Multisensory integration attenuates visually induced oculomotor inhibition of return

Xiaoyu Tang
School of Psychology, Liaoning Collaborative Innovation Center of Children and Adolescents Healthy Personality Assessment and Cultivation, Liaoning Normal University, Dalian, China ✉

Mengying Yuan
School of Psychology, Liaoning Collaborative Innovation Center of Children and Adolescents Healthy Personality Assessment and Cultivation, Liaoning Normal University, Dalian, China ✉

Zhongyu Shi
School of Psychology, Liaoning Collaborative Innovation Center of Children and Adolescents Healthy Personality Assessment and Cultivation, Liaoning Normal University, Dalian, China ✉

Min Gao
School of Psychology, Liaoning Collaborative Innovation Center of Children and Adolescents Healthy Personality Assessment and Cultivation, Liaoning Normal University, Dalian, China ✉

Rongxia Ren
Interdisciplinary Science and Engineering in Health Systems, Okayama University, Okayama, Japan ✉

Ming Wei
School of Psychology, Liaoning Collaborative Innovation Center of Children and Adolescents Healthy Personality Assessment and Cultivation, Liaoning Normal University, Dalian, China ✉

Yulin Gao
Department of Psychology, Jilin University, Changchun, China ✉

Inhibition of return (IOR) is a mechanism of the attention system involving bias toward novel stimuli and delayed generation of responses to targets at previously attended locations. According to the two-component theory, IOR consists of a perceptual component and an oculomotor component (oculomotor IOR [O-IOR]) depending on whether the eye movement system is activated. Previous studies have shown that multisensory integration weakens IOR when paying attention to both visual and auditory modalities. However, it remains unclear whether the O-IOR effect attenuated by multisensory integration also occurs when the oculomotor system is activated. Here, using two eye movement experiments, we investigated the effect of multisensory integration on O-IOR using the exogenous spatial cueing paradigm. In Experiment 1, we found a greater visual O-IOR effect compared with audiovisual and auditory O-IOR in divided modality attention. The relative multisensory response enhancement (rMRE) and violations of Miller's bound showed a greater magnitude of multisensory integration in the cued location compared with the uncued location. In Experiment 2, the magnitude of the audiovisual O-IOR effect was significantly less than that of the visual O-IOR in single visual modality selective attention. Implications for the effect of multisensory integration on O-IOR were discussed under conditions of oculomotor system activation, shedding new light on the two-component theory of IOR.

Citation: Tang, X., Yuan, M., Shi, Z., Gao, M., Ren, R., Wei, M., & Gao, Y. (2022). Multisensory integration attenuates visually induced oculomotor inhibition of return. *Journal of Vision*, *22*(4):7, 1–17, https://doi.org/10.1167/jov.22.4.7.

https://doi.org/10.1167/jov.22.4.7 Received June 17, 2021; published March 17, 2022 ISSN 1534-7362

Introduction

In a spatial cueing paradigm, a transient and spatially uninformative peripheral cue is presented before a target appears at either the same (cued) or opposite (uncued) location as the cue. Previous studies found that the response to a target at the cued location was slower than that at the uncued location when the cue-target stimuli onset asynchrony (SOA) was greater than 300 ms (Klein, 2000; Klein, 1988; Posner & Cohen, 1984), which is termed inhibition of return (IOR). The IOR has also been found in eye movement experiments in which the saccade reaction time (SRT) to a target presented at a cued location is slower than that at an uncued location (Abrams & Dobkin, 1994; Jayaraman, Klein, Hilchey, Patil, & Mishra, 2016; Posner, Rafal, Choate, & Vaughan, 1985).

The two-component inhibition theory has been proposed and suggests that IORs may originate from different components or processing stages between suppression and activation of the oculomotor system (Hilchey, Hashish, MacLean, Satel, Ivanoff, & Klein, 2014a; Hilchey, Klein, & Satel, 2014b; Hunt & Kingstone, 2003; Jayaraman et al., 2016; Łukasz, Michalczyk, Jacek, & Bielas, 2019; MacInnes, 2017; Posner et al., 1985). Specifically, when the oculomotor system was suppressed, the perceptual component – the processing stage of response selection – was inhibited. When the oculomotor system is activated, the oculomotor component – the processing stage of response execution – is inhibited (oculomotor IOR [O-IOR]; Henderson & Luke, 2012; Jayaraman et al., 2016; Redden, Hilchey, & Klein, 2018; Ro, Pratt, & Rafal, 2000). A few studies have noted that the perceptual component and oculomotor component are difficult to separate (Kavyani, Farsi, Abdoli, & Klein, 2016; Souto & Kerzel, 2009). However, most studies demonstrate that the two components are independent of each other (Hilchey, Ivanoff, Taylor, & Klein, 2011; Hilchey et al., 2014a; Łukasz et al., 2019; Sumner, Nachev, Vora, Husain, & Kennard, 2004). For example, Hilchey et al. (2014a) used eye movement monitoring and further demonstrated that suppressing eye movements in keypress tasks changes the origin of IOR from output (response executive process) to input (response selective process).

In the condition of suppressing the oculomotor system, the IOR effect has been observed to range from unimodal to cross-modal or bimodal stimuli (Tang, Gao, Yang, Ren, Wu, Zhang, & Wu, 2019; Van der Stoep, Spence, Nijboer, & Van der Stigchel, 2015; Xu, Yang, Zhou, & Ren, 2020). In studies of bimodal stimuli, the information of simultaneous spatial and temporal visual and auditory stimuli was integrated into a coherent perceptual object via the process of multisensory integration (MSI; Atilgan, Town, Wood, Jones, Maddox, Lee, & Bizley, 2018; Colonius & Diederich, 2012; Diederich & Colonius, 2019; Meredith & Stein, 1986; Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010). The influences of multisensory integration on the IOR effect have been demonstrated (Hanna, Schneider, Engel, & Daniel, 2012; Santangelo, Ho, & Spence, 2008; Van der Stoep, Van der Stigchel, & Nijboer, 2015; Santangelo et al., 2008; Tang et al., 2019). Using the visual spatial cues paradigm, Van der Stoep, Van der Stigchel, Nijboer, and Spence (2017) found an IOR effect with visual targets but not audiovisual targets. Tang et al. (2019) further found that the audiovisual IOR effect was attenuated by multisensory integration only when participants were asked to pay attention to multiple modalities rather than a single modality.

IOR is a mechanism for biasing the visual system to facilitate the efficiency of searching for novel information (Klein, 1988; Klein, 2000; Posner & Cohen, 1984). The biased processing of response selection results in fewer attentional resources and then impairs the perceptual component of targets at the cued location in conditions of suppressing the oculomotor system (Mcdonald, Hickey, Green, & Whitman, 2009; Mcdonald, Ward, & Kiehl, 1999; Prime & Ward, 2004; Prime & Ward, 2006; Satel, Hilchey, Wang, Story, & Klein, 2013). However, the integration of visual and auditory stimuli can attract more attention and enhance the perceptual salience of targets (Talsma et al., 2010; Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008; Van der Burg, Talsma, Olivers, Hickey, & Theeuwes, 2011). When IORs (decreasing perceptual salience) encounter multisensory integration (increasing perceptual salience), a reduced IOR effect elicited by an audiovisual target has been found (Tang et al., 2019). Based on these findings, multisensory integration attenuated the processing of response selection of IORs when the oculomotor system was suppressed under the condition of paying attention to multiple modalities. However, the following question remains: does multisensory integration only affect the process of response selection but not the process of response execution of IOR?

In the condition of activating the oculomotor system, Makovac et al. used an exogenous spatial cue paradigm to explore the effect of multisensory integration on the execution of saccades. They found that the automatically integrated audiovisual target signals contained more excitatory inputs at the location of the saccade target so that the suppression of distractors could be eliminated more quickly (Makovac, Buonocore, & Mcintosh, 2015). In other words, multisensory integration promotes a saccadic execution response. According to the two-component theory, IOR includes the perceptual component of inhibiting response selection and the oculomotor component of inhibiting response execution, which are independent of each other (Chica, Taylor, Lupiáñez, & Klein, 2010; Taylor & Klein, 2000). Consequently, we hypothesize that multisensory integration could affect oculomotor components of IOR. A previous study found that the enhanced effect of multisensory integration on the perceptual process could encounter the attenuated effect of IOR on the selective response process to produce the attenuated effect of multisensory integration on IOR when the oculomotor system was suppressed (Van der Stoep et al., 2015; Tang et al., 2019; Tang, Wang, Peng, Li, Zhang, Wang, & Zhang, 2021). Thus, we hypothesize that the O-IOR effect could also be affected by multisensory integration when the oculomotor system was activated. The enhanced effect of multisensory integration on saccadic executive processes could encounter the attenuated effect of IOR on the response execution process to produce the attenuated effect of multisensory integration on O-IOR.

In the present eye movement study, we adapted the exogenous spatial cueing paradigm in multisensory integration contexts. We aimed to confirm the effect of multisensory integration on O-IOR under the condition of activating the oculomotor system by comparing audiovisual O-IOR and visual O-IOR. In this procedure, we manipulated the modalities of the target stimuli (including visual and audiovisual modalities) and cue validity (cued and uncued). According to our previous study on eye movement suppression, a weaker IOR effect with audiovisual targets compared with visual targets was found in divided modality attention (attending to multiple modalities) but not in modality-specific selective attention (attending to a single visual modality; Tang et al., 2019). Therefore, we conducted two experiments to study the influence of multisensory integration on O-IOR in divided modality attention (Experiment 1) and modality-specific selective attention (Experiment 2). Due to the enhancement effect of multisensory integration, we hypothesized that it could attenuate O-IOR under conditions of oculomotor system activation.

