Acta Psychologica 162 (2015) 20–28

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

On the relative contributions of multisensory integration and crossmodal exogenous spatial attention to multisensory response enhancement

N. Van der Stoep [a,*], C. Spence [b], T.C.W. Nijboer [a,c], S. Van der Stigchel [a]

[a] Utrecht University, Department of Experimental Psychology, Helmholtz Institute, Utrecht, The Netherlands
[b] Oxford University, Department of Experimental Psychology, Oxford, United Kingdom
[c] Brain Center Rudolf Magnus, and Center of Excellence for Rehabilitation Medicine, University Medical Center Utrecht and De Hoogstraat Rehabilitation, The Netherlands

ARTICLE INFO

Article history:
Received 13 May 2015
Received in revised form 22 September 2015
Accepted 24 September 2015
Available online xxxx

Keywords:
Multisensory response enhancement
Multisensory integration
Exogenous spatial attention
Race model
Cueing
Crossmodal

ABSTRACT

Two processes that can give rise to multisensory response enhancement (MRE) are multisensory integration (MSI) and crossmodal exogenous spatial attention. It is, however, currently unclear what the relative contribution of each of these is to MRE. We investigated this issue using two tasks that are generally assumed to measure MSI (a redundant target effect task) and crossmodal exogenous spatial attention (a spatial cueing task). One block of trials consisted of unimodal auditory and visual targets designed to provide a unimodal baseline. In two other blocks of trials, the participants were presented with spatially and temporally aligned and misaligned audiovisual (AV) targets (0, 50, 100, and 200 ms SOA). In the integration block, the participants were instructed to respond to the onset of the first target stimulus that they detected (A or V). The instruction for the cueing block was to respond only to the onset of the visual targets. The targets could appear at one of three locations: left, center, and right. The participants were instructed to respond only to lateral targets. The results indicated that MRE was caused by MSI at 0 ms SOA. At 50 ms SOA, both crossmodal exogenous spatial attention and MSI contributed to the observed MRE, whereas the MRE observed at the 100 and 200 ms SOAs was attributable to crossmodal exogenous spatial attention, alerting, and temporal preparation. These results therefore suggest that there may be a temporal window in which both MSI and exogenous crossmodal spatial attention can contribute to multisensory response enhancement.

1. Introduction

It is now commonly acknowledged that our senses do not operate independently and that what is perceived via one sense will often (for better or for worse) influence what is perceived via another. For example, when a sound attracts attention to the perceived location of its source, it can facilitate the processing of any visual information that happens to be presented from that location as compared to other locations (i.e., crossmodal exogenous spatial attention; e.g., Driver & Spence, 1998; McDonald, Teder-Sälejärvi, Di Russo, & Hillyard, 2003, 2005; McDonald & Ward, 2000; Spence & Driver, 2004). It is often suggested that what we hear can also be integrated with what we see (i.e., multisensory integration (MSI); e.g., Meredith & Stein, 1986; Molholm et al., 2006; Stein & Meredith, 1993; Stein & Stanford, 2008), often resulting in improved sensory information processing (e.g., Laurienti, Burdette, Maldjian, & Wallace, 2006; Leone & McCourt, 2013; Miller, 1982; Stevenson, Fister, Barnett, Nidiffer, & Wallace, 2012; Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008). Thus, both crossmodal exogenous spatial attention and multisensory integration can facilitate sensory information processing. It is, however, currently unclear what the relative contributions of crossmodal shifts of exogenous spatial attention and multisensory integration are to multisensory response enhancement (MRE; i.e., shorter RTs to multisensory stimuli as compared to the shortest RT to either of the unimodal component stimuli).

What both the crossmodal exogenous spatial attention and the multisensory integration accounts have in common is the suggestion that the benefits of multisensory stimulation are most pronounced when the unimodal components of a multisensory stimulus are spatially aligned (i.e., presented from the same spatial location, the spatial rule; McDonald, Teder-Sälejärvi, & Hillyard, 2000; Spence & Driver, 2004; Spence & McDonald, 2004; Störmer, McDonald, & Hillyard, 2009; though see Spence, 2013) as compared to when they are spatially misaligned (that is, presented from different spatial positions). The principle of spatial alignment seems to hold true for both the horizontal and depth dimension in the case of both crossmodal exogenous spatial attention (e.g., Ngo & Spence, 2010; Van der Stoep, Nijboer, & Van der Stigchel, 2014) and multisensory integration (e.g., Canzoneri, Magosso, & Serino, 2012; Sambo & Forster, 2009; for a review, see Van der Stoep, Nijboer, Van der Stigchel, & Spence, 2015).

It is not surprising to find that there is a debate, here, about whether these processes are essentially the same or not (see McDonald, Teder-

* Corresponding author at: Experimental Psychology, Helmholtz Institute, Heidelberglaan 1, 3584 CS, Utrecht, The Netherlands.
E-mail address: N.VanderStoep@uu.nl (N. Van der Stoep).

http://dx.doi.org/10.1016/j.actpsy.2015.09.010
0001-6918/

Sälejärvi, & Ward, 2001; and Spence, 2010, pp. 183–184), given the similarities between crossmodal exogenous spatial attention and multisensory integration. Interestingly, several ways of differentiating between them have already been proposed in a technical commentary by McDonald et al. (2001). Despite this commentary, studies directly investigating the difference between crossmodal spatial attention and MSI are currently still lacking. McDonald et al. (2001) argued that one of the ways in which to differentiate between them is by looking at the time-course of their behavioral effects. In terms of the temporal alignment/misalignment of sound and light, crossmodal exogenous spatial attention and multisensory integration show very different temporal profiles behaviorally. The beneficial effects of crossmodal exogenous shifts of attention are often most pronounced when there is an interval between the presentation of the auditory and the visual stimulus (i.e., at stimulus onset asynchronies (SOA) of between ~50 and ~300 ms; e.g., Berger, Henik, & Rafal, 2005; McDonald & Ward, 1999, 2000; Spence & Driver, 1997; Spence & McDonald, 2004). In contrast, the behavioral benefits of multisensory integration are often most pronounced when the auditory and visual stimuli are presented in close temporal alignment (SOAs between 0 and ±50 ms; e.g., Leone & McCourt, 2013; Stevenson et al., 2012; though see, for example, King & Palmer, 1985) with the behavioral benefits decreasing more or less symmetrically as the SOA increases (e.g., Leone & McCourt, 2013; though see Vroomen & Keetels, 2010). Thus, time is needed for crossmodal exogenous spatial attention to shift to the location of the cue in order to facilitate the processing of the target, whereas there is a more specific (narrow) time window within which auditory and visual stimuli need to be presented for multisensory integration to occur.[1] The differing temporal profiles of MSI and crossmodal exogenous spatial attention provide support for the notion that multisensory integration and crossmodal exogenous spatial attention are fundamentally different processes.

Further support for this distinction comes from those studies that have indicated that multisensory integration can occur pre-attentively: as multisensory integration can occur before attention has had its effect, this indicates that they are indeed two separate processes (e.g., Soto-Faraco, Navarra, & Alsius, 2004; Spence & Driver, 2000; Vroomen, Bertelson, & De Gelder, 2001a; see McDonald et al., 2001). Furthermore, it has also been shown recently that exogenous crossmodal spatial attention modulates multisensory integration (Van der Stoep, Van der Stigchel, & Nijboer, 2015). When an exogenous spatial auditory cue was presented some time before (SOA: 200–250 ms) and at the same location as a multisensory target, multisensory integration was reduced as compared to when the cue was presented from a different location. This result indicates that exogenous spatial attention can act independently of multisensory integration when there is enough time for exogenous spatial attention to shift to the location of the cue (cf. Vroomen, Bertelson, & de Gelder, 2001b). Lastly, integrated auditory and visual cues can attract spatial attention to their location even under conditions of high perceptual load, whereas unimodal exogenous cues do not (see Spence, 2010, and Spence & Santangelo, 2009, for reviews). Taken together, behavioral effects of multisensory integration and crossmodal exogenous spatial attention not only have different temporal profiles, but also can act independently of and modulate each other.

