presented in either the valid (black solid) or invalid (red dotted) locations are shown in panels (a), (b), (e), and (f). The scalp topographies of the P1, N1, N2 and P3 components in the valid and invalid waveforms are shown in panels (c), (d), (g), and (h), respectively. Valid-invalid scalp topographies represent the IOR effect.

$p < 0.001$; $\eta_p^2 = 0.29$] between target type and cortical region revealed that the P1 amplitude was reduced for the A + V targets relative to AV targets over all five regions ($p < 0.01$).

N1 component (140–180 ms). For the N1 amplitude, the analysis yielded a main effect of the target type [$F(1, 19) = 29.157$; $p < 0.001$; $\eta_p^2 = 0.605$] and cortical region [$F(4, 76) = 21.221$; $p < 0.001$; $\eta_p^2 = 0.528$], and the amplitude of the AV targets ($-5.767 \mu\text{V}$) was smaller than the sum of the amplitudes of A + V ($-8.154 \mu\text{V}$). A significant interaction was also found between the target type and cortical region [$F(4, 76) = 4.634$; $p < 0.01$; $\eta_p^2 = 0.196$], and simple effect analysis revealed a multisensory integration effect in all five regions ($p < 0.01$).

N2 component (220–290 ms). Similarly, the results revealed that both the target type [$F(1, 19) = 7.246$; $p < 0.05$; $\eta_p^2 = 0.335$] and cortical region [$F(4, 76) = 9.581$; $p < 0.001$; $\eta_p^2 = 0.335$] were significant, with smaller negative amplitudes of the AV targets ($2.576 \mu\text{V}$) than the sum of the amplitudes of A + V ($1.444 \mu\text{V}$). However, the interaction between the target type and cortical region was not significant [$F(4, 76) = 1.191$; $p = 0.322$; $\eta_p^2 = 0.059$].