Experiment 1: Divided modality attention

Experiment 1a

Method

Participants: Twenty-eight undergraduate students (3 men and 25 women; age 18–26 years) participated in the present experiment. Sample size calculations performed using G*Power software (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007) revealed that a sample of 20 participants was required to detect a medium effect size of $\eta^2 = 0.25$ ($\alpha = 0.05$; $1\text{-}\beta = 0.80$) in Experiment 1. All participants were naive to the goal of the experiment. They received some rewards for their participation. All subjects had normal or corrected-to-normal vision and had no hearing problems. All participants were right-handed and had no history of neurological or psychiatric disorders. The experimental protocol was approved by the Ethics Committee of Liaoning Normal University.

Apparatus and stimuli: Participants were seated in a dark soundproofed room at approximately 60 cm from a 19-inch monitor (60 Hz refresh rate, resolution 1024 × 768) using a chin and forehead rest. A Dell p1914s PC running E-prime 2.0 controlled stimulus presentation. Movements of the left eye were registered using a desktop mounted Eyelink 1000 Plus infrared reflection system (SR Research, 12 Canada), sampling at 1000 Hz. A standard nine-point grid calibration was performed at the beginning of each block, and the calibration was assessed at the beginning of each trial. The experimenter initiated each trial by pressing a specified button on a keyboard.

All visual stimuli were presented on a black background. The fixation stimulus was a filled gray circle (0.67 degrees in diameter, red, green, blue [RGB] = 128, 128, and 128). A white square acted as the peripheral cue (2.5 degrees, RGB = 255, 255, and 255), and the larger and lighter fixation dot was extended to 1.42 degrees to serve as a central cue. The visual target was a white solid rectangle (0.58 degrees), and the auditory target was a pure tone (65 dB, 1000 Hz) generated and handled by Sound Engine Free. The simultaneously presented solid rectangle and tone were audiovisual targets that appeared in the left or right location 10 degrees away from the center cue (Figure 1a).

Procedure and design: Participants were informed that their head and body should remain as still as possible in the same group of tests. Participants performed the experimental test in target modality conditions of visual (V) target, auditory (A) target and audiovisual (AV; visual and auditory stimuli simultaneously appeared) targets. There were 10 blocks of each modality condition in which targets were presented to the left or the right of the fixation dot in a randomized order. The participants were asked to detect targets in three modal conditions, which were presented on the cued or uncued locations followed by the preceding visual peripheral cues.

Each trial began with a drift correction followed by the fixation cross presented at the center of the screen for 800 to 1000 ms. Participants were instructed to keep their eyes on the fixation stimulus until a target appeared. Subsequently, a visual peripheral cue appeared for 100 ms and was randomly located to the left (50%) or right (50%) of the fixation point. Then, consistent with our previous research, we increased the central reorientation such that it is more effective to generate IOR effects (Tang et al., 2019). The fixation point became larger and lighter with the flicker to

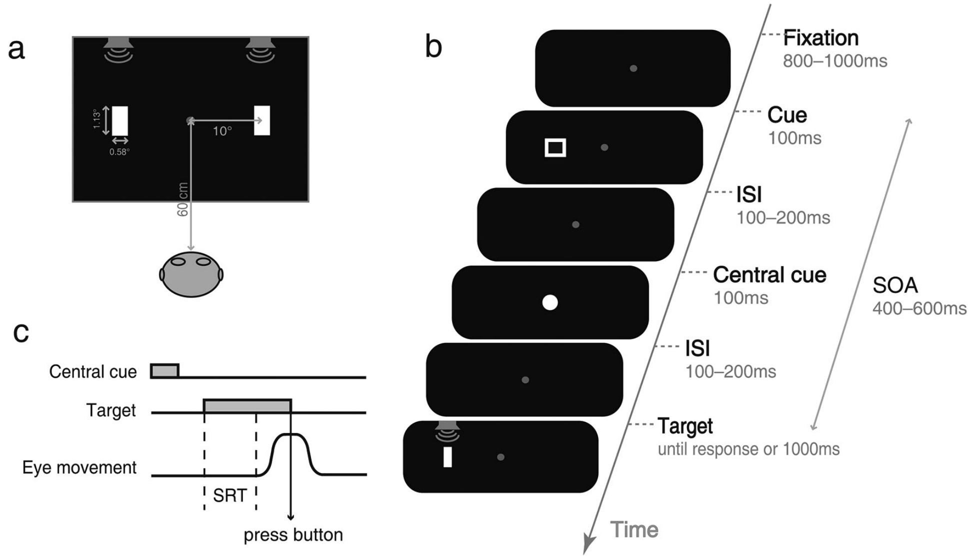

Figure 1. Illustration of the procedure and stimuli of experiment 1a. (**A**) Example of the distance of stimuli and participants. (**B**) The sequence of stimuli comprising a trial is shown from the beginning of the trial to a response. The figure additionally shows the size and position of the experiment. The cue and target were presented to the left or right location, and participants were asked to ignore the sound stimuli. ISI refers to the interstimulus interval. (**C**) Example of the saccade reaction time (SRT), which was defined as the saccade latency of the time from the appearance of the visual target to the beginning of saccade.

serve as a central cue to summon attention back to the central location for 300 to 500 ms. There were two interstimulus intervals for 100 to 200 ms and a central cue for 100 ms followed by a peripheral cue. The SOA, which is defined as the time period between the onset of the cue and the onset of the target, was 400 to 600 ms. After the 400- to 600-ms delay, a visual, auditory, or audiovisual target randomly appeared in one of the cue locations until the participant made a saccade and key press response or 1000 ms had elapsed (Figure 1b illustrates the procedure). Participants were asked to make a saccadic localization response to targets. When the target stimuli appeared, participants moved their eyes from the central fixation to the visual target as quickly and accurately as possible. In addition, we asked subjects to make manual keypress responses to ensure that the subjects completed the auditory localization task well. The visual target and AV target also required keypress responses to certify the consistency of the task under various experimental conditions. After the end of the saccades, participants were asked to press the "F" key with the index finger of the left hand or the "J" key on the keyboard with the index finger of the right hand to indicate whether the target was presented at left location or right location, respectively. The target location (left and right) and target modality (A, V, and AV) were pseudorandomly presented to ensure that no more than three repeated stimuli appeared continuously, which avoided the error of the habitual reactions.

The design was completely within-subject with the following factors randomly intermixed within each block of trials: target modality (with 3 levels: A target, V target, and AV target) and cue validity (with 2 levels: cued and uncued, referring to whether the target's location or direction was the same as or different from the location of the cue). In each block of trials, there were 36 trials in this design. Prior to the experimental trials, the practice trials were designed to familiarize the participants with the test process, including 16 trials, which could be repeated if participants were unfamiliar with the process. After the end of each block, given the correct number of feedbacks, the participants could take a short break between each block. The total duration of the experiment was approximately 30 minutes.

Data recording and analysis: For each trial, participants were instructed to saccade toward the V target, A target, and AV target presented in the left or right location as quickly and accurately as possible and then press the response button to locate targets well. The auditory eye movement data could not be further analyzed due to inferior availability (9%, the area of interest at the same position as the visual target was 2 degrees; 20%, the area of interest expanded to 5 degrees). The saccade latency (SRT) and eye movement amplitudes were used to analyze multisensory response enhancement (MRE) under different cue validity conditions. The velocity threshold to detect saccadic eye movement was set to 30 degrees/second, and the saccadic latencies were computed by subtracting the time at which the eye movement exceeded the velocity threshold from the time at which the imperative stimulus appeared onscreen (Hilchey et al., 2014a; see Figure 1c). The saccade

		Experiment 1a		Experiment 1b	
Target type	Validity	SRT (ms)	Amplitude (°)	SRT (ms)	Amplitude (°)
AV	Cued	382 (40)	9.80 (0.53)	172 (57)	9.74 (0.59)
	Uncued	379 (44)	9.70 (0.55)	155 (50)	9.53 (0.43)
	O-IOR	4		17^{***}	
V	Cued	417 (46)	9.74 (0.57)	233(75)	9.55 (0.58)
	Uncued	392 (44)	9.65 (0.54)	200 (72)	9.35 (0.50)
	O-IOR	25^{***}		33^{***}	
A	Cued			229 (66)	9.58 (0.56)
	Uncued			210 (60)	9.21 (0.46)
	O-IOR			19^{**}	

Table 1. Mean saccade reaction time (SRT, ms), amplitude (degrees), and standard deviation (SD) for each condition. AV, audiovisual target; V, visual target; A, auditory target.
"O-IOR" was obtained by subtracting the saccade reaction time in the uncued location from that in the cued location (ms), that is, cued minus uncued ($^{**}p < 0.01$; $^{***}p < 0.001$).

amplitude following target onset was considered for analysis, which corresponds to the difference between the initial and final eye positions. Trials with SRT more than three standard deviations above the mean in each condition for a given subject were excluded.

Only the trials with the correct manual response and available saccadic response (within a 2 degree area of the target of interest) were retained. Two participants were excluded from the analyses due to less than 75% valid data. For each effective participant, we calculated the SRT and saccade amplitudes for further analysis. To ensure the validity of the data, trials were rejected when (1) the trials had an incorrect manual response (1.90%), (2) trials were unrecorded (7.4%), and (3) the trials had saccade reaction times less than 100 ms and greater than 1000 ms (0.03%) because the results were assumed to be due to anticipation or not paying attention to the task, respectively. For amplitude, individual trials were manually removed when (1) the trials of incorrect manual response (1.90%) and (2) the saccade amplitude was less than 5 degrees and greater than 15 degrees (less than or greater than half of the target amplitude, 1.66%).

