Visually-induced inhibition of return affects the integration of auditory and visual information

N. Van der Stoep[1], S. Van der Stigchel[1], T. C. W. Nijboer[1, 2], & C. Spence[3]

[1] Utrecht University, Department of Experimental Psychology,

Helmholtz Institute, Utrecht, The Netherlands

[2] Brain Center Rudolf Magnus, and Center of Excellence for Rehabilitation Medicine,

University Medical Center Utrecht and De Hoogstraat Rehabilitation, The Netherlands

[3] Oxford University, Department of Experimental Psychology, Oxford, United Kingdom

Address for correspondence:

Nathan Van der Stoep

Experimental Psychology, Helmholtz Institute

Heidelberglaan 1, 3584 CS, Utrecht, The Netherlands

E-mail: N.VanderStoep@uu.nl

Abstract (201 words)

Multisensory integration (MSI) and exogenous spatial attention can both speed-up responses to perceptual events. Recently, it has been shown that audiovisual integration at exogenously attended locations is reduced relative to unattended locations. This effect was observed at short cue-target intervals (200-250 ms). At longer intervals, however, the initial benefits of exogenous shifts of spatial attention at the cued location are often replaced by response time costs (also known as Inhibition of Return, IOR). Given these opposing cueing effects at shorter versus longer intervals, we decided to investigate whether MSI would also be affected by IOR. Uninformative exogenous visual spatial cues were presented between 350 and 450 ms prior to the onset of auditory, visual, and audiovisual targets. As expected, IOR was observed for visual targets (invalid cue RT < valid cue RT). For auditory and audiovisual targets neither IOR, nor any spatial cueing effects, were observed. The amount of relative multisensory response enhancement and race model inequality violation was larger for uncued as compared to cued locations. It can be concluded that IOR delayed visual information processing, resulting in an increase in the difference between unimodal auditory and visual processing latencies, and a reduction in multisensory integration.

Keywords: multisensory integration, inhibition of return, exogenous, spatial attention, race model.

In everyday life, our senses are often flooded with sensory information. At first glance, it might seem that this abundant information would hinder our perception of the environment. However, generally-speaking, we do not experience everything that we see and hear as separate events. Furthermore, neither do we perceive everything that stimulates our sense organs at each moment. Two mechanisms that help in combining and processing sensory information are multisensory integration (MSI) and crossmodal exogenous spatial attention (Calvert, Spence, & Stein, 2004; Spence & Driver, 2004).

The detection, identification, and localization of multisensory stimuli is often enhanced as a result of MSI (Stein & Meredith, 1993; Stein & Stanford, 2008). Research has highlighted a number of factors that can facilitate MSI. First, MSI is often most pronounced when information from the different senses is presented from approximately the same location (this is known as the spatial rule, Stein & Meredith, 1990; though see Spence, 2013). A second factor that has been shown to facilitate MSI is close temporal proximity (this is known as the temporal rule; e.g., Stein & Meredith, 1990; Stevenson, Fister, Barnett, Nidiffer, & Wallace, 2013). Perfect temporal alignment between the stimuli from different modalities is not essential, as long as sensory inputs arrive within a certain temporal binding window (it has been suggested that the discharge trains for visual and auditory inputs should overlap in multisensory neurons; e.g., King & Palmer, 1985; Meredith, Nemitz, & Stein, 1987). Third, MSI seems to be more pronounced for those stimuli that, when presented individually, are weakly, rather than strongly, effective (Holmes, 2007; Holmes & Spence, 2005; Meredith & Stein, 1983).