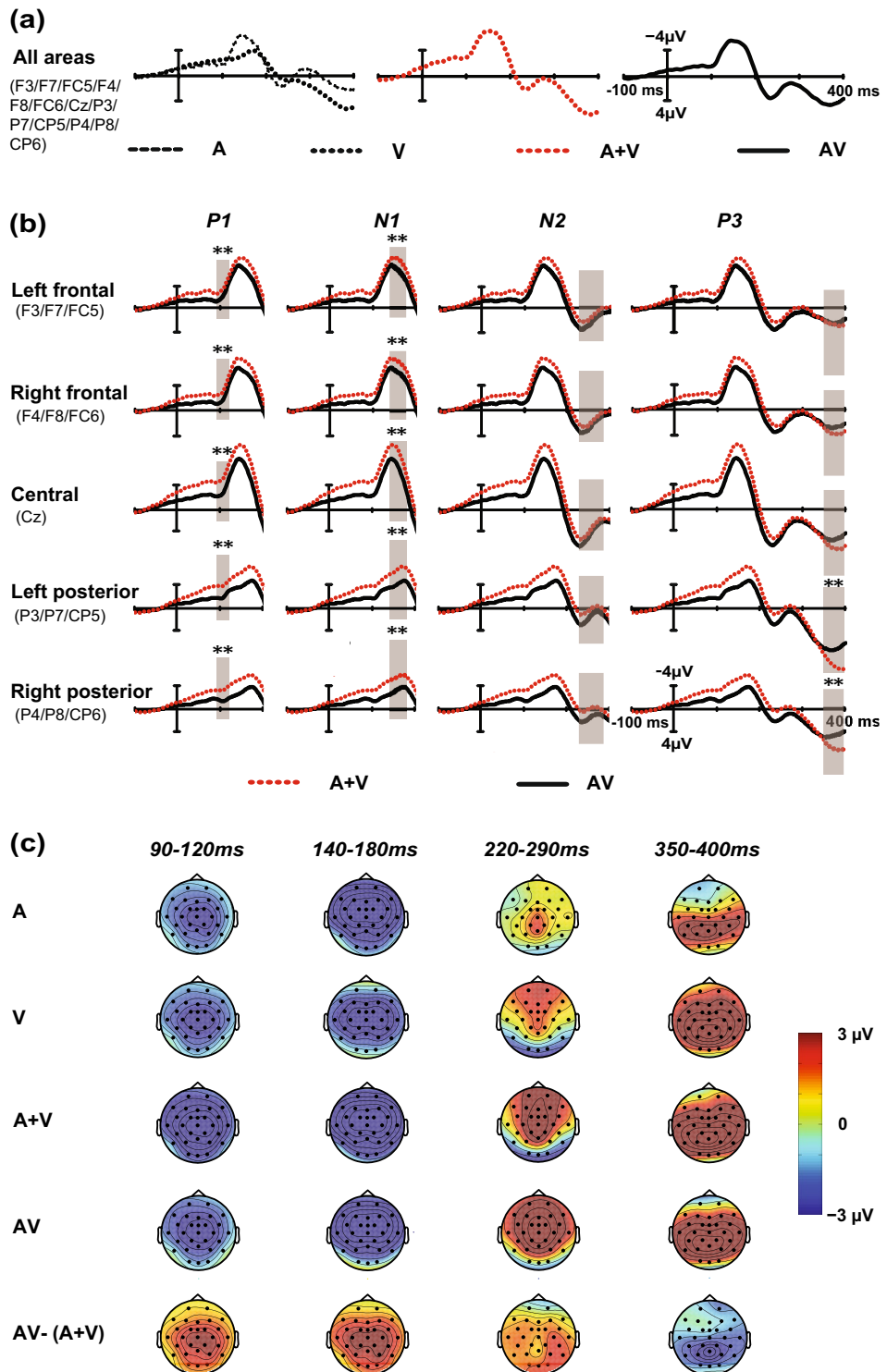


Figure 5. Multisensory integration effect for the AV targets. Panel (a) displays the mean ERP waveforms of 13 electrodes to A (black broken), V (black dotted), A + V (red dotted) and AV (black solid). The five regions of ERP waves in panel (b) depicts the distribution of the multisensory integration effect in four windows, shown separately for the 90–120 ms, 140–180 ms, 220–290 ms, and 350–400 ms windows. Significant difference between the bimodal AV (black solid) and summed unimodal A + V (red dotted) ERP waveforms (sig. two-tailed; 95% confidence interval, CI), reflecting a multisensory integration effect for the AV targets and marked by the symbol ‘**’ ($p < 0.01$). Scalp topographies of unimodal A, unimodal V, summed unimodal A + V, and bimodal AV and the AV-(A + V) processing are depicted in panel (c).

P3 component (350–400 ms). The P3 amplitude was expressed in either a main effect of the target type [$F(1, 19) = 5.267$; $p < 0.05$; $\eta_p^2 = 0.217$] or cortical region [$F(4, 76) = 10.852$; $p < 0.001$; $\eta_p^2 = 0.364$], with a stronger P3 amplitude in the A + V (6.243 μV) targets than in the AV targets (4.549 μV). Additionally, the two-way interaction between the target type and electrode was significant [$F(4, 76) = 4.244$; $p < 0.01$; $\eta_p^2 = 0.183$], resulting from a greater positive amplitude on the A + V waveforms than on the AV waveform at the left posterior region ($p < 0.01$) and right posterior region ($p < 0.01$).

In summary, multisensory integration emerged in three processing stages: the earliest integration appeared in all five cortical areas and over 90–120 ms. The second integration appeared at 140–180 ms and covered these five areas, and the last integration mainly occurred in the left posterior region and the right posterior region for the time window of 350–400 ms.

Discussion

This study aimed to examine the electrophysiological underpinning of behaviour IOR differences between visual targets and audiovisual targets. Our behavioural results confirmed previous studies⁷ in which the IOR effect of the visual targets was larger than that of the audiovisual targets when participants focused on both the visual and auditory modalities, while no auditory IOR effect was observed. By comparing the ERP results of the auditory targets, visual targets and audiovisual targets, the visual IOR appeared at two processing stages (P1, N2), while only the late stage (P3) had a significant audiovisual IOR. The ERP results suggested that the time courses of visual IOR and audiovisual IOR were different.

As expected, we found no auditory IOR at the four processing stages. Although the existence of the auditory IOR has now been demonstrated, little is known about the factors that might limit this effect. However, much more is known about the factors that limit the visual IOR⁵⁵. The visual IOR has usually been attributed to the operation of an exogenous (involuntary) attentional mechanism. Rafal et al.^{56,57} showed that the visual IOR is eliminated when the likelihood that the cue provides accurate spatial information is better than chance. This happens because the predictability of spatial relations may activate the endogenous (voluntary) orienting mechanism, while the higher-level endogenous orienting mechanism system appears to be able to modify the effects of the lower-level exogenous attentional mechanism⁵⁸. By contrast, the results of Mondor's⁵⁵ study showed that predictability substantially influences the time course of the inhibitory component of the auditory IOR. An auditory IOR effect became apparent at a longer SOA when either the temporal or the spatial relationship was predictable, rather than when either was unpredictable. The auditory and visual inhibitory seems to have different bases for selective attention. In this study, we used a cue–target paradigm to examine the effect of uninformative visual cues on IOR effects. Moreover, we only set a single SOA that was randomly presented in 600–900 ms. According to Mondor's views, either the temporal or spatial relationship is unpredictable, which might be the key factors in failure to identify the auditory IOR. Coincidentally, Prime and Word's⁵⁹ work demonstrates that such auditory IOR effects are not found in the target–target paradigm or cue–target paradigm requiring Go/NoGo discrimination of the targets. The latter is the paradigm used in current experiments. Not surprisingly, the auditory IOR was missed. We were interested in revealing the neural processing truth of the visual IOR and audiovisual IOR of different magnitudes.