We used the relative amount of multisensory response enhancement (rMRE) to investigate the amount of speedup in the bimodal condition compared with the unimodal condition (Van der Stoep et al., 2015; Stevenson, Ghose, Fister, Sarko, Altieri, Nidiffer, Kurela, Siemann, James, & Wallace, 2014; Tang et al., 2019; Tang, Wang, Peng, Li, Zhang, Wang, & Zhang, 2021). The rMRE for each participant in cued and uncued conditions was calculated based on the following formula using median SRT:

$$rMRE = \frac{\text{median}(\text{SRT}_\text{V}) - \text{median}(\text{SRT}_\text{AV})}{\text{median}(\text{SRT}_\text{V})} \times 100\% \quad (1)$$

$$rMRE = \frac{\min(\text{median}(\text{SRT}_\text{A}), \text{median}(\text{SRT}_\text{V})) - \text{median}(\text{SRT}_\text{AV})}{\min(\text{median}(\text{SRT}_\text{A}), \text{median}(\text{SRT}_\text{V}))} \times 100\% \quad (2)$$

In short, the SRT and saccade amplitudes of each participant were compared using a two (target modality: V target and AV target) × 2 (cue validity: cued and uncued) repeated-measures ANOVA. Using the abovementioned formulas, the amount of rMRE was calculated based on the cued and uncued conditions. The paired *t*-test was used to compare differences in rMRE between cued and uncued conditions.

Results

Saccade response time: The mean SRTs are shown in Table 1. A 2 × 2 repeated-measures ANOVA was used to analyze the SRT with the factors target modality (V target and AV target) and cue validity (cued and uncued). A main effect of target modality ($F(1,27) = 72.46$, $p < 0.001$, $\eta_p^2 = 0.73$) was noted, and responses to the AV target ($M = 380.43$ ms, $SD = 42.08$) were significantly faster than those to the V target ($M = 404.27$ ms, $SD = 44.90$). The main effect of cue validity was significant ($F(1, 27) = 29.27$, $p < 0.001$, $\eta_p^2 = 0.52$), indicating that the responses in the cued condition ($M = 399.07$ ms, $SD = 44.03$) were slower than those in the uncued condition ($M = 384.49$ ms, $SD = 43.92$). This finding suggested that IOR occurred. A significant interaction was also between target modality and cue validity ($F(1, 27) = 61.60$, $p < 0.001$, $\eta_p^2 = 0.70$). Post hoc analyses of the target modality times cue validity interaction revealed that when the target modality was a visual stimulus, subjects exhibited a faster response at an uncued location compared with a cued location ($p < 0.001$). This finding suggested that IOR occurred in

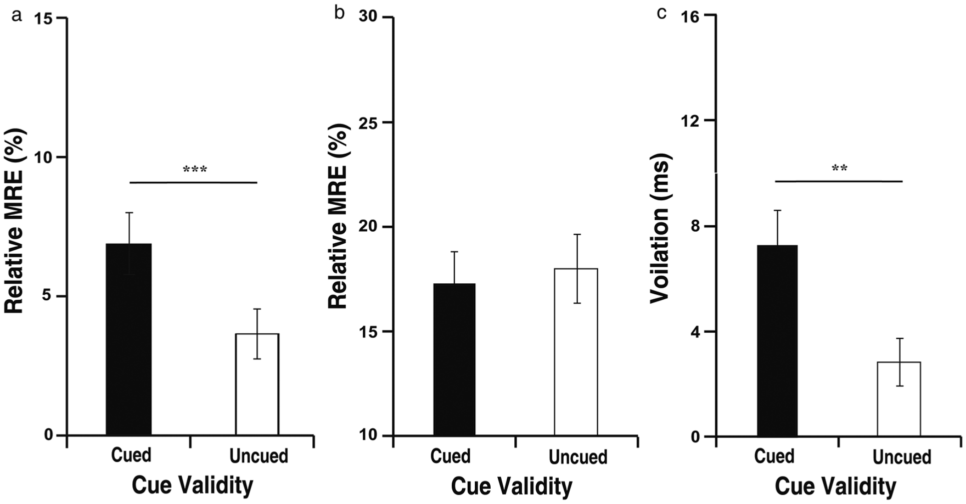

Figure 2. The result of the multisensory integration effect in Experiment 1. (**A**) Magnitude of the relative multisensory response enhancement (rMRE) for each condition in experiment 1a. (**B**) Magnitude of the rMRE in both cue validity conditions in experiment 1b. (**C**) The violation of Miller's bound in experiment 1b. The error bars represent the standard errors of the mean (**$p < 0.01$, ***$p < 0.001$).

the visual target condition. However, when the target modality was an audiovisual stimulus, no significant difference was noted between the response of the cued position and the uncued position ($p = 0.25$), suggesting that the IOR disappeared in the audiovisual target condition.

Multisensory response enhancement: One-sample t-tests showed that the amount of the rMRE was significantly different from zero in both cued (7.97%, $t(27) = 8.66$, $p < 0.001$) and uncued (3.77%, $t(27) = 4.36$, $p < 0.001$) conditions, as shown in Figure 2a. Differences in rMRE between the cued and uncued conditions were analyzed using a paired samples t-test, revealing a greater rMRE (7.97% vs. 3.77%; $t(27) = 4.98$, $p < 0.001$, $d = 0.95$) in the cued condition compared with the uncued condition.

Saccade amplitude: A 2 × 2 repeated-measures ANOVA was used to analyze the amplitude data with the target modality (V target and AV target) and cue validity (cued and uncued) factors. The effect of target modality was significant ($F(1,27) = 7.46$, $p = 0.011$, $\eta_p^2 = 0.22$), and responses to the AV target ($M = 9.75$ degrees, $SD = 0.47$) were longer than those to the V target ($M = 9.70$ degrees, $SD = 0.45$). The main effect of cue validity was significant ($F(1, 27) = 18.50$, $p < 0.001$, $\eta_p^2 = 0.41$), indicating that the responses in the cued condition ($M = 9.77$ degrees, $SD = 0.48$) were significantly longer than those in the uncued condition ($M = 9.68$ degrees, $SD = 0.47$). A significant interaction between target modality and cue validity was not observed ($F(1, 27) = 0.22$, $p = 0.64$, $\eta_p^2 = 0.01$).

Experiment 1b

In experiment 1a, we investigated the effect of multisensory integration on oculomotor IOR in divided modality attention. The results showed a significant visual O-IOR effect but not a significant audiovisual O-IOR effect. These findings may suggest that the experimental requirements of the dual task (saccadic and manual response) did not induce an audiovisual O-IOR effect. To this end, in experiment 1b, we added placeholders on both sides of the screen where the cue and target will appear. Subjects were required to saccade toward the auditory target (saccade toward the placeholder), visual target and audiovisual target as quickly and accurately as possible without pressing keys. Thus, in this case, the impact of the dual task could be excluded, and the saccadic data of the auditory condition could be recorded, further verifying the existence of multisensory integration.

Method

Twenty-two undergraduate students (8 men and 14 women; age 20–30 years) participated in the experiment. All participants were naive to the goal of the experiment. They received some rewards for their participation. All subjects had normal or corrected-to-normal vision and had no hearing problems. Participants had no history of neurological or psychiatric disorders. The procedure and design were similar to those in experiment 1a. In contrast to experiment 1a, the stimuli, apparatus and task were adjusted in the current experiment 1b. Placeholders were added on both sides of the screen. The cue becomes a white rectangle to fill the placeholder, and visual targets appear in the placeholders. Participants were required to saccade toward the stimuli in the placeholders as quickly and accurately as possible regardless of whether the target was an auditory modality, visual modality, or audiovisual modality (saccade toward placeholder when the target was invisible auditory stimulus). The target disappears automatically after 800 to 1000 ms.

We used SRT to analyze violations of Miller's bound, following the approach described with the RSE-box (Miller, 1982; Otto, 2019; Liu & Otto, 2020). First, we ordered SRTs from the fastest to the slowest and computed corresponding cumulative probabilities to obtain cumulative distribution functions (CDFs) in each condition for each participant (using the RSE-box function getCP). Then, we downsampled these distributions to 50 quantiles using linear interpolation (using the RSE-box function interpCDF). To obtain Miller's bound, we summed the corresponding unisensory CDFs (using the RSE-box function getMiller). Violations of Miller's bound are quantified using a geometrical approach, which is averaged at each quantile (Colonius & Diederich, 2006). This final step is obtained using the RSE-box function getViolation, which quantifies the size of the violation area using the following formula (for more details about the RSE-box, see Otto, 2019):

$$\text{Violation} = \frac{1}{50}\sum\nolimits_{q=1}^{50} \max\left(Miller_q - AV_q,\ 0\right) \quad (3)$$

In the formula, Miller_q was obtained by Miller's bounds at the same 50th quantile as the empirical CDFs (where q is the index of the 50 quantiles). Using audiovisual signals, we obtained a vector of quantile SRTs named AV_q. For statistical analysis, the differences in violations between the cued and uncued conditions were analyzed using a paired samples *t*-test.