Generally, two different types of tasks are used to measure the effect of multisensory integration and crossmodal exogenous spatial attention: the redundant target effect (RTE) task and crossmodal spatial cueing tasks (e.g., the orthogonal spatial cueing task, Driver & Spence, 1998; or the implicit spatial discrimination task, McDonald & Ward, 1999; Ward, McDonald, & Lin, 2000). These two paradigms are sometimes referred to as the crossmodal signals paradigm (RTE task) and the focused attention paradigm (crossmodal spatial cueing task; e.g., Colonius & Diederich, 2012). In previous studies, the effects of multisensory stimulation on (saccadic) response times (RTs) for different SOAs in the spatial cueing and RTE paradigms have been modeled within the Time Window of Integration (TWIN) framework (Colonius & Diederich, 2004, 2012; Diederich & Colonius, 2008, 2011). The TWIN model predicts the pattern of multisensory response enhancement (MRE) for a broad range of different SOAs for both paradigms.

Although insights from the TWIN model are specifically helpful in thinking about the optimal time window of multisensory integration under different conditions, it does not provide information about the relative contributions of various crossmodal processes that might contribute to MRE (i.e., crossmodal exogenous spatial attention and multisensory integration). The aim of the present study was therefore to systematically investigate the relative contribution of multisensory integration and crossmodal exogenous spatial attention to MRE at different temporal intervals between an auditory and a visual stimulus (at SOAs of 0, 50, 100, and 200 ms, auditory lead). To do this, two tasks were used that are generally considered to measure the effects of either crossmodal exogenous spatial attention (the implicit spatial discrimination task, e.g., McDonald & Ward, 1999) or multisensory integration (an RTE task; e.g., Laurienti et al., 2006; Miller, 1986; Stevenson et al., 2012; Van der Stoep, Van der Stigchel, & Nijboer, 2015) using exactly the same auditory and visual stimuli. By comparing the results from the two tasks, it was possible to explore the stimulus intervals at which MRE was caused by multisensory integration, exogenous crossmodal spatial attention, or both processes. Based on the above-mentioned literature, it was hypothesized that MRE is caused by multisensory integration at the shortest SOAs (0 ms), by crossmodal exogenous spatial attention and multisensory integration at intermediate SOAs (50 ms), and by crossmodal exogenous spatial attention at longer SOAs (100–200 ms).

2. Materials and methods

2.1. Participants

Twenty-four participants were tested in this experiment (mean age = 26.6 years, *SD* = 3.3, 10 male, 14 female). All of the participants reported a normal sense of hearing and normal or corrected-to-normal visual acuity. They all signed an informed consent form prior to taking part in the study and were rewarded with £10 sterling for their participation. All of the participants took part in the current study and another study of multisensory interactions in one session that lasted for approximately 1.5 h. The order in which the experiments were conducted was counterbalanced across participants. The study was reviewed and approved by the Central University Research Ethics Committee of the University of Oxford.

2.2. Apparatus and stimuli

A custom-built audiovisual stimulus generator (see also Van der Stoep, Van der Stigchel, & Nijboer, 2015) connected to a PC running MATLAB was used to present the auditory stimuli through different loudspeakers (e-audio black 4″ Full Range Mini Box Speaker, dimensions: 120 × 120 × 132 mm, frequency response: 80–20,000 Hz) and visual stimuli through different Light Emitting Diodes (LEDs, Forge Europa, bulb size: 5 mm, viewing angle: 65°, tri-colored LED color: red, green, and blue). The loudspeaker array consisted of three loudspeakers placed at eye-level. One loudspeaker was positioned directly in front of the participant at eye-level at a distance of 64 cm, and two loudspeakers were positioned 26.1° to the left and right of the central loudspeaker. The auditory targets consisted of a 100 ms white noise burst [15 ms rise and fall time, ~65 dB(A)]. The use of tri-colored LEDs

[1] For multisensory integration to occur, it seems especially important that the responses to auditory and visual stimuli in a multisensory neuron overlap (e.g., King & Palmer, 1985; Meredith et al., 1987). Relatively small differences in temporal onset can occur while still resulting in multisensory integration. The overlap in auditory and visual discharge trains is dependent on stimulus intensity, the distance between the stimuli and the observer, and the time it takes for visual and auditory input to reach a multisensory neuron.

enabled the presentation of red, green, and blue stimuli from the same LED. Each LED was mounted on the center of a loudspeaker. The fixation point was a blue light presented for a random duration of between 750–1250 ms (14.76 cd/m^2) and the visual target was a green light that was presented for 100 ms (130.2 cd/m^2; the intensities were measured at a distance of 5 cm from the LED with the lights turned off).

2.3. Task and procedure

The main experiment consisted of three blocks of trials: a unimodal block, which acted as a baseline measure, an Integration block, and a Cueing block. The order in which the blocks were presented was counterbalanced across participants.

The participants were seated in front of the loudspeaker array in a dark soundproofed room and were instructed to place their chin in a chin-rest to ensure a distance of 64 cm between the participants and the stimuli. The participants were instructed to respond as rapidly and accurately as possible to targets appearing on the left and right of the central location by pressing a button with the index finger of their dominant hand (Go trials), and to withhold their response when a target appeared at the central location (No-go trials). At the start of each block of trials, the participants received a verbal instruction delivered from the central loudspeaker prior to the start of the task while the central LED was blue. The instruction varied between blocks. In the unimodal and Integration blocks, the participants were instructed to respond to both auditory and visual targets, but in the Cueing block they were instructed to respond only to the visual targets. After the instruction, there was a short practice session containing one trial of each condition in the experiment at the start of each block (presented in a random order). The experimenter stayed in the room with each participant during the practice trials in order to make sure that the instructions were clear and the task was performed correctly.

Every trial started with the presentation of the central fixation light for a random duration of between 750 and 1250 ms. In the unimodal block, either a unimodal auditory or unimodal visual target was presented from the left, center, or right for 100 ms after the offset of the fixation light. There were 80 Go trials containing 40 unimodal visual (20 left, 20 right) and 40 unimodal auditory target trials (20 left, 20 right). There were 10 No-go trials (central presentation, 11% No-go trials) containing 5 unimodal visual and 5 auditory target trials. Overall, the unimodal block contained 90 trials. A schematic overview of all possible trials from the different blocks is depicted in Fig. 1. In both the Integration and Cueing blocks, the auditory and visual stimuli were always presented after the offset of the fixation light with a variable SOA of 0, 50, 100, or 200 ms. The sound was always presented first.

The Integration block consisted of 80 Go trials for each SOA including 40 spatially aligned (20 left, 20 right) and 40 spatially misaligned trials (i.e., visual left, auditory right and vice versa). The participants had to respond to both sound and light presented from either the left or right (whichever they detected first), so there were no trials in which a Go location (left or right) was combined with a No-go location (the central location; e.g., no occurrence of a visual stimulus on the left and an auditory stimulus at the center). There were 13 No-go trials for each SOA containing spatially aligned center AV stimuli (14% No-go trials). In total, the Integration block contained 320 Go trials and 52 No-go trials. The response window was set to 1500 ms after target onset, followed by the start of the next trial.

In the Cueing block, the auditory stimulus acted as an exogenous spatial cue as participants were instructed to respond only to the visual stimulus. There were 90 Go trials for each SOA divided into 40 Valid Cue (20 left cue, 20 right cue), 40 Invalid Cue (20 left cue, 20 right cue), and 10 Center cue trials. The No-go condition consisted of 15 trials for each SOA (5 left cue, 5 center cue, and 5 right cue; 14% No-go trials). There were 360 Go trials and 60 No-go trials in the Cueing block. The location of the auditory stimulus was not predictive of the visual stimulus location in either the Integration or the Cueing block.