Crossmodal exogenous spatial attention has also enhance perception. For example, when attention is automatically attracted to a certain spatial location by the sudden onset of a sound, the perception of visual information that is presented shortly thereafter at the same location is often facilitated as compared to when the same stimulus is presented from an another, unattended, location (e.g., McDonald, Teder-Salejarvi, Di Russo, & Hillyard, 2005; McDonald, Teder-Sälejärvi, & Hillyard, 2000; Störmer, McDonald, & Hillyard, 2009). Furthermore, response times (RTs) to exogenously attended targets are generally faster than to unattended targets (e.g., McDonald, Teder-Sälejärvi, Di Russo, & Hillyard, 2003; Spence & Driver, 2004; Spence, Nicholls, Gillespie, & Driver, 1998; Ward, McDonald, & Lin, 2000). This initial speed-up of responses at cued locations due to exogenous shifts of spatial attention at short cue-target intervals (100-300 ms) often reverses at longer intervals (>300 ms), especially in those tasks requiring simple speeded detection responses. This inhibitory after-effect is often labeled 'Inhibition of Return' (IOR; see Klein, 2000, for a review). IOR is characterized by slower responses to targets appearing at cued as compared to uncued locations. While IOR was first described in the visual modality (i.e., visual cues preceding visual targets; Posner & Cohen, 1984), it has subsequently been documented between all possible combinations of auditory, visual, and tactile stimuli (McDonald & Ward, 1999; Spence, Nicholls, Gillespie, & Driver, 1998; Ward, McDonald, & Lin, 2000; Spence & Driver, 1998a, 1998b; Spence, Lloyd, McGlone, Nicholls, & Driver, 2000). One of the suggestions that has been put forward is that IOR facilitates efficient visual search as it inhibits the exploration of previously attended locations and encourages the exploration of novel locations (Klein, 2000). As for the underlying effects of IOR, there is evidence to suggest that both attentional and motor processes are affected by IOR (e.g., Hilchey, Hashish, MacLean, Satel, Ivanoff, & Klein, 2014; Hilchey, & Klein, 2014; Kingstone & Pratt, 1999; McDonald, Hickey, Green, & Whitman, 2009; Taylor, Klein, & Satel, 2000; Tian & Yao, 2008). Interestingly, crossmodal (auditory-visual) IOR does not seem to delay sensory processing but rather decreases the brain's response to stimuli at cued locations (i.e., decreases the amplitude of the neuronal response to cued stimuli; e.g., Dorris, Klein, Everling, & Munoz, 2002; McDonald, Hickey, Green, & Whitman, 2009).

Given that both MSI and crossmodal exogenous spatial attention are able to enhance perceptual processing, Van der Stoep et al. (2015) recently set out to investigate the interaction between these two mechanisms (see also McDonald, Teder-Sälejärvi, & Ward, 2001). In that study, auditory exogenous spatial cues were presented before the onset of unimodal auditory (A), unimodal visual (V), and audiovisual (AV) targets. The cumulative distributive functions (CDF) RTS in the A and V condition were used to calculate how fast responses in the AV condition would be expected to become due to statistical facilitation (i.e., independent processing as indicated by the race model inequality (RMI), see Miller, 1982, 1986). If responses to AV stimuli are faster than predicted by statistical facilitation then this indicates that an interaction between the senses must have taken place (MSI). As expected, responses to exogenously attended AV targets were faster than to unattended AV targets, but MSI (i.e., the amount of RMI violation) was *reduced* at exogenously attended locations as compared to unattended locations. One explanation offered for this particular pattern of results was that exogenous spatial attention might affect MSI by increasing the perceptual sensitivity at attended locations as compared to unattended locations (effectively similar to an increase in contrast and/or intensity, Ling & Carrasco, 2006). According to the principle of inverse effectiveness, the effects of MSI are most pronounced for weakly effective information as compared to strongly effective information (Meredith & Stein, 1983, though see Holmes, 2007). Consequently, MSI may be reduced at exogenously attended as compared to unattended locations.

Presenting an exogenous spatial cue before a multisensory target at longer cue-target intervals may, however, have a very different effect on MSI once IOR comes into play. IOR is known to delay the processing of, and/or responses to, information at cued locations at longer cue-target intervals, especially in speeded detection tasks. As IOR has most often been observed in the visual modality, we investigated the effects of visual exogenous spatial cues on the processing of A, V, and AV targets using the implicit spatial discrimination paradigm (McDonald & Ward, 1999; Ward, McDonald, & Lin, 2000) and cue-target intervals in the range of 350-450 ms. Our hypothesis was that this stimulus interval would lead to IOR for visual cue/visual target pairs, but not necessarily for visual cue/auditory target pairs given that a pattern of visual-auditory IOR is often only observed at longer cue-target intervals (Ward, McDonald, & Lin, 2000; Spence, Lloyd, McGlone, Nicholls, & Driver, 2000). If at cue-target intervals between 350-450 ms, the amount of exogenous spatial attention at the cued location is reduced relative to uncued locations (McDonald, Hickey, Green, & Whitman, 2009) for all sensory modalities, then MSI might be enhanced at cued locations as compared to uncued locations (cf. Van der Stoep, Van der Stigchel, & Nijboer, 2015). On the other hand, if IOR were, indeed, only to occur in the visual modality, then MSI could decrease at cued locations due to an increase in the difference in auditory and visual processing latencies. In the latter case, visual, but not auditory, processing is inhibited, thus making auditory information more dominant. Previous studies have shown that such sensory dominance can affect the outcome of MSI greatly (Ernst & Banks, 2002; Mozolic, Hugenschmidt, Peiffer, & Laurienti, 2008; Spence, Shore, & Klein, 2001).