Results

Saccade response times: The mean SRTs are shown in Table 1. A 3 × 2 repeated-measures ANOVA was used to analyze the SRT with the factors target modality (A target, V target, and AV target) and cue validity (cued and uncued). A main effect of target modality ($F(2,42) = 62.71$, $p < 0.001$, $\eta_p^2 = 0.75$) was noted, and responses to the AV target ($M = 163.49$ ms, $SD = 53.48$) were significantly faster than those to the V target ($M = 216.34$ ms, $SD = 69.66$) and A target ($M = 219.41$ ms, $SD = 63.17$). However, no significant difference was noted between visual and auditory responses. The main effect of cue validity was significant ($F(1, 21) = 34.99$, $p < 0.001$, $\eta_p^2 = 0.63$), indicating that the responses in the cued condition ($M = 211.09$ ms, $SD = 66.13$) were slower than those in the uncued condition ($M = 188.40$ ms, $SD = 58.08$), suggesting that IOR occurred. A significant interaction was noted between target modality and cue validity ($F(2, 42) = 5.78$, $p = 0.006$, $\eta_p^2 = 0.22$). Post hoc analyses of the target modality times the cue validity interaction revealed that subjects exhibited a faster response at the uncued location compared with the cued location (p values < 0.01) in three target modalities, suggesting that O-IOR occurred in the auditory target condition (19.15), visual target condition (32.44), and audiovisual target condition (16.48). The oculomotor IOR with the auditory target and audiovisual target was significantly less than that obtained for the visual target (p values < 0.5).

Multisensory response enhancement: One-sample *t*-tests showed that the amount of rMRE was significantly different from zero in both cued (17.24%, $t(21) = 10.98$, $p < 0.001$) and uncued (17.97%, $t(21) = 10.92$, $p < 0.001$) conditions, as shown in Figure 2b. Differences in rMRE between the cued and uncued conditions were analyzed using a paired samples *t*-test, which did not find a significant difference between the cued condition and the uncued condition ($t(21) = -0.39$, $p = 0.70$, $d = 0.08$).

Miller's bound: To test the effectiveness of multisensory integration, we investigated violations of Miller's bound (Liu & Otto, 2020; Miller, 1982; Otto, Dassy, & Mamassian, 2013). One-sample *t*-tests showed that the violations of Miller's bound were significantly different from zero in both cued (7.24 ms, $t(21) = 5.31$, $p < 0.001$) and uncued (2.83 ms, $t(21) = 3.21$, $p = 0.004$) conditions, as shown in Figure 3. Differences in violations between the cued and uncued conditions were analyzed using a paired samples *t*-test in Figure 2c, revealing that the violations of Miller's bound in the cued condition were significantly greater than those in the uncued condition ($t(21) = 2.91$, $p = 0.009$, $d = 0.63$). According to the results of rMRE and violations, we demonstrated the existence of multisensory integration, and the magnitudes of the multisensory integration effect in the cued condition were greater than those in the uncued condition.

Saccade amplitude: A 3 × 2 repeated-measures ANOVA was used to analyze the amplitude data with the target modality (A target, V target, and AV target) and cue validity (cued and uncued) factors. The effect of target modality was significant ($F(2,42) = 8.29$, $p = 0.001$, $\eta_p^2 = 0.28$), and responses to the AV target ($M = 9.64$ degrees, $SD = 0.51$) were longer than those to the A target ($M = 9.39$ degrees, $SD = 0.51$) and V target ($M = 9.45$ degrees, $SD = 0.54$). The main effect of cue validity was significant ($F(1, 21) = 8.23$, $p = 0.009$, $\eta_p^2 = 0.28$), indicating that the responses in the cued condition ($M = 9.63$ degrees, $SD = 0.58$) were significantly longer than those in the uncued condition ($M = 9.36$ degrees, $SD = 0.47$). A

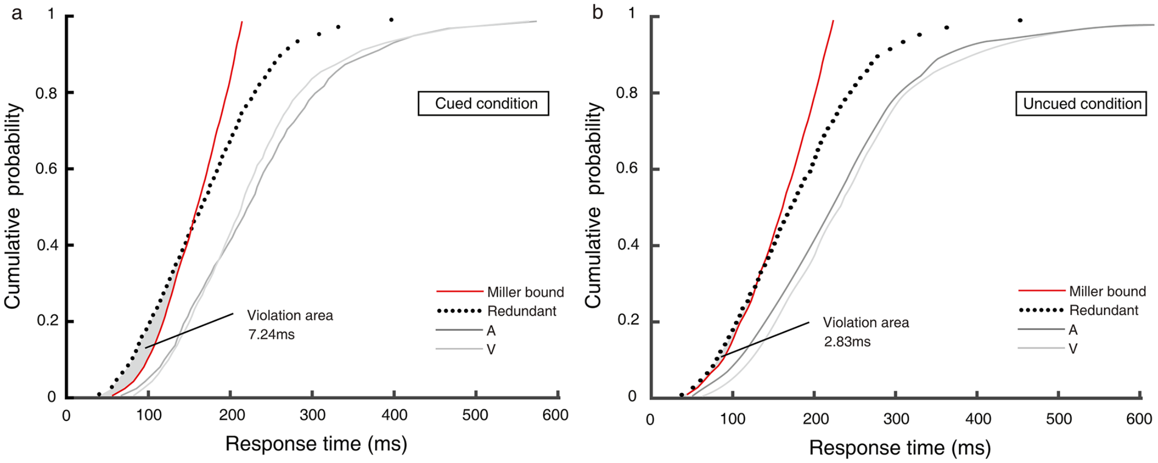

Figure 3. (**A**) Miller's bound in a cued condition (ms). The empirical SRT distribution with redundant signals is given by black dots. Violations of Miller's bound are quantified by the area between the empirical distribution and the bound (gray area). (**B**) Violations of Miller's bound in an uncued condition (ms).

significant interaction was not noted between target modality and cue validity ($F(2, 42) = 2.23$, $p = 0.12$, $\eta_p^2 = 0.096$).

Discussion

The main purpose of Experiment 1 was to determine the effect of multisensory integration on the O-IOR effect in divided modality attention. In experiment 1a, a significant visual O-IOR effect was noted, whereas a significant audiovisual O-IOR effect was not observed. The results of no significance on O-IOR of audiovisual targets in Experiment 1 are consistent with Van der Stoep et al. (2015) and inconsistent with Tang et al. (2019). The present research on the IOR effect using location tasks is different from Tang et al. (2019), who used simple detection tasks, and Van der Stoep et al. (2015), who used Go-NoGo detection tasks. Klein and Taylor observed that detection responses inherently involve some spatial localization and that inhibition of return should occur with simple detection responses as well as choice responses involving spatial localization (Klein & Taylor, 1994). Thus, it seems unlikely that the difference in tasks could account for the disappearance of the audiovisual IOR effect. Some evidence suggests that adding a visual central re-orienting cue is prone to induce the IOR effect by summoning attention back to the central location (Peng, Chang, Li, Wang, & Tang, 2019; Tang et al., 2019). The study of Van der Stoep et al. (2015) did not add a central re-orientation to attract participants' attention back to the central location in the exogenous cue-target paradigm. It is possible that the cuing effect found by Van der Stoep and his colleagues would not be intense enough to resist the enhancement of audiovisual integration. In the condition of adding central reorienting cues, Tang et al. (2019) only found that the audiovisual IOR effect was less than the visual IOR effect, but we found that the audiovisual IOR disappeared, showing that the central reorientation cue did not completely explain the discrepancy in the results of audiovisual IOR.

In experiment 1b, the response to the audiovisual target was significantly faster than those to the visual target and auditory target. There was no significant difference between the response to the visual target and the auditory target. We found a significant O-IOR effect with the auditory target, visual target, and audiovisual target. The magnitude of the O-IOR effect with the visual target was larger than that with the visual target and audiovisual target. In addition, the difference in the audiovisual O-IOR effect between experiment 1a and experiment 1b may be caused by the requirement of dual tasks (saccade and manual response). Experiment 1a asked subjects to saccade toward the target and then press the keys on the keyboards to locate targets. Such a dual task may lead to a slower saccade response to target recognition, and the audiovisual O-IOR effect may be submerged in the longer response. In experiment 1b, the audiovisual O-IOR effect was found when we asked participants to saccade toward targets and excluded the manual response.

Importantly, the rMRE and the violations of Miller's bound both showed the existence of a multisensory integration effect. The magnitude of the multisensory integration effect in the cued condition was greater than that in the uncued condition (Colonius & Diederich, 2012; Diederich & Colonius, 2019; Miller, 2016; Wang, Blohn, Huang, Boehnke, & Munoz, 2017), which was inconsistent with prior findings (Tang et al., 2019; Van der Stoep et al., 2015; Van der Stoep et al., 2015). According to the hypothesis of perceptual sensitivity caused by exogenous cues, exogenous spatial attention could decrease perceptual sensitivity at cued locations compared with uncued locations in long SOAs (400–600 ms; Peng et al., 2019; Tang et al., 2019; Van der Stoep et al., 2015). As a result, the perceptual sensitivity at the cued location was weaker than that at the uncued location. Consistent with the principle of inverse effectiveness, the benefit of audiovisual integration is greater for weaker stimuli compared with stronger stimuli (Otto et al., 2013; Senkowski, Saint-Amour, Höfle, & Foxe, 2011). Then, the benefit of audiovisual integration is greater for stimuli presented at cued locations than for those presented at uncued locations.

In the present experiment, the oculomotor system was activated in divided modality attention. The audiovisual O-IOR effect could restrain the oculomotor components of the inhibiting execution stage. The multisensory integration effect will guide the subject's attention to the location of the visual stimulus (Talsma et al., 2010). Moreover, auditory signals not only enhance the early perception processing of simultaneous visual events (Van der Burg et al., 2008; Van der Burg et al., 2011) but also enhance the late execution processing of simultaneous visual events (Mondor, Terrio, & Hurlburt, 2000). In short, multisensory integration enhances the perception process and the execution process, offsetting the inhibition of the execution process and leading to the attenuation or even disappearance of O-IOR. Therefore, we suggest that the attenuation or even disappearance of audiovisual O-IOR could be caused by multisensory integration in divided modality attention. Experiment 2 was designed to confirm whether selective attention to single modalities modulates the effect of multisensory integration on O-IOR.