In our study, we found no early P1 neural modulation for the audiovisual IOR, likely because of multisensory integration. There is converging evidence to suggest that it leads to multisensory integration when auditory and visual stimuli are simultaneously presented⁶⁰. More importantly, multisensory integration could affect IOR components elicited by exogenous cues⁶¹. Combined with our studies, the audiovisual stimuli were a combination of auditory and visual stimuli presented simultaneously, indicating that multisensory integration could occur. To illustrate the effect of multisensory integration on IOR, we analysed the time course of the multisensory integration of audiovisual targets. As expected, significant differences between the ERP amplitude of AV and combined ERP amplitude of A + V were found in the three time windows of P1, N1 and P3, i.e., multisensory integration occurred at multiple stages of processing. Previous studies have confirmed that multisensory integration more frequently appeared at the P50 stage but also existed at a relatively late stage of processing^{16,48}. We found early multisensory integration in all five cortical regions, and this strong integration indeed caused the absence of an early audiovisual IOR in all brain regions. Only the left posterior region and right posterior region induced the late multisensory integration effect, which cannot offset the IOR effect in the right frontal region (F8/FC6) or central region (Cz), resulting in the partial retention of the late audiovisual IOR. Instead, the unimodal targets do not have the conditions for multisensory integration, and the visual IOR effect can appear on the relatively early P1 and N2 components. The current results confirm the hypothesis that the power of enhanced perceptual salience by audiovisual integration offsets the reduced perceptual processing due to IOR, but not in every stage. When the IOR (decreasing salience) encounters the integration of attended visual and attended auditory stimuli (increasing perceptual salience), a reduced IOR effect elicited by an audiovisual target is found. Taken together, multisensory integration of the audiovisual targets provides some explanations for the magnitude and time course differences between the visual IOR and audiovisual IOR.

However, multisensory integration does not fully explain the time course differences between the visual IOR and the audiovisual IOR. First, audiovisual targets have neither MSI nor the IOR associated with the N2 component, which is a more reliable neural marker of IOR³⁰. For the late response-related N2 component, the IOR was affected by the proportion of Go/NoGo stimuli. The N2 effects were reduced with the decreasing probability of Go stimuli¹⁸. Thus, the IOR effect related to the N2 component could occur in visual and audiovisual targets because of the high probability of 80% Go stimuli in this experiment. As mentioned, an IOR effect similar to that of the visual targets should occur in the N2 component in the absence of an integration effect but did not. Thus, multisensory integration is not the only factor. Additionally, IOR was detected only on the P3 component of the audiovisual targets rather than the visual targets, reflecting that IOR has been affected by the target modality,

perceptual factors known to influence the cueing effects^{45,53}. In other words, the neural processing of the audiovisual IOR is also different from the visual IOR in addition to the influence of multisensory integration factors. Previous studies on the combined regulation of N2 and P3 can account for this difference. Edmund et al.⁶² found that some elderly individuals had a loss of inhibitory control (absence of N2), and the lack of inhibitory control might induce the unnecessary deep evaluation of irrelevant stimuli (presence of P3b). Combined with the current results, the visual targets showed a significant IOR effect on the N2 component and required no more P3 processing. However, we failed to observe the IOR effect associated with the N2 component for the audiovisual targets, leading to P3 processing. As described by Edmund et al., the emergence of the P3 component reflects additional processing. If the time course differences of the IOR effect are simply a function of integration, then the neural processing in the time window without integration should be the same. However, in fact, the N2 component is not similar in these two targets, leading to different P3 processing. Therefore, the neural processing of the two types of IOR were fundamentally different.

Specifically, the neural processing of visual IOR reflects two components of perceptual suppression and reaction inhibition. The former may be related to the suggestion that IOR is associated with changes in perceived quality, which is reflected in the reduced P1 amplitude in extrastriate visual areas^{45,63}. The latter may be related to the suggestion that the IOR may be a response criterion marked by N2⁶⁴, representing reluctance to respond to the previously valid location. P1 and N2 were both related to the visual IOR in our study, supporting that the visual IOR may comprise multiple components (attentional and reaction) that act flexibly on different stages of cognitive processing^{38,65,66}. Whether the IOR triggers attentional or reaction components has been reported to depend on the response preference determined by the task requirements³⁶. For example, a speedy response requirement (reducing the quality of stimulus) may reveal more of the criterion-shift component of IOR, whereas guidance that encourages correctly responding (and emphasises a deeper processing of the stimulus) may reveal more of the attentional component of IOR. In some tasks, such as ours, in which speed and accuracy are equally stressed, both components of IOR may operate within different stages of the time distribution⁶⁷. Thus, the effects of IOR on the response criterion and attention are independent, but they operate cooperatively to achieve the same goal—that is fast and accurate performance.