The aim of the current study was to see how multisensory target processing would be affected by IOR induced by visual exogenous spatial cues. By calculating the amount of multisensory response enhancement and RMI violation (Raab, 1962; Miller, 1986;

Ulrich, Miller, & Schröter, 2007) for cued and uncued locations, the effect of IOR on MSI could be examined.

Methods

Participants

Twenty-four participants were tested in this experiment (mean age = 26 years, *SD* = 3.4, 11 male, 13 female). All of the participants took part in this study and a study of the relative contribution of MSI and exogenous spatial attention to multisensory response enhancement (Van der Stoep, Spence, Nijboer, & Van der Stigchel, 2015) in a single experimental session that lasted for approximately 1.5 hours. The order in which the studies were conducted was counterbalanced across participants and showed no interaction with any of the factors in the analyses. All of the participants reported a normal sense of hearing and normal or corrected-to-normal vision. The participants signed an informed consent form prior to their taking part in the study and were rewarded for their participation with £10 sterling. The study was reviewed and approved by the Central University Research Ethics Committee of the University of Oxford.

Apparatus

A custom built stimulus generator was used to present the auditory and visual stimuli with an accuracy of 1 ms in terms of their onset and offset. Auditory stimuli were presented via loudspeakers (e-audio black 4” Full Range Mini Box Speaker, dimensions: 120x120x132 mm, frequency response: 80-20,000 Hz) and the visual stimuli consisted of the illumination of a Light Emitting Diode (LED; Forge Europa, bulb size: 5 mm, viewing angle: 65°, tri-coloured LED) that was positioned at the center of each loudspeaker. The loudspeaker and LED array was placed at eye-level. Two loudspeakers and LEDs were positioned 26.1° to the left and right of the central loudspeaker and LED, which was placed in front of the participant at a distance of 64 cm. The auditory targets consisted of a white noise burst (100 ms, 15 ms rise and fall of the signal, ~65dB(A)). Each LED could emit red, green, and blue light that was used to differentiate between the fixation (blue, 14.76 cd/m^2), cue (red, 172.8 cd/m^2), and target light (green, 130.2 cd/m^2). The audiovisual target consisted of a combination of the A and V target. The participants were instructed to respond to the targets using a custom response device connected to the audiovisual stimulus generator to allow the precise recording of response times.

Stimuli, Task, and Procedure

At the start of the experiment, the participants were seated on a chair in front of the loudspeaker array in a dark soundproofed room. The LEDs that were illuminated over the course of a trial were the only source of illumination in the experimental room. A verbal instruction was presented from the central loudspeaker while the central LED emitted blue light. The participants were asked whether they had understood the instructions. Next there was a practice session that contained one trial of each condition in the experiment (presented in a random order). The experimenter stayed in the room with each participant during the practice trials and left when the actual experiment started after making sure that the participants were correctly performing the task.

On each trial, the blue central fixation LED was lit up for a random duration of between 750 and 1250 ms. At the offset of the fixation light, the visual spatial cue was presented (red LED) for 100 ms from one of the three locations (left, center, or right). After a random cue-target onset asynchrony (CTOA) of between 350-450 ms, an A, V, or AV target was presented from one of the three locations. The participants were instructed to press a single response key whenever a sound, a green light, or the combined (audiovisual) target was presented to the right or left of center (Go trials) and to withhold their response when a target stimulus (A, V, or AV) appeared at the central location (No-go trial). The unimodal components of the AV target were always presented spatially and temporally aligned. The response window was set to 2000 ms after target onset, after which time the next trial started automatically.

The cues and targets could be presented from either the left, central, or right location, and cues could be presented from either the same or different locations as the target. There were three Cue Types for Go trials: Valid (i.e., same lateral location), Invalid (i.e., opposite lateral locations), and Central Cue (i.e., invalid, but the cue was presented from the center). On the No-go trials, the cues could also be presented from the left, center, and right, but here center cues were considered as valid, and the left and right cues were considered as invalid. The experiment consisted of 540 trials: 360 Go trials, and 180 No-go trials (33%). There were 120 Valid Cue Go trials, 120 Invalid Cue Go trials, and 120 Center Cue Go trials. Of the 180 No-go trials, 60 were Valid Cue No-go trials and 120 were Invalid Cue No-go trials (left and right cue with a target presented in the center). All conditions contained an equal number of A, V, and AV target trials.

