Into the depths of spatial attention and multisensory integration

Nathan van der Stoep

ISBN 978-90-6464-909-7

Cover illustratie: Nathan van der Stoep

Lay-out: Ferdinand van Nispen, Citroenvlinder DTP&Vormgeving, my-thesis.nl,
Bilthoven, the Netherlands

Printed by: GVO drukkers & vormgevers B.V. | Ponsen & Looijen, Ede, The Netherlands

Into the depths of spatial attention and multisensory integration

Ruimtelijke aandacht en multisensorische integratie in de diepte-dimensie
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht

op gezag van de rector magnificus, prof. dr. G.J. van der Zwaan, ingevolge het besluit
van het college voor promoties in het openbaar te verdedigen op

vrijdag 2 oktober 2015 des ochtends te 10.30 uur

door

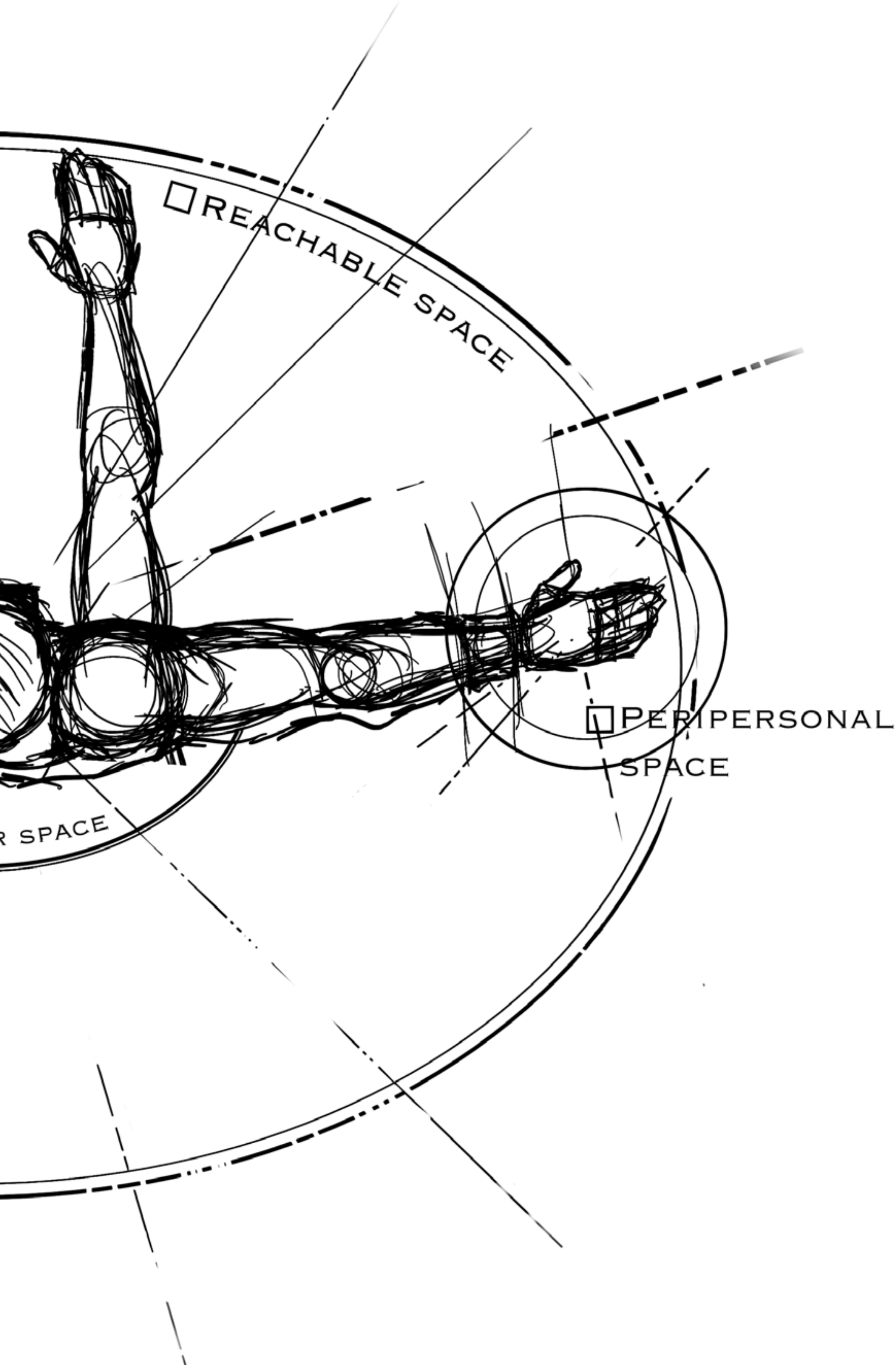
Nathan van der Stoep

geboren op 22 augustus 1987
te Dordrecht

Promotor:
Prof. dr. A. Postma

Copromotoren:
Dr. S. Van der Stigchel
Dr. T.C.W. Nijboer

Contents


☐ Reachable space
☐ Peripersonal space
r space

Chapter 1

Introduction

During our daily lives our senses are flooded with information. We can see, hear, feel, smell, and taste all at the same time, but we generally do not experience all these sensations as separate events. Although we are not aware of it, our brain is helping us to make sense of this abundant information by combining information from different senses. Of all of our senses, only vision and audition allow us to perceive information that is currently out of reach[1]. Whereas vision enables us to determine the spatial location of something we see in *frontal* space only (i.e., left/right, up/down, near/far; though with the use of mirrors we can see behind us), audition helps us with localizing sounds *all around* us (i.e., front/back, left/right, up/down, near/far). In the last couple of decades, there has been a growing interest in how vision can affect audition, and vice versa (see, for example chapter 6, Figure 1). Two processes through which such interactions can occur are *crossmodal exogenous spatial attention* and *multisensory integration*. These processes are essential for multisensory perception and spatial orienting, and are central to the studies that are described in the current thesis. After a brief introduction of the concepts of crossmodal exogenous spatial attention and multisensory integration, I will highlight how the studies described in this thesis provided an in-depth look into multisensory spatial processing leading to a deeper understanding of these processes and their contribution to multisensory spatial perception. Not only do they add to the understanding of human multisensory perception, they also provide a foundation for the application of these findings to more practical, real-life situations (see the final chapter).