By contrast, the IOR effect of audiovisual targets was only detected at the P3 stage. P3 component is thought to index aspects of perceived stimulus relevance or context updating^{68–70}. As another neural marker of IOR, P3 was increased when a Go stimulus was cued⁷¹ or was otherwise expected⁷² and usually showed a larger amplitude in the valid locations than in the invalid locations, contrary to the trend of the earlier neural component. Joseph et al.²⁴ proposed that P3 has a greater influence on the stimuli of valid locations, indicating that reflexive attention mechanisms cause valid location stimuli to be treated as more salient or potentially significant at higher stages of stimuli evaluation. Additionally, P3 is typically sensitive to infrequent stimuli⁷³, suggesting a higher level of attention affecting information processing, perhaps by tagging novel events as having greater potential relevance than other stimuli. Other studies also confirmed that unexpected stimuli increase the P3 component of the wave¹⁹. These results provided some evidence that the neural processing of the audiovisual IOR in the current study occurs late and contrasts the trend of early neural marks. Combined with the target type after the visual cue in this study, the space specific attention bias of the modality can be explained accordingly. For example, a visual cue on the left might effectively "prime" the visual modality for that side, over other sensory modalities^{74,75}. When the cues are visual stimuli, we might expect the subsequent target to be visual as well. When the visual targets are accompanied by the auditory targets, the audiovisual targets may exceed the participants' expectations, and their presence at the valid locations may result in an increase in the P3 amplitude. This point was supported by the evidence of the P3 component from auditory targets. Figures 2f,h and 4f,h show P3 enhancement in the valid locations but not in the invalid locations for both auditory targets and audiovisual targets, rather than for visual targets, and this P3 enhancement was not significant for auditory targets. Thus, the P3 component related to the audiovisual IOR effect was not accidental.

Additionally, we found no visual IOR or audiovisual IOR effects associated with the early N1 component. It is debatable whether the N1 component marks IOR. In the discrimination or Go/NoGo tasks, reductions in the N1 component for valid location compared with those for invalid location trials were associated with IOR. In the detection or localisation task, the absence of N1 modulations was associated with IOR. Differences in the experimental paradigm could be crucial for understanding these discrepant results. Martín-Arévalo et al.⁷⁶ pointed out that previous studies do not provide a consistent association between modulations of the N1 component and the observed behaviour IOR effect; if any association exists, it seems highly dependent on some variables such as the task at hand. In this study, the discrimination task combined with the Go targets/NoGo non-targets setting was adopted. We did not observe the IOR effect of N1 modulation, and the related causes need to be further studied.

To our best knowledge, the present study provides the first electrophysiological evidence for the behaviour IOR difference between visual targets and audiovisual targets. The visual IOR was generated from the P1 and N2 components, while the audiovisual IOR was only derived from the P3 component. Multisensory integration from audiovisual targets causes great effects on audiovisual IOR by offsetting perceptual salience reduced at multiple processing stages. Additionally, the information processing patterns of the visual IOR and the audiovisual IOR are fundamentally different besides the influence of integration. In summary, the difference in early and late processing between the visual IOR and audiovisual IOR implies that the two types of targets may have different inhibitory orientation mechanisms.