Data Preprocessing

Response times shorter than 100 ms and those greater than 1000 ms were removed from further analysis because they were assumed to be the result of anticipation or not paying attention to the task, respectively. Only Go trials with a correct response were used in the RT analysis. The median RT of each participant in each condition was used in the analysis of the RT data. For the analysis of the accuracy data, both the accuracy on Go and No-go trials were calculated. Three participants were removed from further analysis because their accuracy was equal to, or fell below, 50% in at least one of the conditions. In total, 2.7% of the data was removed: 1.4% of the Go trials (on average 5 Go trials per participant) and 5.2% of the No-go trials (on average 9 No-go trials per participant) were discarded.

To investigate the amount of speed-up in the multisensory condition compared to the unimodal condition, the relative amount of Multisensory Response Enhancement (rMRE) was calculated for each participant and each condition using the following formula:

$$rMRE = \frac{\min(median(RT_A), median(RT_V)) - \ median(RT_{AV})}{\min(median(RT_A), median(RT_V))} \times \ 100\%$$

To investigate whether a speed-up in the multisensory condition could be explained by statistical facilitation or by MSI, the AV cumulative distributive function (CDF) of RTs was compared with the sum of the unimodal (A and V) CDFs for each Cue Type at the 10th, 20th, 30th up to the 90th percentile (Raab, 1962; Miller, 1986; Ulrich, Miller, & Schröter, 2007). To do so, the RMI was tested:

$$P(RT_{AV} < t) \leq \ P(RT_A < t) + \ P(RT_V < t)$$

Our main theoretical interest was in the effects of IOR as reflected in the difference in RT, MRE, and RMI violation between Validly and Invalidly cued targets. The Center Cue condition was included to ensure a balanced design and make the cues uninformative of the target (left and right) location. The interpretation of the results of the Center Cue condition in terms of cueing effects is difficult, however, as the participants had to withhold their response to targets presented at this central location. Therefore, we did not go into detail on the results of the center cue condition, but we do present the results of this condition in the figures for the purpose of completeness. Adding the Center Cue condition to the analyses of RTs, rMRE, and RMI did not change the outcome of the current study.

Results

Accuracy

The participants were generally very accurate in terms of their responding with an average hit rate of .99 (*SE* = .005) on Go trials and an average correct rejection rate of .95 (*SE* = .007) on No-go trials. We therefore decided to not further analyze the accuracy data.

Response times

A 2 x 3 repeated measures ANOVA was used to analyze the RT data on the Go trials with the factors Cue Type (Valid, Invalid) and Target Modality (Auditory, Visual, Audiovisual). There was a main effect of Target Modality [$F(1.5, 29.993) = 67.478$, $p < .001$, $\varepsilon = .750$, $\eta p^2 = .771$]. Responses to audiovisual targets (*M* = 390 ms, *SE* = 17) were significantly faster than to either visual (*M* = 466 ms, *SE* = 18, $t(20) = 18.171$, $p < .001$, d = .904) or auditory targets (*M* = 427 ms, *SE* = 17, $t(20) = 7.186$, $p < .001$, d = .476), as expected. Furthermore, the responses to auditory targets were faster than to visual targets [$t(20) = -5.924$, $p < .001$, d = -.466].

There was no main effect of Cue Type [$F(1, 20) = .151$, $p = .702$, $\eta p^2 = .008$]. The average of median RTs for A, V, and AV targets in the valid, invalid, and center cue condition are shown in Figure 1A.