Experiment 2: Modality-specific selective attention

Experiment 2a

Method

Participants: Twenty-six undergraduate students (10 men and 16 women; age 18–26 years) participated in the experiment. A sample size calculation conducted using G*Power software (Faul et al., 2007; Faul et al., 2009) revealed that a sample of 24 participants was required to detect a medium effect size of $\eta^2 = 0.25$ ($\alpha = 0.05$; $1\text{-}\beta = 0.80$) in Experiment 2. All participants were naive to the goal of the experiment. Participants received some rewards for their participation. All subjects had normal or corrected-to-normal vision and no hearing problems. Participants were all right-handed and had no history of neurological or psychiatric disorders. The experimental protocol was approved by the Ethics Committee of Liaoning Normal University.

Experiment 2 was performed to examine the O-IOR effect when selectively paying attention to visual targets. The apparatus and procedure used were identical to Experiment 1, except for the target modality. Only visual targets and audiovisual targets were presented in the current eye movement localization task. Participants were required to pay attention to visual stimuli but to ignore auditory stimuli in modality-specific selective attention.

Data recording and analysis: The data recording and analysis were identical to Experiment 1. Six participants were excluded from the analyses because less than 75% valid data were available. To ensure the validity of the data, trials were rejected when (1) the trials had an incorrect manual response (2.2%), (2) the trials were unrecorded (6.1%), and (3) the trials had saccade reaction times less than 100 ms and greater than 1000 ms (0.03%) because they were assumed to be the result of anticipation or not paying attention to the task, respectively. In total, 8.36% of SRT was removed. For saccade amplitude, individual trials were removed when (1) the trials included an incorrect manual response (2.89%) and (2) saccade amplitude was less than 5 degrees and greater than 15 degrees (less than or greater than half of the target amplitude, 2.66%). In short, the SRT and saccade amplitude data were compared using a 2 (target modality: V target and AV target) × 2 (cue validity: cued and uncued) repeated-measures ANOVA. The amount of aMRE and rMRE was calculated based on the cued and uncued conditions using the formulas mentioned in Experiment 1. A paired t-test was used to compare differences in aMRE or rMRE between cued and uncued conditions.

Results

Saccade response time: The SRT data are shown in Table 2. A 2 × 2 repeated-measures ANOVA was used to analyze the SRT with the factors target modality (V target and AV target) and cue validity (cued and uncued). A main effect of target modality ($F(1,25) = 29.46$, $p < 0.001$, $\eta_p^2 = 0.54$) was noted, and responses to the AV target ($M = 356.50$ ms, $SD = 47.26$) were significantly faster than those to the V target ($M = 382.72$ ms, $SD = 54.72$). The main effect of cue validity was significant ($F(1, 25) = 70.95$, $p < 0.001$, $\eta_p^2 = 0.74$), indicating that the responses in the cued condition ($M = 377.88$ ms, $SD = 52.29$) were slower than those in the uncued condition ($M = 361.30$ ms, $SD = 49.74$). This finding suggested that IOR occurred. A significant interaction was noted between target modality and cue validity ($F(1, 25) = 16.20$, $p < 0.001$, $\eta_p^2 = 0.39$), and the inhibition of return was greater for the V target ($M = 22.66$ ms, $SD = 18.70$)

		Experiment 2a		Experiment 2b	
Target type	Validity	SRT (ms)	Amplitude (°)	SRT (ms)	Amplitude (°)
AV	Cued	362 ± 48	9.80 (0.53)	113 (35)	9.82 (0.80)
	Uncued	351 ± 48	9.70 (0.55)	105 (37)	9.43 (0.60)
	O-IOR	11^{***}		8^{*}	
V	Cued	394 ± 58	9.74 (0.57)	164 (56)	9.75 (0.74)
	Uncued	372 ± 53	9.65 (0.54)	134 (43)	9.27 (0.76)
	O-IOR	22^{***}		30^{***}	

Table 2. Mean saccade reaction time (SRT, ms), amplitude (degrees) and standard deviation (SD) for each condition. AV, audiovisual target; V, visual target.
"O-IOR" was obtained by subtracting the saccade reaction time in the uncued location from that in the cued location (ms), i.e. cued minus uncued (*$p < 0.05$; ***$p < 0.001$).

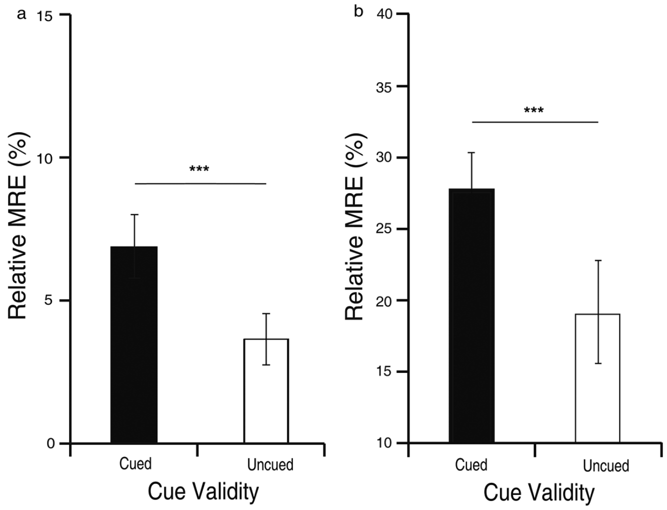

Figure 4. The result of the multisensory integration effect in Experiment 2. (**A**) Magnitude of relative multisensory response enhancement (rMRE) for each condition in experiment 2a. (**B**) Magnitude of rMRE in both cue validity conditions in experiment 2b. The error bars represent the standard errors of the mean (***$p < 0.001$).

compared with the AV target ($M = 11.05$ ms, $SD = 16.12$) condition ($t(25) = -4.03$, $p < 0.001$, $d = 0.79$).

One-sample t-tests showed that the amount of rMRE was significantly different from zero in both cued (6.89%, $t(25) = 6.22$, $p < 0.001$) and uncued (3.65%, $t(25) = 4.07$, $p < 0.001$) conditions, as shown in Figure 4a. Differences in rMRE between the cued and uncued conditions were analyzed using a paired samples t-test. The paired t-test found larger rMRE (6.89% vs. 3.65%; $t(25) = 4.21$, $p < 0.001$, $d = 0.87$) in the cued condition compared with the uncued condition.

Saccade amplitude: A 2 × 2 repeated-measures ANOVA was used to analyze the saccade amplitudes with the target modality (V target and AV target) and cue validity (cued and uncued). The effect of target modality was significant ($F(1,25) = 6.85$, $p = 0.015$, $\eta_p^2 = 0.22$), and responses to the AV target ($M = 9.71$ degrees, $SD = 0.35$) were longer than those to the V target ($M = 9.63$ degrees, $SD = 0.33$). The main effect of cue validity was significant ($F(1, 25) = 11.94$, $p = 0.002$, $\eta_p^2 = 0.32$), indicating that the responses in the cued condition ($M = 9.74$ degrees, $SD = 0.34$) were significantly longer than those in the uncued condition ($M = 9.60$ degrees, $SD = 0.33$). We did not find a significant interaction between target modality and cue validity ($F(1, 25) = 0.95$, $p = 0.34$, $\eta_p^2 = 0.004$).

Experiment 2b

The experimental protocol was similar to that in experiment 2a, and the stimulus parameters were similar to those in experiment 1b. Experiment 2b aimed to exclude the effect of dual tasks and further verify the effect of multisensory integration on oculomotor IOR in single visual modality attention.

Method

Twenty-two undergraduate students (8 men and 14 women; age 20–30 years) participated in the experiment. All participants were naive to the goal of the experiment. Participants received some rewards for their participation. All subjects had normal or corrected-to-normal vision and no hearing problems. Participants had no history of neurological or psychiatric disorders. The experimental protocol was approved by the Ethics Committee of Liaoning Normal University. The procedure was similar to that described for experiment 2a, except that the subjects were not required to press the keys on the keyboard after saccades. We conducted the current experiment using the same stimulus parameters as noted for experiment 1b.

Results

Saccade response times: The SRT data are shown in Table 2. A 2 × 2 repeated-measures ANOVA was used to analyze the SRT with the factors target modality (V target and AV target) and cue validity (cued and uncued). A main effect of target modality ($F(1,21) = 73.59$, $p < 0.001$, $\eta_p{}^2 = 0.78$) was observed, and responses to the AV target ($M = 109.02$ ms, $SD = 36.14$) were significantly faster than those to the V target ($M = 148.85$ ms, $SD = 49.72$). The main effect of cue validity was significant ($F(1, 21) = 24.06$, $p < 0.001$, $\eta_p{}^2 = 0.53$), indicating that the responses in the cued condition ($M = 138.67$ ms, $SD = 45.64$) were slower than those in the uncued condition ($M = 119.20$ ms, $SD = 40.22$). This finding suggested that O-IOR occurred. A significant interaction was noted between target modality and cue validity ($F(1, 21) = 19.28$, $p < 0.001$, $\eta_p{}^2 = 0.48$), and the O-IOR effect was greater for the V target ($M = 30.54$ ms, $SD = 25.54$) than for the AV target ($M = 8.40$ ms, $SD = 17.90$) condition ($t(21) = -4.39$, $p < 0.001$, $d = 0.94$).

One-sample t-tests showed that the amount of rMRE was significantly different from zero in both cued (27.74%, $t(21) = 10.97$, $p < 0.001$) and uncued (19.10%, $t(25) = 5.30$, $p < 0.001$) conditions, as shown in Figure 4b. Differences in rMRE between the cued and uncued conditions were analyzed using a paired samples t-test. The paired t-test found larger rMRE (27.74% vs. 19.10%; $t(21) = 3.38$, $p = 0.003$, $d = 0.72$) in the cued condition compared with the uncued condition.