Received: 12 October 2020; Accepted: 22 March 2021

Published online: 13 April 2021

References

- Posner, M. & Cohen, Y. Components of visual orienting. In *Attention and Performance X: Control of Language Processes* Vol. 32 (eds Bouma, H. & Bouwhuis, D. G.) 531–556 (Psychology Press, 1984).
- Klein, R. Inhibition return. *Trends Cogn. Sci.* **4**, 138–147. [https://doi.org/10.1016/S1364-6613\(00\)01452-2](https://doi.org/10.1016/S1364-6613(00)01452-2) (2000).
- Pierce, A. M., McDonald, J. J. & Green, J. J. Electrophysiological evidence of an attentional bias in crossmodal inhibition of return. *Neuropsychologia* **114**, 11–18. <https://doi.org/10.1016/j.neuropsychologia.2018.04.007> (2018).
- Lupiáñez, J., Milán, E. G., Tornay, F. J., Madrid, E. & Tudela, P. Does IOR occur in discrimination tasks? Yes, it does, but later. *Percept. Psychophys.* **59**, 1241–1254. <https://doi.org/10.3758/BF03214211> (1997).
- Spence, C. & Driver, J. Auditory and audiovisual inhibition of return. *Percept. Psychophys.* **60**, 125–139. <https://doi.org/10.3758/BF03211923> (1998).
- McDonald, J. J., Ward, L. M. & Kiehl, K. A. An event-related brain potential study of inhibition of return. *Percept. Psychophys.* **61**, 1411–1423. <https://doi.org/10.3758/BF03206190> (1999).
- Tang, X. *et al.* Bimodal-divided attention attenuates visually induced inhibition of return with audiovisual targets. *Exp. Brain Res.* **237**, 1093–1107. <https://doi.org/10.1007/s00221-019-05488-0> (2019).
- Van der Burg, E., Talsma, D., Olivers, C. N. L., Hickey, C. & Theeuwes, J. Early multisensory interactions affect the competition among multiple visual objects. *Neuroimage* **55**, 1208–1218. <https://doi.org/10.1016/j.neuroimage.2010.12.068> (2011).
- Krause, H., Schneider, T. R., Engel, A. K. & Daniel, S. Capture of visual attention interferes with multisensory speech processing. *Front. Integr. Neurosci.* **6**, 67. <https://doi.org/10.3389/fnint.2012.00067> (2012).
- Spence, C. & Santangelo, V. Capturing spatial attention with multisensory cues: A review. *Hearing Res.* **258**, 134–142. <https://doi.org/10.1016/j.heares.2009.04.015> (2009).
- Belopolsky, A. V. & Theeuwes, J. Inhibition of saccadic eye movements to locations in spatial working memory. *Atten. Percept. Psychophys.* **71**, 620–631. <https://doi.org/10.3758/APP.71.3.620> (2009).
- van Martijn Gerbrand, K. Hemispheric asymmetry in the remapping and maintenance of visual saliency maps: A TMS study. *J. Cogn. Neurosci.* **22**, 1730–1738. <https://doi.org/10.1162/jocn.2009.21356> (2010).
- Tanaka, E., Kida, T., Kakigi, R. & Hoshiyama, M. Neuroscientific evidence for multisensory convergence and interaction. *J. Phys. Fitness Sport* **6**, 301–310. <https://doi.org/10.7600/jpfsm.6.301> (2017).
- Macaluso, E. & Driver, J. Spatial attention and crossmodal interactions between vision and touch. *Neuropsychologia* **39**, 1304–1316. [https://doi.org/10.1016/S0028-3932\(01\)00119-1](https://doi.org/10.1016/S0028-3932(01)00119-1) (2001).
- Calvert, G. A. & Thesen, T. Multisensory integration: Methodological approaches and emerging principles in the human brain. *J. Physiol. Paris* **98**, 191–205. <https://doi.org/10.1016/j.jphysparis.2004.03.018> (2004).
- Koelewijn, T., Bronkhorst, A. & Theeuwes, J. Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychol.* **134**, 372–384. <https://doi.org/10.1016/j.actpsy.2010.03.010> (2010).
- Hickey, C., McDonald, J. J. & Theeuwes, J. Electrophysiological evidence of the capture of visual attention. *J. Cogn. Neurosci.* **18**, 604–613. <https://doi.org/10.1162/jocn.2006.18.4.604> (2006).
- Prime, D. J. & Jolicoeur, P. Response-selection conflict contributes to inhibition of return. *J. Cogn. Neurosci.* **21**, 991–999. <https://doi.org/10.1162/jocn.2009.21105> (2009).
- Chica, A. B. & Lupiáñez, J. Effects of endogenous and exogenous attention on visual processing: An inhibition of return study. *Brain Res.* **1278**, 75–85. <https://doi.org/10.1016/j.brainres.2009.04.011> (2009).
- McDonald, J. J., Hickey, C., Green, J. J. & Whitman, J. C. Inhibition of return in the covert deployment of attention: Evidence from human electrophysiology. *J. Cogn. Neurosci.* **21**, 725–733. <https://doi.org/10.1162/jocn.2009.21042> (2009).