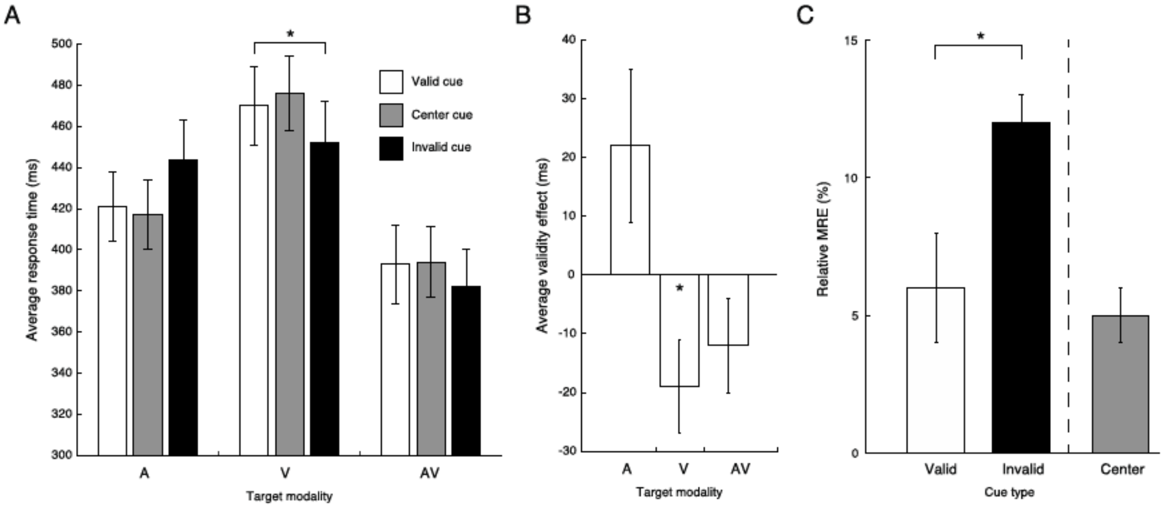

Figure 1. A: The average of median RTs for each Target Modality and Cue Type. Only significant validity effects are indicated with an asterisk ($p < .05$). B: The size and direction of the validity effect for each Target Modality. Positive values indicate that responses to validly cued targets were faster than to invalidly cued targets, whereas negative values indicate the opposite. Error bars represent the standard error of the mean. The asterisk inside the bar indicates a significant difference from zero. C: The average amount of rMRE for each Cue Type. Significant differences are indicated with an asterisk ($p < .05$).

There was a significant interaction between Cue Type and Target Modality [$F(1.460, 29.201) = 6.030$, $p = .001$, $\varepsilon = .730$, $\eta p^2 = .232$], with the cues exerting a different effect on performance as a function of the modality of the target. In particular, responses to validly cued *visual* targets (M = 470 ms, SE = 19) were significantly slower than to invalidly cued *visual* targets (*M* = 452 ms, *SE* = 20, *t*(20) = 2.394, *p* = .027, d = .204) indicating an inhibitory aftereffect of the cue (often labeled IOR; see Figure 1A and 1B). There was no difference in RTs between validly (*M* = 421 ms, *SE* = 17) and invalidly cued *auditory* targets (*M* = 444 ms, *SE* = 19, *t*(20) = -1.743, *p* = .097), but the pattern of RTs was in the direction of there being a spatial cuing effect (valid RTs < invalid RTs, mean difference = 22 ms, *SE* = 13).

No differences in RTs were observed between validly (*M* = 393 ms, *SE* = 19) and invalidly cued (*M* = 382 ms, *SE* = 18) *audiovisual* targets [t(20) = 1.428, p = .169]. In sum, IOR was observed for visual targets, but not for auditory or audiovisual targets.

Multisensory response enhancement

The amount of rMRE was significantly different from zero for all Cue Types (*t*'s > 2.3, *p*'s < .05). Figure 1C depicts the average rMRE for the different cue types. Difference in rMRE between the Valid and Invalid cue condition was analyzed using a paired samples t-test. The amount of rMRE was significantly larger in the Invalid (*M* = 11%, *SE* = 1) as compared to the Valid Cue condition (*M* = 6%, *SE* = 2, t(20) = -2.145, *p* = .044, d = -0.649). These results indicate that the amount of speed-up attributable to multisensory stimulation was significantly larger for uncued as compared to cued targets.

Race model inequality violation

In order to investigate whether the speed-up in the multisensory conditions could be explained by an independent processing model, or by co-activation (i.e., MSI), violations of the RMI were analyzed for each Cue Type. Figure 2A shows the average amount of race model equality violation for each percentile bin for all Cue Types.

One-tailed one sample t-tests on the difference between the audiovisual and the race model CDF were performed at each of the 9 percentiles for each Cue Type (not all percentiles are shown in Figure 2A because only positive RMI violations are meaningful). Significant violations of the RMI were observed in the Invalid Cue condition for the 10th to 30th percentiles (*t*'s > 3, *p*'s < .05, corrected for multiple comparisons using the Bonferroni correction method). The RMI was not violated in the Valid Cue condition (*t*'s < 1.8, *p*'s > .4).