2.4. Data analysis

Reaction times (RTs) of less than 100 ms (anticipation) or greater than 1000 ms (not paying attention to the task) were removed from further analysis. The median RT of each participant during Go trials in each condition was used in the analysis of the RT data. Only correct Go trials were used in the RT analyses. In total, 2.30% of the data was removed. In the Unimodal block 0.21% (on average, ~2 trials) of the Go and 11.67% (on average, ~1 trial) of the No-go trials were removed. In the Integration block 0.65% (on average, ~2 trials) of the Go and 13.72% (on average, ~7 trials) of the No-go trials, and in the Cueing block 0.71% (on average, ~3 trials) of the Go and 12.01% (on average, ~7 trials) of and No-go trials were removed. Both the percentage of responses on Go trials and the percentage of correctly withheld responses on No-go trials were calculated.

To investigate the magnitude of the speed-up in the multisensory condition as compared to the unimodal condition, the absolute amount of Multisensory Response Enhancement (MRE) was calculated for each participant and each condition in the Integration and the Cueing block using the following formula: MRE $=$ min[median (RT_A), median (RT_V)] – median (RT_{AV}).

In theory, the participants could also respond to the auditory stimulus in the Cueing block. Such responses cannot be distinguished from responses to visual targets at the 0 ms SOA. Therefore, multisensory response enhancement was calculated in relation to the fastest unimodal A and V RT, and not the unimodal V RT alone in the Cueing block. It is unlikely that participants would confuse the instructions

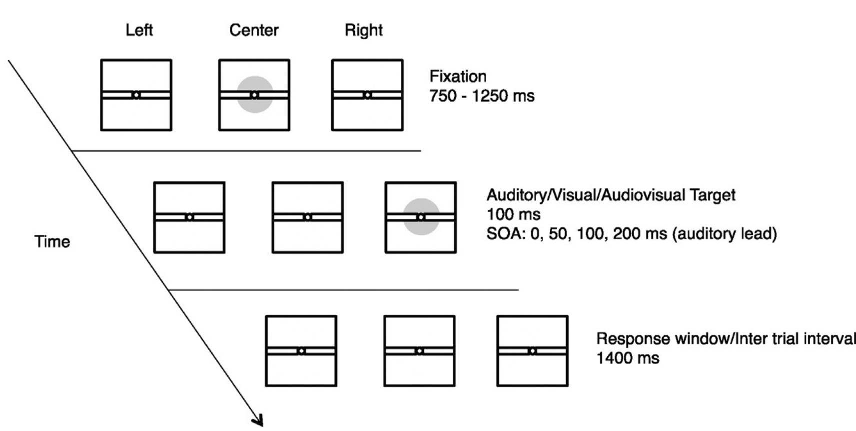

Fig. 1. Schematic representation of a trial. The grey filled circles indicate the onset of the fixation light and one of three possible locations of the onset of the visual target. Unimodal auditory and visual targets only appeared in the Unimodal block and bimodal targets only in the Integration and Cueing blocks. The response window/intertrial interval started at the offset of the auditory stimulus.

from the integration (respond to A or V, whichever is detected first) and the cueing task (respond to V only) because the tasks were presented in separate blocks of trials each with their own set of practice trials.

To investigate whether any speeding-up in the multisensory condition could be explained by statistical facilitation or by multisensory integration, the audiovisual cumulative distributive function (CDF) was compared with the race model (the sum of the unimodal CDFs) for each SOA and spatial alignment condition (Miller, 1986; Raab, 1962; Stevenson et al., 2014). Violations of the race model inequality indicate the occurrence of multisensory integration (i.e., co-activation): $P(RT_{AV} < t) \leq P(RT_A < t) + P(RT_V + SOA < t)$.[2]

Race model violations were analyzed using one-tailed paired samples t-tests for the first four quantiles of the difference between the race model and the AV CDF in the 0, 50, 100, and 200 ms SOA conditions in the Integration block and in the 0 and 50 ms SOA condition in the Cueing block (see the section on race model violation below). Race model violations were only statistically analyzed in a limited percentile range (the 10th, 20th, 30th, and 40th percentile) to avoid Type 1 error accumulation (see Kiesel, Miller, & Ulrich, 2007; Ulrich, Miller, & Schröter, 2007).

3. Results

3.1. Task performance

Overall, participants were well able to perform the task. They responded on ~99% ($SE = .1$) of the Go trials and correctly withheld their response on ~88% ($SE = 2.8$) of the No-go trials in the Unimodal block. In the Integration block, the participants responded on average on ~99% ($SE = 0.5$) of the Go trials and withheld their response on ~86% ($SE = 1.6$) of the No-go trials. Performance was similar in the Cueing block with ~99% ($SE = 0.5$) of the Go trials responded to. The participants withheld their responses on ~88% ($SE = 1.9$) of the No-go trials in the Cueing block.

3.2. Response times

3.2.1. Unimodal block

RTs in the Unimodal block were analyzed using a paired samples t-test. The difference in RT between Auditory ($M = 371$ ms, $SE = 15$) and Visual target trials was not significant ($M = 383$ ms, $SE = 13$, $t(23) = -1.824$, $p = .081$, see Fig. 2).

3.2.2. Integration block

RTs in the Integration block were analyzed using a 2 × 4 repeated measure ANOVA with Spatial Alignment (aligned, misaligned) and SOA as factors. There was a significant main effect of Spatial Alignment [$F(1, 23) = 54.527$, $p < .001$, partial $\eta^2 = .703$]. The participants' RTs were shorter when the auditory and visual target stimuli were presented from the same spatial location ($M = 365$ ms, $SE = 15$) as compared to from different spatial locations ($M = 387$ ms, $SE = 17$). The average RTs for the aligned and misaligned condition for each SOA are shown in Fig. 2A.

The main effect of SOA was also significant [$F(1.352, 31.103) = 40.173$, $p < .001$, $\varepsilon = .451$, partial $\eta^2 = .636$]. The average RT increased as the SOA increased (i.e., participants were slower at larger SOAs). Pairwise comparisons indicated significant differences between all pairs of SOAs (t's < -2.3, p's $< .001$) except for the difference in RT between the SOA_0 and SOA_{50} conditions [$t(23) = -2.350$, $p = .157$]. There was no interaction between Spatial Alignment and SOA [$F(2.024, 46.550) = 61.398$, $p = 891$, $\varepsilon = .675$, partial $\eta^2 = .005$].

Overall, participants responded most rapidly when the A and V targets were presented at approximately the same time (0 and 50 ms SOA) and from the same location.

3.2.3. Cueing block

A 3 × 4 repeated measure ANOVA with the factors Cue Type (Valid, Invalid, and Center cue) and SOA was used to analyze the RT data from the Cueing block. A main effect of Cue type (Valid, Invalid, Center cue) was observed [$F(1.242, 28.575) = 74.176$, $p < .001$, $\varepsilon = .621$, partial $\eta^2 = .763$]. Responses to validly cued visual targets were faster ($M = 324$ ms, $SE = 13$) than to invalidly cued ($M = 352$ ms, $SE = 12$, $t(23) = -13.008$, $p < .001$, $d = -.411$) and centrally cued visual targets ($M = 381$ ms, $SE = 14$, $t(23) = -10.111$, $p < .001$, $d = -.846$). RTs to invalidly cued targets were shorter than to center cued targets [$t(23) = -5.249$, $p < .001$, $d = -.404$]. The difference in RT between validly and invalidly cued visual targets indicates that exogenous spatial attention was shifted to the cued location and that this led to a speeding-up of responses to targets appearing at the cued location as compared to the uncued locations (i.e., invalid and center cue conditions). The observation of the longest RTs in the center cue condition might be explained by the fact that in this case the No-go location was cued. This could have led to an initial inhibition of a motor response that had to be overcome as soon as the visual target appeared at one of the two Go locations (left or right). The average RTs for each Cue Type and for each SOA are shown in Fig. 2B.