- Van der Lubbe, R. H. J., Vogel, E. & Postma, A. Different effects of exogenous cues in a visual detection and discrimination task: Delayed attention withdrawal and/or speeded motor inhibition?. *J. Cogn. Neurosci.* **17**, 1829–1840. <https://doi.org/10.1162/089892905775008634> (2005).
- Prime, D. J. & Ward, L. M. Cortical expressions of inhibition of return. *Brain Res.* **1072**, 161–174. <https://doi.org/10.1016/j.brainres.2005.11.081> (2006).
- Prime, D. J. & Jolicoeur, P. On the relationship between occipital cortex activity and inhibition of return. *Psychophysiology* **46**, 1278–1287. <https://doi.org/10.1111/j.1469-8986.2009.00858.x> (2009).
- Hopfinger, J. B. & Mangun, G. R. Tracking the influence of reflexive attention on sensory and cognitive processing. *Cogn. Affect. Behav. Neurosci.* **1**, 56–65. <https://doi.org/10.3758/CABN.1.1.56> (2001).
- Prime, D. J. & Ward, L. M. Inhibition of return from stimulus to response. *Psychol. Sci.* **15**, 272–276. <https://doi.org/10.1111/j.0956-7976.2004.00665.x> (2004).
- Satel, J., Hilchey, M., Wang, Z., Reiss, C. & Klein, R. In search of a reliable electrophysiological marker of oculomotor inhibition of return. *Psychophysiology* **51**, 1037–1045. <https://doi.org/10.1111/psyp.12245> (2014).
- Falkenstein, M., Hoormann, J. & Hohnsbein, J. ERP components in Go/Nogo tasks and their relation to inhibition. *Acta Psychol.* **101**, 267–291. [https://doi.org/10.1016/S0001-6918\(99\)00008-6](https://doi.org/10.1016/S0001-6918(99)00008-6) (1999).
- Sanger, K. L. & Dorjee, D. Mindfulness training with adolescents enhances metacognition and the inhibition of irrelevant stimuli: Evidence from event-related brain potentials. *Trends Neurosci. Educ.* **5**, 1–11. <https://doi.org/10.1016/j.tine.2016.01.001> (2016).
- Folstein, J. R. & Van Petten, C. Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology* **45**, 152–170. <https://doi.org/10.1111/j.1469-8986.2007.00602.x> (2008).
- Satel, J., Wang, Z., Hilchey, M. D. & Klein, R. M. Examining the dissociation of retinotopic and spatiotopic inhibition of return with event-related potentials. *Neurosci. Lett.* **524**, 40–44. <https://doi.org/10.1016/j.neulet.2012.07.003> (2012).
- Gutiérrez-Domínguez, F. J. *et al.* Vertical asymmetries and inhibition of return: Effects of spatial and non-spatial cueing on behavior and visual ERPs. *Int. J. Psychophysiol.* **91**, 121–131. <https://doi.org/10.1016/j.ijpsycho.2013.12.004> (2014).
- Martín-Arévalo, E., Chica, A. B. & Lupiáñez, J. Electrophysiological modulations of exogenous attention by intervening events. *Brain Cogn.* **85**, 239–250. <https://doi.org/10.1016/j.bandc.2013.12.012> (2014).
- Posner, M. I., Rafal, R. D., Choate, L. S. & Vaughan, J. Inhibition of return: Neural basis and function. *Cogn. Neuropsychol.* **2**, 211–228. <https://doi.org/10.1080/02643298508252866> (1985).
- Spalek, T. M. & Di Lollo, V. The time required for perceptual (nonmotoric) processing in IOR. *Psychon. Bull. Rev.* **14**, 327–331. <https://doi.org/10.3758/BF03194072> (2007).
- Ivanoff, J. & Klein, R. M. The presence of a nonresponding effector increases inhibition of return. *Psychon. Bull. Rev.* **8**, 307–314. <https://doi.org/10.3758/BF03196166> (2001).
- Taylor, T. L. & Klein, R. M. Visual and motor effects in inhibition of return. *J. Exp. Psychol. Hum.* **26**, 1639–1656. <https://doi.org/10.1037//0096-1523.26.5.1639> (2000).
- Zhang, M., Tang, X. & Wu, J. Blocking the link between stimulus and response at previously attended locations: Evidence for inhibitory tagging mechanism. *Neurosci. Biomed. Eng.* **1**, 13–21. <https://doi.org/10.2174/2213385211301010004> (2013).
- Hunt, A. R. & Kingstone, A. Inhibition of return: Dissociating attentional and oculomotor components. *J. Exp. Psychol. Hum.* **29**, 1068–1074. <https://doi.org/10.1037/0096-1523.29.5.1068> (2003).
- Abrams, R. A. & Dobkin, R. S. Inhibition of return: Effects of attentional cuing on eye movement latencies. *J. Exp. Psychol. Hum.* **20**, 467–477. <https://doi.org/10.1037//0096-1523.20.3.467> (1994).