Saccade amplitude: A 2 × 2 repeated-measures ANOVA was used to analyze the saccade amplitudes with the target modality (V target and AV target) and cue validity (cued and uncued). The effect of target modality was not significant ($F(1,21) = 3.10$, $p = 0.093$, $\eta_p{}^2 = 0.13$). The main effect of cue validity was significant ($F(1, 21) = 18.42$, $p < 0.001$, $\eta_p{}^2 = 0.47$), indicating that the responses in the cued condition ($M = 9.78$ degrees, $SD = 0.77$) were significantly greater than those in the uncued condition ($M = 9.35$ degrees, $SD = 0.68$). A significant interaction between target modality and cue validity was not observed ($F(1, 21) = 0.54$, $p = 0.47$, $\eta_p{}^2 = 0.025$).

Discussion

Identical results of the O-IOR effect were found in experiment 2a and experiment 2b. A significant O-IOR effect with visual targets and audiovisual targets in modality-specific selective attention was noted. The magnitude of the audiovisual O-IOR effect was less than that of the visual O-IOR effect. The results of Experiment 2 showed that the O-IOR effect may be attenuated by multisensory integration in modality-specific selective attention, which is inconsistent with previous studies on the IOR effect under the condition of suppressing the oculomotor system (Tang et al., 2019, see more discussion in the "General Discussion" section).

According to results from the two experiments, the requirements of the dual task and simple saccade task have slight effects on the O-IOR effect in divided modality attention and single visual modality attention. On the one hand, following saccade toward targets, the manual response was performed, which includes eye movement data. The current research focused more on saccade data, which could reflect the results more accurately (Reimer, Tudge, & Schubert, 2021). On the other hand, a similar trend in the O-IOR effect (the magnitude of the O-IOR effect with the visual target was larger than that with the audiovisual target) was noted between the dual task and saccadic response. A recent study found that the manual response led to a delay of saccade execution in saccade eye movements in dual tasking but not to an impairment of the spatial planning of the saccade trajectory in dual tasks (Reimer et al., 2021). We used the spatial cue-target paradigm to discuss the O-IOR effect, which acts through the mechanism of biasing attention toward novel spatial locations. The delay of the dual task only has a slight effect on spatial-based IOR.

In addition, the difference between the responses of cued and uncued conditions may be affected by the response priming effect and Simon effect. Previous studies have proposed that both the response priming effect and Simon effect are involved in the extraction of identity information and manual key press (Bédard, El Massioui, Pillon, & Nandrino, 1993; Hommel, 1993; Hilchey, Leber, & Pratt, 2018; Simon, 1990; Simon, Acosta, & Mewaldt, 1975; Spence & Driver, 1994). However, in the present study, we found that the response of the cued condition was greater than that of the uncued condition in the dual response task, and the same results in saccadic response times were noted when the manual key press was excluded. We believe that the slower saccadic response times to targets at previously attended locations are a mechanism of the attentional system involving bias toward novel stimuli or nonattentional locations (i.e. oculomotor inhibition of the return effect; Abrams & Dobkin, 1994; Klein, 2000; Klein, 1988; Posner et al., 1985).

Our findings in two experiments both showed that the saccade amplitude for the audiovisual target was greater than that for the visual target. This result may indicate that information from multiple modalities (vision and audition) could promote the recognition of targets and extend the cognitive scope (Irwin, 1998; Rayner, 2009). In addition, as a mechanism for biasing the visual system to acquire novel information, O-IOR prevents the execution of the saccade and increases the saccade latency of the cued location compared to the uncued location. The lower top-down attentional activity conflicts with the task of the bottom-up saccade target; thus, a larger saccade amplitude in the cued location (Godijn & Theeuwes, 2002a; Godijn & Theeuwes, 2002b; Mokler & Fischer, 1999). Participants voluntarily resisted the conflict by extending the search execution intensity of the target in the cued location, resulting in a greater amplitude in the cued location. We suggest that future research should observe saccade amplitudes, indicating the inhibition of the return effect during eye movement experiments, which is an effective eye movement indicator in the spatial cue-target paradigm.

General discussion

The primary goal of the current study was to investigate the effect of multisensory integration on the O-IOR effect in divided modality attention and modality-specific selective attention under the condition of an activated oculomotor system. We addressed this question using an exogenous spatial cue paradigm in which subjects were required to determine the spatial location of targets and perform a saccadic localization task. In divided modality attention, we found a significant multisensory integration effect and a smaller O-IOR effect with the audiovisual target compared with the visual O-IOR effect. The results suggest that O-IOR was attenuated by multisensory integration under conditions of activating the oculomotor system when paying attention to multiple modalities. A larger O-IOR effect with the visual target than with the audiovisual target was found in single visual modality selective attention. The effect of multisensory integration on O-IOR was discussed under conditions of oculomotor system activation.

Tang et al. (2019) used the same paradigm and found a comparable IOR effect with visual and audiovisual targets in modality-specific selective attention, demonstrating that multisensory integration has no efficient impact on IOR. The attenuated effect of audiovisual integration on IOR depends on the subject attending to both visual and auditory modalities simultaneously (Lunn, Sjoblom, Ward, Sotofaraco, & Forster, 2019; Mozolic, Hugenschmidt, Peiffer, & Laurienti, 2008; Talsma, Doty, & Woldorff, 2007; Tang et al., 2019; Tang, Wu, & Shen, 2016). In the current eye movement study, however, we not only found that multisensory integration has a weakening effect on O-IOR in specific modality selective attention but also detected the attenuated or even counteracted effect of O-IOR by multisensory integration in divided modality attention. On the one hand, several studies demonstrate that multisensory integration occurs at early or late processing stages depending on the attention resources available (Baart, Stekelenburg, & Vroomen, 2014; Calvert & Thesen, 2004; Koelewijn, Bronkhorst, & Theeuwes, 2010; Talsma, 2015; Talsma & Woldorff, 2005; Tang et al., 2021). Talsma et al. (2007) showed that attending to both modalities is a prerequisite for automatic early integration of the bottom-up process, whereas late integration is not an automated process and requires top-down attention to resource allocation (Koelewijn et al., 2010; Talsma et al., 2007). In current specific modality selective attention, early integration may not occur; only late integration occurred during delayed eye movement overt execution. Regarding divided modality attention, early integration as well as late integration potentially occur. On the other hand, according to the two-component theory, IOR includes a perceptual component of inhibiting the early response selection process in suppressing the oculomotor system and an oculomotor component of inhibiting the late response execution process in activating the oculomotor system (Abrams & Dobkin, 1994; Chica et al., 2010; Hilchey et al., 2014a; Hunt & Kingstone, 2003; Jayaraman et al., 2016; Łukasz et al., 2019; MacInnes, 2017; Souto & Kerzel, 2009).

In divided modality attention, the integration of visual and auditory stimuli occurs in both early and late stages (Koelewijn et al., 2010; Talsma et al., 2007; Talsma et al., 2010; Tang et al., 2021). Multisensory integration can enhance both perceptual salience and the executive process of audiovisual targets. The O-IOR effect might weaken the response execution process of the audiovisual target when activating the oculomotor system. The enhanced effect of perceptual salience and the executive process by multisensory integration could attenuate or even counteract the weakened effect of the response execution process by O-IOR, resulting in the attenuation or even disappearance of IOR with audiovisual targets. In specific modality selective attention, the execution process enhanced by late integration could not affect the early perceptual component of IOR when suppressing the activation of the oculomotor system. Therefore, a similar magnitude of visual IOR and audiovisual IOR was observed (Tang et al., 2019). However, when activating the oculomotor system, late multisensory integration can enhance not only the early perceptual but also the late executive stage to resist the inhibition process; thus, a smaller audiovisual O-IOR is found.

Several researchers hold that selective attention to a single modality eliminated response enhancements associated with multisensory stimuli and weakened the occurrence of multisensory integration (Mozolic et al., 2008; Santangelo et al., 2008). We only observed the relative multisensory response enhancement in the two-cue validity condition with SRT, but the results did not support race model inequality violation in single visual selective attention. The multisensory response enhancement effect in these conditions may not be the result of multisensory integration (Innes & Otto, 2019; Miller, 2016; Otto, 2019; Liu & Otto, 2020). According to this notion, it is very likely that the greater O-IOR effect for the visual target compared with the audiovisual target may not be caused by multisensory integration but by other factors, such as the activation of the oculomotor and the requirement of the dual task (see more details in the discussion of Experiment 2). Previous studies have demonstrated that saccade toward the sound location promoted attention to unattended auditory stimuli (Reuter-Lorenz & Rosenquist, 1996; Mondor, Terrio, & Hurlburt, 2000). The ignored auditory stimuli are still intimately related to the occurrence of the oculomotor system, thereby resisting the audiovisual O-IOR effect and resulting in a smaller magnitude of the audiovisual O-IOR effect compared to the visual O-IOR effect.