RTs decreased as the SOA increased as shown by a main effect of SOA [$F(1.575, 36.230) = 29.113$, $p < .001$, $\varepsilon = .525$, partial $\eta^2 = .559$]. The longer interval between the cue and the target resulted in faster responses possibly because of increased effects of alerting and response preparation (e.g., Diederich & Colonius, 2008). Pairwise comparisons revealed that responses in the SOA_0 condition were significantly slower than to all of the other SOAs (50, 100, and 200 ms SOA; t's > 6.0, p's $< .001$). The difference in RT between SOA_{50} and SOA_{200} was also significant [$t(23) = 3.380$, $p = .018$, $d = .286$], but the other comparisons failed to reach statistical significance after correcting for multiple comparisons (t's < 3.4, p's $> .1$).

The interaction between Cue Type and SOA was also significant [$F(2.560, 58.875) = 9.426$, $p < .001$, $\varepsilon = .427$, partial $\eta^2 = .291$]. This interaction was driven by an increase in the size of the spatial alignment effect (RT Invalid cue − RT Valid cue) as the SOA increased. The effects of spatial alignment were further explored in a separate analysis, as described below in Section 3.3.

Overall, responses were fastest to multisensory stimuli when the A and V targets were spatially and temporally aligned in the Integration block. A different pattern of results was observed in the Cueing block: responses were fastest when the A and V targets were spatially aligned but temporally misaligned.

3.3. Spatial alignment effects

The difference in RT between the spatially aligned and misaligned conditions is often interpreted as the result of a shift of crossmodal exogenous spatial attention at SOAs larger than 0 ms in crossmodal cueing tasks such as that presented in the Cueing block. The effect of spatial alignment in a redundant target task such as in the Integration block is often interpreted as being the result of overlapping (vs. not overlapping) visual and auditory receptive fields of multisensory neurons. To further investigate the effects of spatial alignment on RT and its relation with temporal (mis)alignment, the spatial alignment effect between the four SOAs for the Integration and the Cueing block were compared. The spatial alignment effect for each SOA is shown in Fig. 3A and B for the Integration and Cueing blocks, respectively. A repeated measures ANOVA with SOA as the factor was used in order to analyze the size of

[2] The race model inequality can also be expressed as $P(RT_{AV} < t) \leq P(RT_A < t) + P(RT_V + SOA < t) - P(RT_A < t) \times P(RT_V + SOA < t)$ in which a maximum negative correlation between the sensory channels is not assumed. Both versions of the race model inequality have been used in the literature, but we chose to use the most conservative version. The data were also analyzed using this version, but the results remained unchanged.

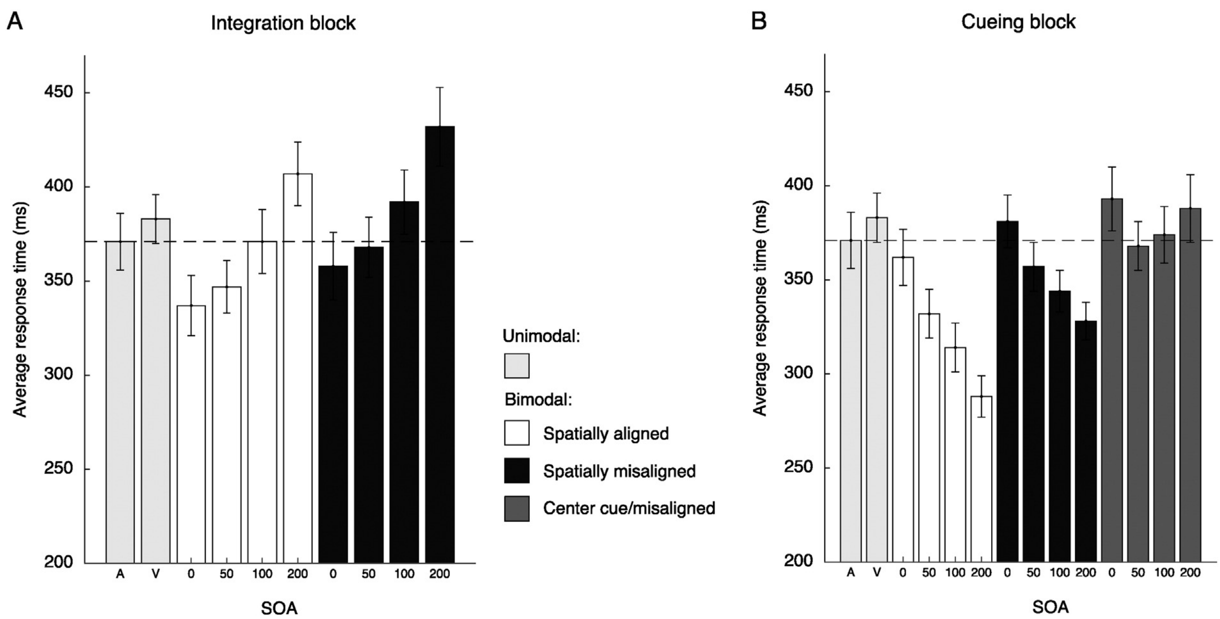

Fig. 2. The average RTs in the unimodal conditions (A and V in panel A and B), and the multisensory conditions in the Integration (panel A) and in the Cueing block (panel B) for each SOA (0, 50, 100, 200 ms). The dashed line indicates the shortest average unimodal response time. Error bars represent standard error of the mean.

the spatial alignment effect across SOAs for the Integration and the Cueing block.

3.3.1. Integration block

The size of the spatial alignment effect did not differ between the four different SOAs in the Integration block ($F(2.024, 46.550) = .118$, $p = .891$, $\varepsilon = .675$, partial $\eta^2 = .005$, see Fig. 3A: mean spatial alignment effect ~ 20 ms). One-sample t-tests indicated that the spatial alignment effect was significantly different from zero at all of the SOA's (t's > 3, p's < .05).

3.3.2. Cueing block

A main effect of SOA was observed for the size of the spatial alignment effect [$F(3, 69) = 5.555$, $p = .002$, partial $\eta^2 = .195$]. The spatial alignment effect was significantly larger in the SOA_{200} condition (M validity effect = 41 ms, $SE = 4$) than in the SOA_0 condition (M validity effect = 19 ms, $SE = 4$; $t(23) = -3.560$, $p = .012$, $d = -1.09$). The other pairwise comparisons between the different levels of SOA did not survive the correction for multiple comparisons (t's < 2.3, p's > .05).

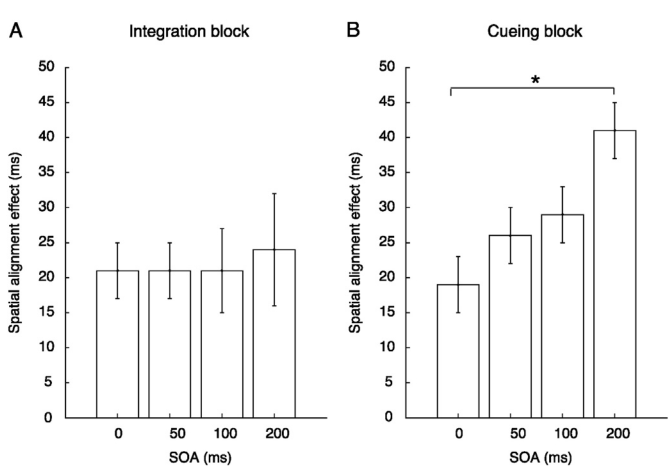

Fig. 3. The effect of spatial alignment (i.e., RT AV misaligned − RT AV aligned) for each SOA in the Integration (panel A) and the Cueing block (panel B). Significant differences are indicated with an asterisk ($p < .05$, corrected for multiple comparisons).

One-sample t-tests indicated that the spatial alignment effect was significantly different from zero at all SOAs (t's > 4.7, p's < .001). These results indicate that the effect of spatial alignment increased as the SOA increased in the Cueing block.