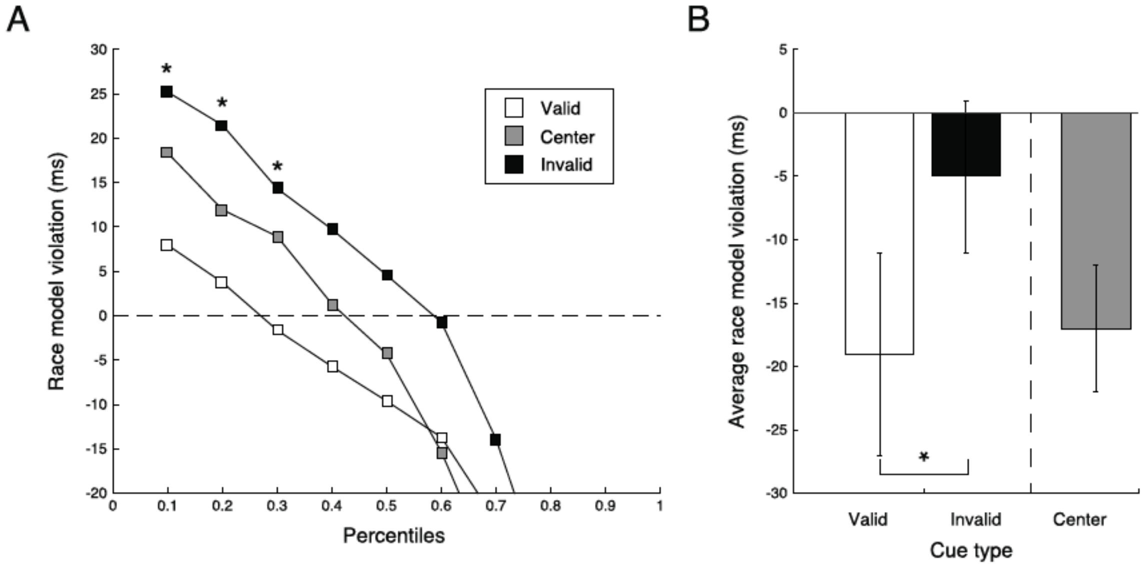

Figure 2. A: RMI violations in the Valid (white), Invalid (black), and Center Cue condition (grey). Significant positive violations of the RMI in the Valid and Invalid cue condition are indicated with an asterisk ($p < .05$). B: The average RMI violation across all percentiles for each Cue Type. Significant differences in RMI violation are indicated with an asterisk ($p < .05$).

The one-sample t-tests for each quantile and each Cue Type allows drawing conclusions about whether MRE could be explained by statistical facilitation or MSI. To further explore and quantify differences in RMI violation, a repeated measures ANOVA with the factors Cue Type (Valid, Invalid) and Percentile (10th to 90th) was conducted. There was a main effect of Cue Type [$F(1, 20) = 7.553$, $p = .012$, ηp^2 = .274] indicating that the average amount of violation was significantly larger in the Invalid Cue condition (M = -5, SE = 6) than in the Valid Cue condition (M = -19, SE = 8, see Figure 2B). The average amount of RMI violation was negative in both the Valid and Invalid cue conditions. Therefore, the main effect only indicates that there is a difference between the two conditions. The one-sample t-tests described above indicate that the RMI is only significantly violated in the Invalid Cue condition. There was also a significant main effect of Percentile [$F(1.285, 25.703) = 36.436$, $p < .001$, $\varepsilon = .161$, ηp^2 = .646], indicating that the amount of violation varied across percentiles, which can be clearly seen in Figure 2A. This pattern of results is often observed when testing for RMI violations and is nothing unusual (Ulrich, Miller, & Schröter, 2007). The interaction between Cue Type and Percentile was not significant [$F(1.873, 37.470) = .218$, $p = .791$, $\varepsilon = .234$, ηp^2 = .011].

Correlation between differences in unimodal response times, rMRE, and RMI violation

As can be seen from Figure 1A, the difference in the average of median RTs between A and V targets was larger for cued (Valid) than for uncued targets (Invalid). It has previously been suggested that equal performance in different sensory modalities is important for MSI and often leads to the largest benefits of multisensory stimulation (e.g., the fastest or most accurate response; Ernst & Banks, 2002; Leone & McCourt, 2013; Otto & Mamassian, 2013). Although the difference between A and V RTs between the Valid (M difference = 58 ms, SE = 8) and Invalid Cue condition (M difference = 38 ms, SE = 6) was only marginally significant [$t(20) = 2.009$, $p = .058$], we wanted post-hoc to test whether the amount of rMRE and race model violation was related to the size of the differences in unimodal RTs. To do so, we correlated the difference in A and V RTs with the amount of rMRE and the average amount of RMI violation across participants. When the differences between unimodal RTs were grouped across the two cue types (Valid, Invalid), there was a significant correlation between the difference in unimodal RTs and the amount of rMRE ($r = -.676$, $p < .001$). In addition, the difference in unimodal RTs was also significantly related to the average amount of RMI violation ($r = -.530$, $p = .001$). Overall, these correlations indicate that when the difference between unimodal RTs was smaller the amount of rMRE and RMI violation was larger.

These observations are in line with the idea that equal performance for different sensory modalities leads to larger benefits of MSI (in this case faster RTs) as compared to unequal performance. The results reported here could certainly be taken to suggest that IOR modulates MSI by changing differences in unimodal processing as it differentially affects processing in the auditory and visual sensory modality at the cue-target intervals that we used here.