In addition, the attenuation and even disappearance of audiovisual O-IOR implies that the perceptual and oculomotor components of IOR could not be independent. Using the psychological refractory period (PRP) paradigm, Klein et al. suggested that IORs inhibit an early response-selective stage of processing (based on input IORs) when eye movement responses are precluded (Kavyani et al., 2016). IOR inhibits a late response execution stage of processing (based on output IOR) when eye movement responses are activated (Klein, Kavyani, Farsi, & Lawrence, 2018). However, several studies have demonstrated that dissociation of oculomotor and attentional components seems implausible because selective attention and the programming of eye movements are tightly coupled (Hilchey, Klein, & Ivanoff, 2012; Souto & Kerzel, 2009). Souto et al. reported that an attention component was found in saccadic inhibition of return (Souto & Kerzel, 2009). Thus, the disappearance of audiovisual O-IOR prompts us to hypothesize that O-IOR may inhibit both the early perception component and the oculomotor component under the condition of eye movement activation. Multisensory integration can occur across multiple neural levels (i.e. at subcortical levels, the level of association cortices and the lowest cortical levels). The two components could be modulated by multisensory integration at different brain neural levels (Tang et al., 2016; Xi, Li, Zhang, Liu, & Wu, 2020). The current research extends the two-component theory, which is worthy of further discussion in future event-related potentials (ERPs) or functional magnetic resonance imaging (fMRI) studies to provide neurophysiological evidence.

Another possible explanation is that the present finding may be caused by multisensory integration more sensitively modulating O-IOR when the oculomotor system was activated. The activation of the oculomotor system is more consistent with the cognitive process in the actual environment, which has higher ecological validity (Greenlee, 2017). When we perform target detection and localization in a natural environment, eye movement is inevitable. Regardless of whether the cognitive system automatically integrates multiple sensory stimuli (MSI) or focuses attention on novel stimuli (IOR), the requirement of eye movement ingratiates the requirements of the natural environment. Thus, we suggest that the present results could reflect the phenomenon that multisensory integration more sensitively modulates O-IOR when the oculomotor system is activated. Future research on the interaction between multisensory integration and exogenous attention should consider the potential effects of oculomotor activation.

Conclusion

The present study investigated the effect of multisensory integration on O-IOR using the exogenous spatial cueing paradigm with two eye movement experiments. The O-IOR effect with the audiovisual target was smaller than that noted with the visual target regardless of whether the participant paid attention to multiple modalities or a single visual modality. Furthermore, the current study demonstrated the existence of a multisensory integration effect using saccadic reaction times. These findings suggest that the O-IOR effect was attenuated by multisensory integration when the oculomotor system was activated. In addition, our results may shed new light on the two-component theory of IOR, suggesting that a nuanced coupled relationship between the perceptual component and oculomotor component was present under the condition of an activated oculomotor system. These findings further verify the interaction between multisensory integration and inhibition of return.

Keywords: oculomotor inhibition of return, multisensory integration, two-component theory, saccade latency, saccade amplitude

Acknowledgments

Supported by a Scientific Research Project of the Education Department of Jilin Province (JJKH20211101KJ, Y.G.); the Sunglory Education Institute (371192001401, Y.G.); the Fundamental Research Funds for the Central Universities (2021ZZ028, Y.G.); the National Natural Science Foundation of China (31871092, X.T.); and the Natural Science Basic Scientific Research Project of Educational Department of Liaoning Province (LJKZ0987, X.T.).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Commercial relationships: none.

Corresponding authors: Yulin Gao, Xiaoyu Tang, Ming Wei.
Emails: gaoyulin@jlu.edu.cn, tangyu-2006@163.com, 870679495@qq.com.
Addresses: Department of Psychology, Jilin University, Changchun 130012, China; School of Psychology, Liaoning Normal University, Dalian 116029, China; School of Psychology, Liaoning Normal University, Dalian 116029, China.

References

Abrams, R. A., & Dobkin, R. S. (1994). Inhibition of return: Effects of attentional cuing on eye movement latencies. *Journal of Experimental Psychology: Human Perception & Performanance, 20*(3), 467–477.

Atilgan, H., Town, S. M., Wood, K. C., Jones, G. P., Maddox, R. K., Lee, A. K. C., . . . Bizley, J. K. (2018). Integration of Visual Information in Auditory Cortex Promotes Auditory Scene Analysis through Multisensory Binding. *Neuron, 97*(3), 640–655. e644.

Baart, M., Stekelenburg, J. J., & Vroomen, J. (2014). Electrophysiological evidence for speech-specific audiovisual integration. *Neuropsychologia, 53*, 115–121.

Bédard, M. A., El Massioui, F., Pillon, B., & Nandrino, J. L.(1993). Time for reorienting of attention: A premotor hypothesis of the underlying mechanism. *Neuropsychologia, 31*, 241–249.

Calvert, G. A., & Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology-Paris, 98*(1–3), 191–205.

Chica, A. B., Taylor, T. L., Lupiáñez, J., & Klein, R. M. (2010). Two Mechanisms Underlying Inhibition of Return. *Experimental Brain Research, 201*(1), 25–35.

Colonius, H., & Diederich, A. (2006). The race model inequality: interpreting ageometric measure of the amount of violation. *Psychological Review, 113*(1), 148–154.

Colonius, H., & Diederich, A. (2012). Focused attention vs. crossmodal signals paradigm: deriving predictions from the time-window-of-integration model. *Frontiers in Integrative Neuroscience, 6*, 62.

Diederich, A., & Colonius, H. (2019). Multisensory integration and exogenous spatial attention: A time-window-of-integration analysis. *Journal of Cognitive Neuroscience, 31*(5), 699–710.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Frens, M. A., Van Opstal, A. J., & Van der Willigen, R. F. (1995). Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Perception & Psychophysics, 57*(6), 802–816.

Godijn, R., & Theeuwes, J. (2002a). Oculomotor capture and Inhibition of Return: Evidence for an oculomotor suppression account of IOR. *Psychological Research, 66*(4), 234–246.

Godijn, R., & Theeuwes, J. (2002b). Programming of endogenous and exogenous saccades: evidence for a competitive integration model. *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1039–1054.

Greenlee, M. (2017). Self-Motion Perception: Ups and Downs of Multisensory Integration and Conflict Detection. *Current Biology, 27*(18), R1006–R1007.

Hanna, K., Schneider, T. R., Engel, A. K., & Daniel, S. (2012). Capture of visual attention interferes with multisensory speech processing. *Frontiers in Integrative Neuroscience, 6*, 67.

Henderson, J. M., & Luke, S. G. (2012). Oculomotor inhibition of return in normal and mindless reading. *Psychonomic Bulletin & Review, 19*(6), 1101–1107.

Hilchey, M., Ivanoff, J., Taylor, T., & Klein, R. (2011). The effect of inhibition of return on the temporal dynamics of spatial information processing. *Acta Psychologica, 136*, 235–244.

Hilchey, M. D., Hashish, M., MacLean, G. H., Satel, J., Ivanoff, J., & Klein, R. M. (2014a). On the role of eye movement monitoring and discouragement on inhibition of return in a go/no-go task. *Vision Research, 96*(3), 133–139.

Hilchey, M. D., Klein, R. M., & Ivanoff, J. (2012). Perceptual and motor inhibition of return: components or flavors? *Attention, Perception, & Psychophysics, 74*(7), 1416–1429.

Hilchey, M. D., Klein, R. M., & Satel, J. (2014b). Returning to "inhibition of return" by dissociating long-term oculomotor IOR from short-term sensory adaptation and other nonoculomotor "inhibitory" cueing effects. *Journal of Experimental Psychology: Human Perception and Performance, 40*(4), 1603–1616.

Hilchey, M. D., Rajsic, J., Huffman, G., & Pratt, J. (2017). Response-mediated spatial priming despite perfectly valid target location cues and intervening response events. *Visual Cognition, 25*(9–10), 888–902.

Hilchey, M. D., Leber, A. B., & Pratt, J. (2018). Testing the role of response repetition in spatial priming in visual search. *Attention, Perception & Psychophysics, 80*(6), 1362–1374.

Ho, C., Santangelo, V., & Spence, C. (2009). Multisensory warning signals: when spatial correspondence matters. *Experimental Brain Research, 195*(2), 261–272.

Hommel, B. (1993). The role of attention for the simon effect. *Psychological Research, 55*(3), 208–222.

Hunt, A. R., & Kingstone, A. (2003). Inhibition of return: Dissociating attentional and oculomotor components. *Journal of Experimental Psychology: Human Perception & Performan, 29*(5), 1068–1074.

Innes, B. R., & Otto, T. U. (2019). A comparative analysis of response times shows that multisensory benefits and interactions are not equivalent. *Scientific Reports, 9*(1), 2921, https://doi.org/10.1038/s41598-019-39924-6.

Irwin, D. E. (1998). Lexical processing during saccadic eye movements. *Cognitive Psychology, 36*(1), 1–27.

Jayaraman, S., Klein, R. M., Hilchey, M. D., Patil, G. S., & Mishra, R. K. (2016). Spatial gradients of oculomotor inhibition of return in deaf and normal adults. *Experimental Brain Research, 234*(1), 323–330.

Kavyani, M., Farsi, A., Abdoli, B., & Klein, R. M. (2016). Using the locus-of-slack logic to determine whether inhibition of return in a cue–target paradigm is delaying early or late stages of processing. *Canadian Journal of Experimental Psychology, 71*(1), 63–70.

Klein,, & R, . (1988). Inhibitory tagging system facilitates visual search. *Nature, 334*(6181), 430–431.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences, 4*(4), 138–147.

Klein, R. M., Kavyani, M., Farsi, A., & Lawrence, M. A. (2018). Using the locus of slack logic to determine whether the output form of inhibition of return affects an early or late stage of processing. *Cortex*, *122*, 123–130.

Klein, R., & Taylor, T. L. (1994). Categories of cognitive inhibition with reference to attention. In: D. Dagenbach, & T.H. Carr (Eds.), *Inhibitory Processes in Attention, Memory, and Language* (pp. 113–150). New York, NY: Academic Press.

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychologica, 134*(3), 372–384.

Laurienti, P. J., Burdette, J. H., Maldjian, J. A., & Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiology of Aging, 27*(8), 1155–1163.