To summarize, the effect of spatial alignment remained the same despite increased temporal misalignment when the participants were allowed to respond to both A and V targets (in the Integration block). However, when the participants were only allowed to respond to visual targets (in the Cueing block), the A stimulus acted as a crossmodal exogenous spatial cue. As the SOA increased, the spatial alignment effect increased in the Cueing block. The simultaneous presentation of the A and V stimulus did not provide sufficient time in which to elicit a crossmodal exogenous shift of spatial attention, resulting in a spatial alignment effect of the same size as in the Integration block at 0 ms SOA. The observation of an increase in the effect of spatial alignment in the Cueing block is in line with previous studies of exogenous spatial attention in which the effects of exogenous spatial attention are often most pronounced around 200 ms SOA (e.g., Berger et al., 2005; Spence & McDonald, 2004).

3.4. Multisensory response enhancement

The amount of MRE was analyzed using a 2 × 4 repeated measures ANOVA with the factors of Spatial Alignment (aligned vs. misaligned) and SOA for the Integration and Cueing blocks. The average MRE for each alignment condition and SOA is shown in Fig. 4.

3.4.1. Integration block

There was a main effect of Spatial Alignment on the amount of MRE [$F(1, 23) = 54.527$, $p < .001$, partial $\eta^2 = .703$]. The amount of MRE was significantly larger when the A and V stimuli were spatially aligned ($M = -2.167$, $SE = 10.820$) as compared to spatially misaligned targets ($M = -23.880$, $SE = 12.342$). The average MRE was negative as it was the average of the four SOAs and thus included multisensory response inhibition at the longer SOAs.

The main effect of SOA was also significant [$F(1.352, 31.103) = 40.173$, $p < .001$, $\varepsilon = .451$, partial $\eta^2 = .636$]. The amount of MRE decreased significantly as the SOA increased. The amount of MRE was not significantly different between the SOA_0 and the SOA_{50} conditions after correction [$t(23) = 2.350$, $p = .157$], but it was between all of

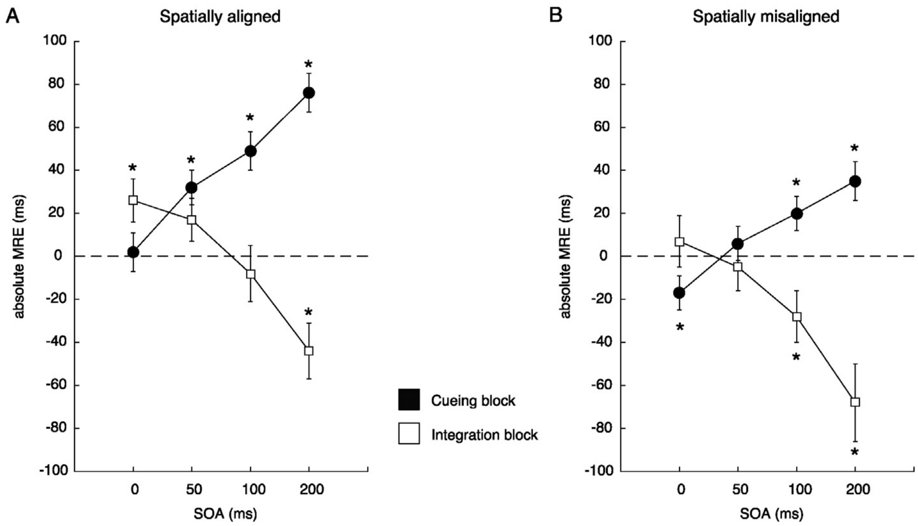

Fig. 4. The average absolute multisensory response enhancement (positive values) and multisensory response interference (negative values) for each SOA in the Cueing (black symbols) and the Integration block (white symbols) in the spatially aligned (panel A) and misaligned (panel B) condition. Significant MRE (different from zero) is indicated with an asterisk ($p < .05$).

the other pairs of SOAs (t's > 5.0, p's < .001). There was no significant interaction between Spatial Alignment and SOA [$F(2.024, 46.550) = .118$, $p = .891$, $\varepsilon = .675$, partial $\eta^2 = .005$].

Separate one-sample t-tests for each SOA in the spatially aligned condition revealed a significant MRE in the SOA_0 condition ($M = 26$ ms, $SE = 10$) when the auditory and visual stimuli were spatially aligned [$t(23) = 2.58$, $p = .017$, d = .745]. There was no significant MRE in the SOA_{50} ($M = 17$ ms, $SE = 10$, $t(23) = 1.705$, $p = .102$) and SOA_{100} conditions ($M = -8$ ms, $SE = 13$, $t(23) = -.589$, $p = .562$). Significant multisensory response inhibition (MRI) was observed in the SOA_{200} condition ($M = -44$ ms, $SE = 13$, $t(23) = -3.252$, $p = .004$, d = −.939).

There was no significant MRE in the SOA_0 ($M = 6$ ms, $SE = 12$, $t(23) = .465$, $p = .646$) and SOA_{50} conditions ($M = -5$ ms, $SE = 11$, $t(23) = -.402$, $p = .691$) when the A and V stimuli were spatially misaligned. There was significant MRI in the SOA_{100} ($M = -28$ ms, $SE = 12$, $t(23) = -2.357$, $p = .027$) and the SOA_{200} conditions ($M = -68$ ms, $SE = 18$, $t(23) = -3.896$, $p = .001$). These results indicate that absolute multisensory response enhancement was largest when the A and V targets were spatially and temporally aligned in the Integration block.

3.4.2. Cueing block

A main effect of Spatial Alignment was observed in the Cueing block [$F(1, 23) = 169.200$, $p < .001$, partial $\eta^2 = .880$]. The amount of MRE was significantly larger in the spatially aligned condition ($M = 39$ ms, $SE = 8$) than in the spatially misaligned condition ($M = 11$ ms, $SE = 8$).

The main effect of SOA was also significant [$F(1.837, 42.247) = 78.991$, $p < .001$, $\varepsilon = .612$, partial $\eta^2 = .774$]. In contrast with the results from the Integration block in which the amount of MRE decreased as the SOA increased, here the amount of MRE increased as the SOA increased. The difference in MRE between the 0 and 200 ms SOA conditions was significant [$t(23) = -3.560$, $p = .012$, d = −.961], but the other comparisons failed to reach statistical significance after correction (t's < 2.8, p's > .05).

There was an interaction between Spatial Alignment and SOA [$F(3, 69) = 5.555$, $p = .002$, partial $\eta^2 = .195$]. This interaction could be explained by a larger increase in the amount of MRE as the SOA increased in the spatially aligned condition as compared to the spatially misaligned condition (this is essentially the increase of the spatial alignment effect as SOA increases, see Fig. 3B).

One-sample t-tests on the absolute MRE for each SOA revealed a significant absolute MRE in the SOA_{50} ($M = 32$ ms, $SE = 8$), SOA_{100} ($M = 49$ ms, $SE = 9$), and SOA_{200} ($M = 76$ ms, $SE = 9$) condition (t's > 3.9, p's < .005) when the auditory and visual stimuli were spatially *aligned*, but not in the SOA_0 condition ($M = 2$ ms, $SE = 9$, $t(23) = .177$, $p = .861$).

When the auditory and visual stimuli were spatially *misaligned* significant MRI was observed in the SOA_0 condition ($M = -17$ ms, $SE = 8$, $t(23) = -2.169$, $p = .041$) and significant MRE at the 100 ms ($M = 20$ ms, $SE = 8$, $t(23) = -2.346$, $p = .028$) and 200 ms SOAs ($M = 35$ ms, $SE = 9$, $t(23) = -3.899$, $p = .001$). There was no MRE in the SOA_{50} condition ($M = -6$ ms, $SE = 8$, $t(23) = .781$, $p = .443$).