Discussion

The present study investigated whether IOR affects the integration of audiovisual information. Visual exogenous spatial cues evoked IOR for V, but not for A and AV targets. Although the lack of IOR for audiovisual targets could be taken to suggest that IOR does not affect multisensory processing, further analysis of multisensory response enhancement suggests otherwise. Importantly, we observed that the amount of relative multisensory response enhancement was decreased at cued as compared to uncued locations. Furthermore, the RMI violation analysis indicated that multisensory response enhancement was due to MSI at uncued locations, but could be explained by statistical facilitation at cued locations. Based on these findings, we conclude that MSI was reduced by IOR as it inhibited visual information processing at cued as compared to uncued locations.

Our findings seem to conflict with the results of a previous study of the effects of exogenous spatial attention on audiovisual integration (Van der Stoep, Van der Stigchel, & Nijboer, 2015). In that study, MSI was also reduced for cued (exogenously attended) as compared to uncued (exogenously unattended) locations. Although the effects of exogenous spatial cueing on MSI are similar for short (facilitation, see Van der Stoep, Nijboer, & Van der Stigchel, 2015) and long CTOAs (the current study), we argue that different processes underlie the decreases in MSI at short and longer CTOAs.

At those (short) CTOAs at which both A and V processing is facilitated at the cued location, multisensory integration is likely reduced because exogenous spatial attention enhances contrast sensitivity at that location (e.g., Carrasco, 2011). Given that the individual signals are enhanced due to attention, MSI could be reduced at cued relative to uncued locations in line with the principle of inverse effectiveness.

At longer CTOAs, when IOR occurs for V stimuli, while A stimuli are still facilitated at the cued location, differences in unisensory processing times may reflect differences in signal strength causing one sense to become dominant. Consequently, the most dominant sense drives the response which reduces integration at cued relative to uncued locations. This idea is consistent with the observation that the benefits of multisensory stimulation are most pronounced when performance in the different modalities is similar as compared to when one modality is dominant (e.g., in terms of estimation accuracy or processing speed; see, for example, Ernst & Banks, 2002; Otto & Mamassian, 2013). This makes sense in that MSI is most beneficial (i.e., the largest increase of multisensory response enhancement) when none of the senses dominate perception. Otherwise the dominant sense drives the response and the contribution of the non-dominant sense is decreased. Interestingly, multisensory response enhancement was still observed for cued locations thus suggesting that multisensory stimulation is still helpful in speeding-up responses to targets at the cued location.

When speculating about the possible neuronal mechanisms underlying the effect of IOR on MSI at this cue-target interval, the concept of the temporal binding window (TBW) may come to mind. The TBW is a temporal interval within which auditory and visual information is integrated in multisensory neurons (e.g., Meredith, Nemitz, & Stein, 1983; Stein & Stanford, 2008). This TBW can be applied to stimulus onset times and the onset of neuronal responses. Both physical and physiological simultaneity seem to be important in determining when multisensory integration is most pronounced (Leone & McCourt, 2013). Generally-speaking, MSI decreases as a function of differences in stimulus onset (presentation) times and (resulting) differences in physiological arrival times (e.g., Leone & McCourt, 2013; Meredith, Nemitz, & Stein, 1987). One could argue that the observed difference in unisensory RTs reflects differences in sensory processing times. This differences would, however, not be large enough to push the visual stimuli outside the TBW which is often considered to be approximately ±100 ms. Nonetheless, small differences in unisensory physiological arrival times may still decrease the response of multisensory neurons (Meredith, Nemitz, & Stein, 1987).

This seems an unlikely scenario though, given that IOR seems to mainly affect the amplitude, not the timing of the response of neurons in the superior colliculus, a structure involved in spatial orienting of attention and multisensory integration (Dorris, Klein, Everling, & Munoz, 2002; Stein & Stanford, 2008). Therefore, we think it is more likely that the difference in unisensory RTs reflects a difference in signal strength rather than differences in sensory processing times. Further support for this notion comes from EEG studies of IOR in humans (e.g., McDonald, Hickey, Green, & Whitman, 2009; Prime & Ward, *2006; Tao & You, 2008). IOR seems to mainly affect the amplitude and not necessarily the latency of perceptual processing of target stimuli. If visual input to a multisensory neuron is reduced due to IOR and auditory input enhanced due to attentional facilitation, then this may consequently reduce the outcome of integration of these inputs in multisensory neurons.*

To conclude, the results of the present study clearly indicate that IOR (elicited by the presentation of an exogenous visual cue) can decrease multisensory integration at cued relative to uncued locations. This modulation of multisensory integration as a result of IOR is likely driven by an increase in the difference in unisensory (A vs. V) signal strength at cued as compared to uncued locations.