40. Santangelo, V., Lubbe, R. H. J. V. D., Belardinelli, M. O. & Postma, A. Multisensory integration affects ERP components elicited by exogenous cues. *Exp. Brain Res.* **185**, 269–277. <https://doi.org/10.1007/s00221-007-1151-5> (2008).
41. Verleger, R. Event-related potentials and cognition: A critique of the context updating hypothesis and an alternative interpretation of P3. *Behav. Brain Sci.* **11**, 343–356. <https://doi.org/10.1017/S0140525X00058015> (1988).
42. Polich, J. & Herbst, K. L. P300 as a clinical assay: Rationale, evaluation, and findings. *Int. J. Psychophysiol.* **38**, 3–19. [https://doi.org/10.1016/S0167-8760\(00\)00127-6](https://doi.org/10.1016/S0167-8760(00)00127-6) (2000).
43. Jones, K. A. *et al.* A cholinergic receptor gene (CHRM2) affects event-related oscillations. *Behav. Genet.* **36**, 627–639. <https://doi.org/10.1007/s10519-006-9075-6> (2006).
44. Keitel, C. & Müller, M. M. Audio-visual synchrony and feature-selective attention co-amplify early visual processing. *Exp. Brain Res.* **234**, 1221–1231. <https://doi.org/10.1007/s00221-015-4392-8> (2016).
45. Tian, Y. & Yao, D. A study on the neural mechanism of inhibition of return by the event-related potential in the Go/NoGo task. *Biol. Psychol.* **79**, 171–178. <https://doi.org/10.1016/j.biopsycho.2008.04.006> (2008).
46. John, J. F. The case for feedforward multisensory convergence during early cortical processing. *NeuroReport* **16**, 419–423. <https://doi.org/10.1097/00001756-200504040-00001> (2005).
47. Magosso, E., Ursino, M. & Cuppini, C. A neurocomputational model of cortical auditory–visual illusions. *Seeing Perceiving* **25**, 115–115. <https://doi.org/10.1163/187847612X647487> (2012).
48. Andres, A. J. D., Oram Cardy, J. E. & Joanisse, M. F. Congruency of auditory sounds and visual letters modulates mismatch negativity and P300 event-related potentials. *Int. J. Psychophysiol.* **79**, 137–146. <https://doi.org/10.1016/j.ijpsycho.2010.09.012> (2011).
49. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191. <https://doi.org/10.3758/BF03193146> (2007).
50. Van der Stoep, N., Van der Stigchel, S., Nijboer, T. C. W. & Spence, C. Visually induced inhibition of return affects the integration of auditory and visual information. *Perception* **46**, 6–17. <https://doi.org/10.1177/0301006616661934> (2017).
51. Satel, J., Hilchey, M. D., Wang, Z., Story, R. & Klein, R. M. The effects of ignored versus foveated cues upon inhibition of return: An event-related potential study. *Atten. Percept. Psychophys.* **75**, 29–40. <https://doi.org/10.3758/s13414-012-0381-1> (2013).
52. Green, J. J. & McDonald, J. J. Electrical neuroimaging reveals timing of attentional control activity in human brain. *PLoS Biol.* **6**, e81. <https://doi.org/10.1371/journal.pbio.0060081> (2008).
53. Wascher, E. & Tipper, S. P. Revealing effects of noninformative spatial cues: An EEG study of inhibition of return. *Psychophysiology* **41**, 716–728. <https://doi.org/10.1111/j.1469-8986.2004.00198.x> (2004).
54. Talsma, D. & Woldorff, M. G. Selective attention and multisensory integration: Multiple phases of effects on the evoked brain activity. *J. Cogn. Neurosci.* **17**, 1098–1114. <https://doi.org/10.1162/0898929054475172> (2005).
55. Mondor, T. A. Predictability of the cue-target relation and the time-course of auditory inhibition of return. *Percept. Psychophys.* **61**, 1501–1509. <https://doi.org/10.3758/BF03213113> (1999).
56. Rafal, R. D., Calabresi, P. A., Brennan, C. W. & Sciolto, T. K. Saccade preparation inhibits reorienting to recently attended locations. *J. Exp. Psychol. Hum.* **15**, 673–685. <https://doi.org/10.1037//0096-1523.15.4.673> (1989).
57. Rafal, R. & Henik, A. The neurology of inhibition: Integrating controlled and automatic processes. In *Inhibitory Processes in Attention, Memory, and Language* (eds Dagenbach, D. & Carr, T. H.) 1–51 (Academic Press, 1994).
58. Müller, H. J. & Rabbitt, P. M. Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *J. Exp. Psychol. Hum.* **15**, 315–330. <https://doi.org/10.1037//0096-1523.15.2.315> (1989).
59. Prime, D. J. & Ward, L. M. Auditory frequency-based inhibition differs from spatial IOR. *Percept. Psychophys.* **64**, 771–784. <https://doi.org/10.3758/BF03194744> (2002).