Taken together, when the A and V stimuli were spatially aligned, MRE was only observed at the 0 and 50 ms SOAs when participants were allowed to respond to both A and V stimuli. These results are in line with the idea that multisensory integration often occurs within a certain temporal binding window (see Footnote 1; Meredith, Nemitz, & Stein, 1987; Stein & Stanford, 2008). This MRE changed into significant response inhibition at the 200 ms SOA. In the cueing task, the opposite pattern of results was observed: that is, there was no MRE at the 0 ms SOA, but multisensory response enhancement started to emerge and increase from the 50 SOA up until the 200 ms SOA. This increase in MRE could be the result of a crossmodal shift of exogenous spatial attention with added non-spatial effects of alerting, and/or response preparation as the auditory stimulus acted as an exogenous spatial cue and a warning signal with SOAs larger than 0 ms.

In the spatially misaligned condition, there was no MRE in the Integration block at any of the SOAs, but MRI in the 100 and 200 ms SOA condition. In the Cueing block, there was significant MRI at the 0 ms SOA, and MRE at the 100 and 200 ms SOAs.

3.5. Race model violation

3.5.1. Integration block

To investigate whether the observed multisensory response enhancement effect could be explained merely in terms of the statistical facilitation or multisensory integration, the amount of race model violation was analyzed for different SOAs and Spatial alignment conditions.

The race model equality was significantly violated in the spatially aligned condition at the 10th percentile in the SOA_0 condition. At the 50 ms SOA, race model violation was observed from the 10th to the 30th percentile (all p's < .05). There was no race model violation in the SOA_{100} and SOA_{200} conditions, nor was there at any of the SOA conditions in the spatially misaligned conditions (all p's > .05). Thus, the amount of MRE observed at the 0 and 50 ms SOAs was the result of multisensory integration. See Fig. 5 for the average amount of race model violation in the spatially aligned and misaligned condition at the different SOAs in the Integration and Cueing block.

3.5.2. Cueing block

In the Cueing block, the auditory stimulus acted as a cue/warning signal for an upcoming visual target and participants were instructed to respond only to the visual targets. Therefore, in the Cueing block, multisensory response enhancement could in part also be the result of crossmodal exogenous spatial attention (e.g., Spence & McDonald, 2004; also see Figs. 3B and 5), alerting, and response preparation (e.g., Los & Schut, 2008; see also Los & Van der Burg, 2013). In the Cueing block, race model violations for those SOA conditions in which there were significant race model violations in the Integration block (the SOA_0 and SOA_{50} condition) are of most interest. This is because additional alerting and preparation do not affect any differences in the amount of race model inequality violation between SOAs in the Integration block, as the participants could always respond to the first stimulus that was presented. This was not the case in the Cueing block, which allowed for additional effects of alerting and response preparation.

In the Cueing block, the race model was not violated in the SOA_0 condition, but it was significantly violated in the spatially aligned condition from the 10th to the 40th percentile in the SOA_{50}, $SOA_{100,}$ and SOA_{200} condition (all p's < .05). In the SOA_{100} and the SOA_{200} condition the race model inequality was violated even in the spatially misaligned condition (all p's < .05). The race model inequality was never violated in the Center Cue condition.

4. Discussion

The aim of the present study was to investigate the relative contributions of multisensory integration and crossmodal exogenous spatial attention to audiovisual multisensory response enhancement (MRE). To do so, two tasks were utilized that are often used to independently measure multisensory integration and crossmodal exogenous spatial attention: an RTE paradigm (e.g., Laurienti et al., 2006; Miller, 1982, 1986; Stevenson et al., 2012; Van der Stoep, Van der Stigchel, & Nijboer, 2015) and a crossmodal exogenous cueing paradigm (see Spence & McDonald, 2004, for a review), respectively. The presentation of the auditory and visual stimuli at four different SOAs (0, 50, 100, and 200 ms) and the addition of a separate unimodal RT baseline in both an RTE task (the Integration block) and a crossmodal cueing task (the Cueing block) allowed us to investigate the timecourse at which integration and crossmodal exogenous spatial attention contribute to MRE.

In line with the literature on multisensory integration, MRE in the Integration block was present only when the auditory and visual stimuli were presented in close spatial and temporal alignment (here at the 0 and 50 ms SOA; cf. Leone & McCourt, 2013; Stevenson et al., 2012). The race model inequality violation analysis indicated that the amount of MRE that was observed at the 0 and 50 ms SOAs in the Integration block could be interpreted as the result of multisensory integration. It is unlikely that the observed MRE at 0 ms was the result of a crossmodal exogenous spatial attention shift as participants were allowed to respond to the first stimulus that they detected.

In some cases, it is possible that MRE at 0 ms SOA can be accounted for in part by the non-spatial effects of alerting (Diederich & Colonius, 2008) or even in full by response preparation (Los & Van der Burg, 2013). For example, it has been shown that response preparation can speed up responses to multisensory stimuli even at an SOA of 0 ms, but only when there are differences in auditory and visual central arrival times and when responses are made towards the slowest signal (Los & Van der Burg, 2013). If auditory processing finishes before visual processing, the auditory stimulus can act as a warning signal and start temporal preparation. Given that the unimodal RTs in the current study were not significantly different, there was (almost) no room for temporal preparation to facilitate responses and contribute to the MRE that was observed here (see Los & Van der Burg, 2013). Furthermore, one would expect temporal preparation to contribute equally to MRE in the spatially aligned and misaligned condition. Given that the MRE and race model inequality violation were affected by the spatial alignment of sound and light, temporal preparation cannot explain the current results.

MRE in the Cueing block could potentially be the result (at least in part) of multisensory integration (see Fig. 5). Therefore, it was especially interesting to take a closer look at the amount of MRE in the

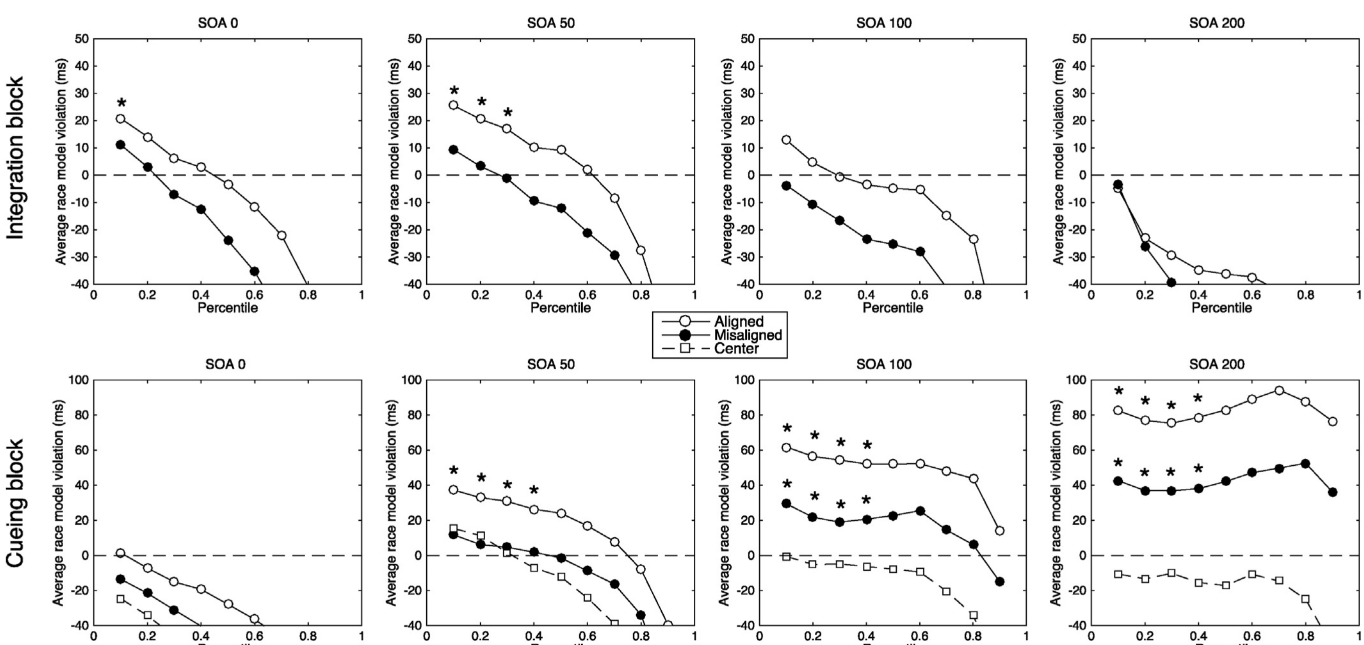

Fig. 5. The average amount of race model violation in milliseconds across the 9 quantiles for the SOA_0, $SOA_{50,}$ SOA_{100}, and SOA_{200} conditions in the spatially aligned (white dots) and spatially misaligned conditions (black dots) in the Integration block (top row) and the Cueing block (bottom row, white squares indicate the center cue condition). Significant race model violations within the 10 to 40 percentile range are indicated with an asterisk ($p < .05$).