References

Calvert, G., Spence, C., & Stein, B. E. (Eds.) (2004). *The handbook of multisensory processes*. Cambridge, MA: MIT press.

Dorris, M. C., Klein, R. M., Everling, S., & Munoz, D. P. (2002). Contribution of the primate superior colliculus to inhibition of return. *Journal of Cognitive Neuroscience, 14*(8), 1256-1263.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429-433.

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

Hilchey, M. D., Klein, R. M., & Satel, J. (2014). 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*, 1603-1616.

Holmes, N. P. (2007). The law of inverse effectiveness in neurons and behaviour: Multisensory integration versus normal variability. *Neuropsychologia*, *45*, 3340-3345

Holmes, N. P., & Spence, C. (2005). Multisensory integration: Space, time, and superadditivity. *Current Biology*, **15**, R762-R764.

King, A. J., & Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research*, *60*, 492-500.

Kingstone, A., & Pratt, J. (1999). Inhibition of return is composed of attentional and oculomotor processes. *Perception & Psychophysics*, *61*, 1046-1054.

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

Leone, L. M., & McCourt, M. E. (2013). The roles of physical and physiological simultaneity in audiovisual multisensory facilitation. *i-Perception*, *4*, 213-228.

Ling, S., & Carrasco, M. (2006). Sustained and transient covert attention enhance the signal via different contrast response functions. *Vision Research*, *46*, 1210-1220.

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., Teder-Sälejärvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature*, *407*(6806), 906-908.

McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., & Hillyard, S. A. (2005). Neural basis of auditory-induced shifts in visual time-order perception. *Nature Neuroscience*, *8*(9), 1197-1202.

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *The Journal of Neuroscience*, *7*, 3215-3229.

Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science, 221*, 389-391.

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

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

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

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance: Control of language processes* (Vol. 10, pp. 531-556). Hillsdale, NJ: Erlbaum.

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

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, *24*, 574-590.

Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Annals of the New York Academy of Sciences*, *1296*, 31-49.

Spence, C., & Driver, J. (1998a). Auditory and audiovisual inhibition of return. *Perception & Psychophysics,* 60, 125-139.

Spence, C., & Driver, J. (1998b). Inhibition of return following an auditory cue: The role of central reorienting events. *Experimental Brain Research*, 118, 352-360.

Spence, C., & Driver, J. (Eds.). (2004). *Crossmodal space and crossmodal attention*. Oxford, UK: Oxford University Press.

Spence, C., Lloyd, D., McGlone, F., Nicholls, M. E. R., & Driver, J. (2000). Inhibition of return is supramodal: A demonstration between all possible pairings of vision, touch and audition. *Experimental Brain Research*, *134*, 42-48.

Spence, C., Nicholls, M. E., Gillespie, N., & Driver, J. (1998). Cross-modal links in exogenous covert spatial orienting between touch, audition, and vision. *Perception & Psychophysics*, *60*, 544-557.

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, *130*, 799-832.

Stein, B.E., & Meredith, M. (1990). Multisensory integration. *Annals of the New York Academy of Sciences*, *608*, 51-70.

Stein, B. E., & Meredith, M. A. (1993). The merging of the senses. Cambridge, MA: MIT Press.

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, *9*, 255-266.

Stevenson, R. A., Fister, J. K., Barnett, Z. P., Nidiffer, A. R., & Wallace, M. T. (2012). Interactions between the spatial and temporal stimulus factors that influence multisensory integration in human performance. *Experimental Brain Research*, *219*, 121-137.

Störmer, V. S., McDonald, J. J., & Hillyard, S. A. (2009). Cross-modal cueing of attention alters appearance and early cortical processing of visual stimuli. *Proceedings of the National Academy of Sciences of the USA*, *106*(52), 22456-22461.

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

Tian, Y., & Yao, D. (2008). A study on the neural mechanism of inhibition of return by the event-related potential in the Go/Nogo task. *Biological Psychology*, *79*(2), 171-178.

Ulrich, R., Miller, J., & Schröter, H. (2007). Testing the race model inequality: An algorithm and computer programs. *Behavior Research Methods*, *39*, 291-302.

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

Van der Stoep, N., Spence, C., Nijboer, T. C. W., & Van der Stigchel, S. (in prep.). On the relative contributions of multisensory integration and crossmodal exogenous spatial attention to multisensory response enhancement.

Ward, L. M., McDonald, J. J., & Lin, D. (2000). On asymmetries in cross-modal spatial attention orienting. *Perception & Psychophysics*, *62*, 1258-1264.