60. Foxe, J. *et al.* Auditory modulation of oscillatory activity in extra-striate visual cortex and its contribution to audio-visual multisensory integration: A human intracranial EEG study. *Seeing Perceiving* **25**, 198–198. <https://doi.org/10.1163/187847612X648279> (2012).
61. Santangelo, V. & Spence, C. Is the exogenous orienting of spatial attention truly automatic? Evidence from unimodal and multisensory studies. *Conscious Cogn.* **17**, 989–1015. <https://doi.org/10.1016/j.concog.2008.02.006> (2008).
62. Wascher, E., Falkenstein, M. & Wild-Wall, N. Age related strategic differences in processing irrelevant information. *Neurosci. Lett.* **487**, 66–69. <https://doi.org/10.1016/j.neulet.2010.09.075> (2011).
63. Hopfinger, J. B. & Mangun, G. R. Reflexive attention modulates processing of visual stimuli in human extrastriate cortex. *Psychol. Sci.* **9**, 441–447. <https://doi.org/10.1111/1467-9280.00083> (1998).
64. Band, G. P. H., Ridderinkhof, K. R. & van der Molen, M. W. Speed-accuracy modulation in case of conflict: The roles of activation and inhibition. *Psychol. Res.* **67**, 266–279. <https://doi.org/10.1007/s00426-002-0127-0> (2003).
65. Chica, A. B., Taylor, T. L., Lupiáñez, J. & Klein, R. M. Two mechanisms underlying inhibition of return. *Exp. Brain Res.* **201**, 25–35. <https://doi.org/10.1007/s00221-009-2004-1> (2010).
66. Smith, D. T., Jackson, S. R. & Rorden, C. An intact eye-movement system is not required to generate inhibition of return. *J. Neuropsychol.* **3**, 267–271. <https://doi.org/10.1348/174866408X324377> (2009).
67. Ivanoff, J. & Klein, R. Inhibition of return: Sensitivity and criterion as a function of response time. *J. Exp. Psychol. Hum.* **32**, 908–919. <https://doi.org/10.1037/0096-1523.32.4.908> (2006).
68. Donchin, E. Surprise!... Surprise?. *Psychophysiology* **18**, 493–513. <https://doi.org/10.1111/j.1469-8986.1981.tb01815.x> (1981).
69. Donchin, E. & Coles, M. G. H. Is the P300 component a manifestation of context updating?. *Behav. Brain Sci.* **11**, 357–374. <https://doi.org/10.1017/S0140525X00058027> (2010).
70. Sander, N. Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychol. Bull.* **131**, 510–532. <https://doi.org/10.1037/0033-2909.131.4.510> (2005).
71. Randall, W. M. & Smith, J. L. Conflict and inhibition in the cued-Go/NoGo task. *Clin. Neurophysiol.* **122**, 2400–2407. <https://doi.org/10.1016/j.clinph.2011.05.012> (2011).
72. Smith, J. L., Smith, E. A., Provost, A. L. & Heathcote, A. Sequence effects support the conflict theory of N2 and P3 in the Go/NoGo task. *Int. J. Psychophysiol.* **75**, 217–226. <https://doi.org/10.1016/j.ijpsycho.2009.11.002> (2010).
73. Duncan-Johnson, C. C. & Donchin, E. On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology* **14**, 456–467. <https://doi.org/10.1111/j.1469-8986.1977.tb01312.x> (1977).
74. Spence, C. & Driver, J. On measuring selective attention to an expected sensory modality. *Percept. Psychophys.* **59**, 389–403. <https://doi.org/10.3758/BF03211906> (1997).
75. Spence, C., Lloyd, D., McGlone, F., Nicholls, M. E. R. & Driver, J. Inhibition of return is supramodal: A demonstration between all possible pairings of vision, touch, and audition. *Exp. Brain Res.* **134**, 42–48. <https://doi.org/10.1007/s002210000442> (2000).
76. Martín-Arévalo, E., Chica, A. B. & Lupiáñez, J. No single electrophysiological marker for facilitation and inhibition of return: A review. *Exp. Brain Res.* **300**, 1–10. <https://doi.org/10.1016/j.bbr.2015.11.030> (2016).

Acknowledgements

This study was supported in part by a National Natural Science Foundation of China (31600882, XT; 31871092, MZ; 31700939, AW; 61773076, QL; 61703069, CZ).

Author contributions

X.T. conceived the study and designed the experiments. Testing and data collection were performed by X.P.; X.W. and C.Z. performed the data analysis and interpretation under the supervision of X.T.; X.T. and X.W. wrote the main manuscript text; Q.L., A.W., M.Z. reviewed the manuscript and all authors approved the final version of the manuscript for submission.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-86999-1>.

Correspondence and requests for materials should be addressed to X.T., X.P., A.W. or M.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021