Cueing block at those SOAs at which multisensory integration was observed in the Integration block. There was no MRE at the 0 ms SOA in the Cueing block, whereas there was significant MRE in the Integration block at this SOA. At the 50 ms SOA, however, there was significant MRE in the Cueing block. As in the Integration block, the race model inequality was violated at the 50 ms SOA in the Cueing block, indicating that the observed MRE was, at least in part, the result of multisensory integration. Importantly, however, unlike in the Integration block in which the size of the spatial alignment effect remained the same across all SOAs (see Fig. 3A), the spatial alignment effect increased as the SOA increased in the Cueing block (see Fig. 3B). It is well-established that the behavioral benefits of exogenous spatial attention increase as the time between the cue and the target increases just as we observed here (up until a certain stimulus onset asynchrony after which inhibition of return can sometimes emerge; e.g., Berger et al., 2005, and Driver & Spence, 1998, for example). It is therefore likely that crossmodal exogenous spatial attention also contributed to the observed MRE at the 50 ms SOA in the Integration block.

Even though we observed race model inequality violations in the 100 and the 200 ms SOA conditions in the Cueing block, we do not think that the observed MRE in these conditions was the result of multisensory integration. The reason for this is that there was no multisensory integration at the 100 and 200 ms SOAs in the Integration block, which makes it unlikely that the observed MRE was the result of integration. We argue that the amount of MRE at the 100 and 200 ms SOAs in the Cueing task was the result of the combined effects of crossmodal exogenous spatial attention (as indicated by the difference between validly and invalidly cued visual targets), alerting, and temporal preparation effects. This is indicated by the fact that the mere presence of an auditory exogenous cue resulted in shorter RTs compared to the shortest unimodal RTs, even in the spatially misaligned condition. Further support for this interpretation of the data can be found in the Central Cue condition in the Cueing task. In this condition, the No-go location was cued, after which a visual target was presented at one of the Go locations. Given that the central location is a No-go location, any response preparation evoked by the cue has to be inhibited in order for the participant to perform the task correctly. This reduces the benefits of any response preparation induced by the center cue, which can clearly be seen in the lack of MRE (see Fig. 2B) and race model inequality violation at any of the SOAs in the Center Cue condition (see Fig. 5).

Using race model inequality violation as a measure of multisensory integration is not without its limitations. One must carefully consider the paradigm that was used when interpreting race model violations, given that the amount of race model violation can be affected by various factors that are perhaps not directly linked to multisensory integration. For example, response preparation and modality switch effects can also contribute to a (relative) speed up of responses to multisensory stimuli (e.g., Gondan, Lange, Rösler, & Röder, 2004; Los & Van der Burg, 2013; Otto & Mamassian, 2012; Otto, Dassy, & Mamassian, 2013; Van der Stoep, Van der Stigchel, & Nijboer, 2015). That being said, in the redundant target effect paradigm (i.e., the integration block) race model violations have been shown to relate to neural measures of multisensory integration (e.g., Gondan, Niederhaus, Rösler, & Röder, 2005; Molholm et al., 2002, 2006; also see Mercier et al., 2015). This provides support for the idea that race model inequality violation reflects multisensory integration when the right paradigm is used.

The lack of any multisensory integration at the 0 ms SOA in the Cueing block may be unexpected. However, the main difference between the Cueing and the Integration blocks was that participants were instructed to respond to both the onset of sound and light in the Integration block and only to the onset of light in the Cueing block. In the current experiment, attention may have been divided between audition and vision in the Integration block and focused on vision in the Cueing block. As a result of selective attention to the visual modality, visual stimulus processing may have been enhanced relative to auditory stimulus processing (e.g., McDonald et al., 2005; Spence & Parise, 2010). It has been shown that selectively attending to only one sense also modulates multisensory integration of sensory information of that and another sensory modality (e.g., Hugenschmidt, Mozolic, & Laurienti, 2009; Mozolic, Hugenschmidt, Peiffer, & Laurienti, 2008; Talsma, Doty, & Woldorff, 2007). In some cases, modality-specific attention even reduced multisensory response enhancement to statistical facilitation (see Mozolic et al., 2008, Fig. 2C). However, there is also evidence that multisensory integration (or association) might still occur at a later stage of processing under certain circumstance (see Busse, Roberts, Crist, Weissman, & Woldorff, 2005; Talsma et al., 2007). The fact that MRE was observed at the 50 ms SOA in the Cueing block seems to contradict this explanation. However, at the 50 ms SOA, auditory stimulus processing is given a head start, which may have reduced the differences in auditory and visual processing times due to modality specific attention, and may have allowed for multisensory integration to occur again (see also Spence, Shore, & Klein, 2001, for similar results using visual and tactile stimuli).

Any difference in RTs between spatially aligned and misaligned conditions is often interpreted as the result of crossmodal exogenous spatial attention in crossmodal spatial cueing paradigms. The current findings indicate that at shorter cue-target intervals, the speeding-up of a participant's responses can also be the result of multisensory integration. This may not be surprising as the auditory and visual target at short SOAs have a higher chance of being presented within the temporal binding window, thus increasing the probability that the two stimuli will be integrated (i.e., the temporal rule; e.g., Colonius & Diederich, 2012; Stein & Meredith, 1990; Stein & Stanford, 2008, for a review). The effect of spatial alignment remained constant across different SOAs when attention was divided between audition and vision (in the Integration block). However, the effect of spatial alignment increased beyond the size of the spatial alignment effect (~20 ms) in the Integration block as the SOA increased in the Cueing block (up to ~40 ms at the 200 ms SOA). The increase in the spatial alignment effect as the SOA increases is in line with the literature on crossmodal exogenous spatial attention (see Spence & McDonald, 2004, for a review). As has been proposed previously, it is possible that spatial alignment effects during multisensory integration and crossmodal exogenous spatial attention effects are mediated by partially overlapping neuronal substrates (e.g., Spence, 2010, pp. 183–184). It is therefore difficult to say which neuronal mechanisms are involved in the spatial alignment effects in the integration and spatial cueing tasks at each SOA.

The present study is one of the first to explore the relative contributions of multisensory integration and crossmodal exogenous spatial attention to multisensory response enhancement. The results indicate that audiovisual multisensory response enhancement resulting from the presentation of simple lights and sounds can be explained by multisensory integration when auditory and visual stimuli are spatially and temporally aligned *and* attention is divided across the two modalities. Our results indicate that at the 50 ms SOA, the amount of MRE likely results from multisensory integration and crossmodal exogenous spatial attention when vision is selectively attended. Alerting effects and response preparation also start to contribute to multisensory response enhancement in addition to multisensory integration at SOAs from 50 ms. At SOAs of 100 and 200 ms, multisensory response enhancement cannot be explained by multisensory integration when vision is selectively attended, and is likely driven by the effects of crossmodal exogenous spatial attention, alerting, and response preparation.

Acknowledgments

This research was funded by two grants from the Netherlands Organisation for Scientific Research (NWO): Grant nos. 451-09-019 (to S.V.d.S.) and 451-10-013 (to T.C.W.N.). Charles Spence would like to thank the AHRC for the Rethinking the Senses grant (AH/L007053/1).