Łukasz,, Michalczyk,, Jacek,, & Bielas, . (2019). The gap effect reduces both manual and saccadic inhibition of return (IOR). *Experimental Brain Research, 237*(7), 1643–1653.

Lunn, J., Sjoblom, A., Ward, J., Sotofaraco, S., & Forster, S. (2019). Multisensory enhancement of attention depends on whether you are already paying attention. *Cognition, 187*, 38–49.

Liu, Y., & Otto, T. U. (2020). The role of context in experiments and models of multisensory decision making. *Journal of Mathematical Psychology, 96*, 102352.

MacInnes, W. J. (2017). Multiple Diffusion Models to Compare Saccadic and Manual Responses for Inhibition of Return. *Neural Computation*, *29*(3), 804–824.

Makovac, E., Buonocore, A., & Mcintosh, R. D. (2015). Audio-visual integration and saccadic inhibition. *Quarterly Journal of Experimental Psychology, 68*(7), 1295–1305.

Mcdonald, J. J., Hickey, C., Green, J. J., & Whitman, J. C. (2009). Inhibition of return in the covert deployment of attention: evidence from human electrophysiology. *Journal of Cognitive Neuroscience. 21*(4), 725–733.

Mcdonald, J. J., Ward, L. M., & Kiehl, K. A. (1999). An event-related brain potential study of inhibition of return. *Percept Psychophys, 61*(7), 1411–1423.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology, 14*(2), 247–279.

Miller, J. (1986). Timecourse of coactivation in bimodal divided attention. *Perception & Psychophysics, 40*(5), 331–343.

Miller, J. (2016). Statistical facilitation and the redundant signals effect: What are race and coactivation models?. *Attention, Perception, & Psychophysics, 78*(2), 516–519.

Meredith, M. A., & Stein, B. E. (1986). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research, 365*(2), 350–354.

Mokler, A., & Fischer, B. (1999). The recognition and correction of involuntary prosaccades in an antisaccade task. *Experimental Brain Research, 125*(4), 511–516.

Mondor, T. A., Terrio, N. A., & Hurlburt, J. (2000). On the role of eye movements and saccade preparation in generating auditory inhibition of return. *Canadian Journal of Experimental Psychology, 54*(4), 326–338.

Mozolic, J. L., Hugenschmidt, C. E., Peiffer, A. M., & Laurienti, P. J. (2008). Modality-specific selective attention attenuates multisensory integration. *Experimental Brain Research, 184*(1), 39–52.

Otto, T. U., Dassy, B., & Mamassian, P. (2013). Principles of multisensory behavior. *Journal of Neuroscience, 33*(17), 7463–7474.

Otto, T. U. (2019). RSE-box: An analysis and modelling package to study response times to multiple signals. *The Quantitative Methods for Psychology*, *15*(2), 112–133.

Peng, X., Chang, R., Li, Q., Wang, A., & Tang, X. (2019). Visually induced inhibition of return affects the audiovisual integration under different SOA conditions. *Acta Psychologica Sinica, 51*(7), 759–771.

Posner, M. I., & Cohen, Y. (1984) Components of visual orienting. In: H Bouma, & DG Bouwhuis (eds) *Attention & performance X: control of language processes* (pp. 531–556). Hillsdale, NJ:Psychology Press Erlbaum.

Posner, M. I., Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology, 2*(3), 211–228.

Prime, D. J., & Ward, L. M. (2004). Inhibition of return from stimulus to response. *Psychological Science, 15*(4), 272–276.

Prime, D. J., & Ward, L. M. (2006). Cortical expressions of inhibition of return. *Brain Research, 1072*(1), 161–174.

Rafal, R., Egly, R., & Rhodes, D. (1994). Effects of inhibition of return on voluntary and visually guided saccades. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Experimentale, 48*(2), 284–300.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506.

Redden, R. S., Hilchey, M. D., & Klein, R. M. (2018). Oculomotor inhibition of return: Evidence against object-centered representation. *Visual Cognition*, 1–15.

Reimer, C. B., Tudge, L., & Schubert, T. (2021). Saccadic eye movements in dual tasking: No impairment of spatial planning, but delayed execution of saccades. *Quarterly Journal of Experimental Psychology (2006), 74*(2), 326–343.

Reuter-Lorenz, P. A., & Rosenquist, J. N. (1996). Auditory cues and inhibition of return: the importance of oculomotor activation. *Experimental Brain Research, 112*(1), 119–126.

Ro, T., Pratt, J., & Rafal, R. D. (2000). Inhibition of return in saccadic eye movements. *Experimental Brain Research, 130*(2), 264–268.

Santangelo, V., Ho, C., & Spence, C. (2008). Capturing spatial attention with multisensory cues. *Psychonomic Bulletin & Review, 15*(2), 398–403.

Sapir, A., Soroker, N., Berger, A., & Henik, A. (1999). Inhibition of return in spatial attention: Direct evidence for collicular generation. *Nature Neuroscience, 2*(11),1053–1054.

Satel, J., Hilchey, M. D., Wang, Z., Story, R., & Klein, R. M. (2013). The Effects of ignored versus foveated cues upon inhibition of return : an event-related potential study. *Attention, Perception, & Psychophysics, 75*(1), 29–40.

Senkowski, D., Saint-Amour, D., Höfle, M., & Foxe, J. J. (2011). Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *Neuroimage, 56*(4), 2200–2208.

Simon, R. J., Acosta, E., & Mewaldt, S. P. (1975). Effect of locus of warning tone on auditory choice reaction time. *Memory & Cognition, 3*(2), 167–170.

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. *In Advances in psychology* (Vol. *65*, pp. 31–86). Princeton, NJ: North-Holland Publishing.

Souto, D., & Kerzel, D. (2009). Evidence for an attentional component in saccadic inhibition of return. *Experimental Brain Research, 197*(2), 205–205.

Spence, C. J., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance, 20*(3), 555.

Stevenson, R. A., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., … et al, . (2014). Identifying and Quantifying Multisensory Integration: A Tutorial Review. *Brain Topography, 27*(6), 707–730.

Sumner, P., Nachev, P., Vora, N., Husain, M., & Kennard, C. (2004). Distinct Cortical and Collicular Mechanisms of Inhibition of Return Revealed with S Cone Stimuli. *Current Biology, 14*(24), 2259–2263.

Talsma, D. (2015). Predictive coding and multisensory integration: an attentional account of the multisensory mind. *Frontiers in Integrative Neuroscience, 9*, 19.

Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cerebral Cortex, 17*(3), 679–690.

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences, 14*(9), 400–410.

Talsma, D., & Woldorff, M. G. (2005). Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *Journal of Cognitive Neuroscience, 17*(7), 1098–1114.

Tang, X., Gao, Y., Yang, W., Ren, Y., Wu, J., Zhang, M., . . . Wu, Q. (2019). Bimodal-divided attention attenuates visually induced inhibition of return with audiovisual targets. *Experimental Brain Research, 237*(4), 1093–1107.

Tang, X., Wu, Y., Peng, X., Wang, A., & Li, Q. (2020). The influence of endogenous spatial cue validity on audiovisual integration. *Acta Psychologica Sinica, 52*(7), 835–846.

Tang, X., Wang, X., Peng, X., Li, Q., Zhang, C., Wang, A., . . . Zhang, M. (2021). Electrophysiological evidence of different neural processing between visual and audiovisual inhibition of return. *Scientific Reports, 11*(1), 8056.

Tang, X., Wu, J., & Shen, Y. (2016). The interactions of multisensory integration with endogenous and exogenous attention. *Neuroscience & Biobehavioral Reviews, 61*, 208–224.

Taylor, T. L., & Klein, R. M. (2000). Visual and motor effects in inhibition of return. *Journal of Experimental Psychology: Human Perception & Performan, 26*(5), 1639–1656.

Van der Burg, E., Talsma, D., Olivers, C.N.L., Hickey, C., & Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects. *Neuroimage, 55*(3), 1208–1218.

Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception & Performance, 34*(5), 1053–1065.

Van der Stoep, N., Spence, C., Nijboer, T. C. W., & Van der Stigchel, S. (2015). On the relative contributions of multisensory integration and crossmodal exogenous spatial attention to multisensory response enhancement. *Acta Psychologica, 162*, 20–28.

Van der Stoep, N., Van der Stigchel, S., & Nijboer, T. C. W. (2015). Exogenous spatial attention decreases audiovisual integration. *Attention Perception & Psychophysics, 77*(2), 464–482.

Van der Stoep, N., Van der Stigchel, S., Nijboer, T. C. W., & Spence, C. (2017). Visually Induced Inhibition of Return Affects the Integration of Auditory and Visual Information. *Perception, 46*, 6–17.

Wang, C. A., Blohm, G., Huang, J., Boehnke, S. E., & Munoz, D. P. (2017). Multisensory integration in orienting behavior: Pupil size, microsaccades, and saccades. *Biological Psychology, 129*, 36–44.

Wilson, D. E., Castel, A. D., & Pratt, J. (2006). Long-term inhibition of return for spatial locations: Evidence for a memory retrieval account. *Quarterly Journal of Experimental Psychology, 59*(12), 2135–2147.

Xi, Y., Li, Q., Zhang, M., Liu, L., & Wu, J. (2020). Characterizing the Time-Varying Brain Networks of Audiovisual Integration across Frequency Bands. *Cognitive Computation, 12*(6), 1154–1169.

Xu, Z., Yang, W., Zhou, Z., & Ren, Y. (2020). Cue–target onset asynchrony modulates interaction between exogenous attention and audiovisual integration. *Cognitive Processing*, *21*(2), 261–270.