References

Berger, A., Henik, A., & Rafal, R. (2005). Competition between endogenous and exogenous orienting of visual attention. *Journal of Experimental Psychology: General*, *134*(2), 207–221.

Busse, L., Roberts, K.C., Crist, R.E., Weissman, D.H., & Woldorff, M.G. (2005). The spread of attention across modalities and space in a multisensory object. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(51), 18751–18756.

Canzoneri, E., Magosso, E., & Serino, A. (2012). Dynamic sounds capture the boundaries of peripersonal space representation in humans. *PLoS ONE*, *7*(9), e44306.

Colonius, H., & Diederich, A. (2004). Multisensory interaction in saccadic reaction time: A time-window-of-integration model. *Journal of Cognitive Neuroscience*, *16*(6), 1000–1009.

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*, 1–10.

Diederich, A., & Colonius, H. (2008). Crossmodal interaction in saccadic reaction time: Separating multisensory from warning effects in the time window of integration model. *Experimental Brain Research*, *186*(1), 1–22.

Diederich, A., & Colonius, H. (2011). Modeling multisensory processes in saccadic responses: Time-window-of-integration model. In M.M. Murray, & M.T. Wallace (Eds.), *The neural bases of multisensory processes* (pp. 253–278). Boca Raton, FL: CRC Press.

Driver, J., & Spence, C. (1998). Cross-modal links in spatial attention. *Philosophical Transactions of the Royal Society of London, Series B*, *353*, 1319–1331.

Gondan, M., Lange, K., Rösler, F., & Röder, B. (2004). The redundant target effect is affected by modality switch costs. *Psychonomic Bulletin & Review*, *11*(2), 307–313.

Gondan, M., Niederhaus, B., Rösler, F., & Röder, B. (2005). Multisensory processing in the redundant-target effect: A behavioral and event-related potential study. *Perception & Psychophysics*, *67*(4), 713–726.

Hugenschmidt, C.E., Mozolic, J.L., & Laurienti, P.J. (2009). Suppression of multisensory integration by modality-specific attention in aging. *Neuroreport*, *20*(4), 349–353.

Kiesel, A., Miller, J., & Ulrich, R. (2007). Systematic biases and Type I error accumulation in tests of the race model inequality. *Behavior Research Methods*, *39*(3), 539–551.

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*(3), 492–500.

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.

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

Los, S.A., & Schut, M.L. (2008). The effective time course of preparation. *Cognitive Psychology*, *57*(1), 20–55.

Los, S.A., & Van der Burg, E. (2013). Sound speeds vision through preparation, not integration. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(6), 1612–1624.

McDonald, J.J., Teder-Sälejärvi, W.A., Di Russo, F., & Hillyard, S.A. (2003). Neural substrates of perceptual enhancement by crossmodal spatial attention. *Journal of Cognitive Neuroscience*, *15*(1), 10–19.

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.

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

McDonald, J. J., Teder-Sälejärvi, W. A., & Ward, L. M. (2001). Multisensory integration and crossmodal attention effects in the human brain. Science, 292(5523), 1791-1791.

McDonald, J.J., & Ward, L.M. (1999). Spatial relevance determines facilitatory and inhibitory effects of auditory covert spatial orienting. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(5), 1234–1252.

McDonald, J.J., & Ward, L.M. (2000). Involuntary listening aids seeing: Evidence from human electrophysiology. *Psychological Science*, *11*(2), 167–171.

Mercier, M.R., Molholm, S., Fiebelkorn, I.C., Butler, J.S., Schwartz, T.H., & Foxe, J.J. (2015). Neuro-oscillatory phase alignment drives speeded multisensory response times: An electro-corticographic investigation. *Journal of Neuroscience*, *35*(22), 8546–8557.

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*(10), 3215–3229.

Meredith, M.A., & Stein, B.E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, *56*(3), 640–662.

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.

Molholm, S., Ritter, W., Murray, M.M., Javitt, D.C., Schroeder, C.E., & Foxe, J.J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Cognitive Brain Research*, *14*(1), 115–128.

Molholm, S., Sehatpour, P., Mehta, A.D., Shpaner, M., Gomez-Ramirez, M., Ortigue, S., ... Foxe, J.J. (2006t). Audio-visual multisensory integration in superior parietal lobule revealed by human intracranial recordings. *Journal of Neurophysiology*, *96*(2), 721–729.

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.

Ngo, M.K., & Spence, C. (2010). Auditory, tactile, and multisensory cues facilitate search for dynamic visual stimuli. *Attention, Perception, & Psychophysics*, *72*(6), 1654–1665.

Otto, T.U., & Mamassian, P. (2012). Noise and correlations in parallel perceptual decision making. *Current Biology*, *22*(15), 1391–1396.

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

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

Sambo, C.F., & Forster, B. (2009). An ERP investigation on visuotactile interactions in peripersonal and extrapersonal space: Evidence for the spatial rule. *Journal of Cognitive Neuroscience*, *21*, 1550–1559.

Soto-Faraco, S., Navarra, J., & Alsius, A. (2004). Assessing automaticity in audiovisual speech integration: Evidence from the speeded classification task. *Cognition*, *92*(3), B13–B23.

Spence, C. (2010). Crossmodal spatial attention. *Annals of the New York Academy of Sciences*, *1191*(1), 182–200.

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. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, 59(1), 1-22.

Spence, C., & Driver, C. (2000). Attracting attention to the illusory location of a sound: Reflexive crossmodal orienting and ventriloquism. *Neuroreport*, *11*(9), 2057–2061.

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

Spence, C., & McDonald, J. (2004). The cross-modal consequences of the exogenous spatial orienting of attention. In G.A. Calvert, C. Spence, & B.E. Stein (Eds.), *The handbook of multisensory processes* (pp. 3–26). Cambridge, MA: MIT Press.

Spence, C., & Parise, C. (2010). Prior entry. *Consciousness and Cognition*, *19*(1), 364–379.

Spence, C., & Santangelo, V. (2009). Capturing spatial attention with multisensory cues: A review. *Hearing Research*, *258*(1), 134–142.

Spence, C., Shore, D.I., & Klein, R.M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, *130*(4), 799–832.

Stein, B.E., & Meredith, M. (1990). Multisensory integration. *Annals of the New York Academy of Sciences*, *608*(1), 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*(4), 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*(1), 121–137.

Stevenson, R.A., Ghose, D., Fister, J.K., Sarko, D.K., Altieri, N.A., Nidiffer, A.R., ... Wallace, M.T. (2014). Identifying and quantifying multisensory integration: A tutorial review. *Brain Topography*, *27*(6), 707–730.

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 United States of America*, *106*(52), 22456–22461.

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.

Ulrich, R., Miller, J., & Schröter, H. (2007). Testing the race model inequality: An algorithm and computer programs. *Behavior Research Methods*, *39*(2), 291–302.

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

Van der Stoep, N., Nijboer, T. C. W., & Van der Stigchel, S. (2014). Exogenous orienting of crossmodal attention in 3-D space: Support for a depth-aware crossmodal attentional system. *Psychonomic Bulletin & Review*, *21*(3), 708–714.

Van der Stoep, N., Nijboer, T.C.W., Van der Stigchel, S., & Spence, C. (2015). Multisensory interactions in the depth plane in front and rear space: A review. *Neuropsychologia*, *70*, 335–349.

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.

Vroomen, J., Bertelson, P., & de Gelder, B. (2001a). Directing spatial attention towards the illusory location of a ventriloquized sound. *Acta Psychologica*, *108*(1), 21–33.

Vroomen, J., Bertelson, P., & de Gelder, B. (2001b). The ventriloquist effect does not depend on the direction of automatic visual attention. *Perception & Psychophysics*, *63*(4), 651–659.

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics*, *72*(4), 871–884.

Ward, L.M., McDonald, J.J., & Lin, D. (2000). On asymmetries in cross-modal spatial attention orienting. *Perception & Psychophysics*, *62*(6), 1258–1264.