Crossmodal exogenous spatial attention

The sudden onset of, for example, a sound (i.e., an auditory cue) can evoke an exogenous (or automatic/reflexive) crossmodal shift of spatial attention to its spatial location and facilitate the processing of visual information that happens to be presented at that same location (Calvert, Spence, & Stein, 2004; Spence & Driver, 2004). By now, it has been shown that such crossmodal shifts of exogenous spatial attention can occur between all pairs of auditory, visual, and tactile stimuli (Spence & McDonald, 2004). The effects of a shift of crossmodal exogenous spatial attention are generally most pronounced when there is some time between the onset of the sound and the light (~100-300 ms, e.g., Berger, Henik, & Rafal, 2005; Spence & Driver, 1997) and when they originate from approximately the same spatial location (Spence & McDonald, 2004).

---

[1] Our sense of touch is relatively constrained in terms of depth perception as we can only feel things that that are within reach. Perceiving touch is thus bound to the body. This causes interesting spatial limitations on multisensory interactions involving touch, which seem to be quite different from audiovisual interactions in different depth-planes. An in-depth review of this matter can be found in chapter 6.

Crossmodal exogenous spatial attention shifts result in faster response times (RTs) and higher detection sensitivity for information appearing at attended as compared to unattended locations (McDonald, Teder-Sälejärvi, & Hillyard, 2000; Santangelo & Spence, 2009).

When there is more time between the onset of the sound and the light (>300 ms), an inhibitory after-effect can be observed which is often labeled 'inhibition of return' (IOR) (Klein, 2000; Spence, Lloyd, McGlone, Nicholls, & Driver, 2000; Spence, Nicholls, Gillespie, & Driver, 1998). Although IOR has first been observed in the visual modality (i.e., an inhibitory after effect for lights appearing some time after the onset of a visual cue at the same location; Posner & Cohen, 1984), it has now been demonstrated between all possible pairings of audition, vision, and touch (Klein, 2000; Spence et al., 1998, 2000). In contrast with *intramodal* visual IOR (a visual cue preceding a visual target), *crossmodal* IOR seems to emerge at somewhat longer cue-target intervals (e.g., Schmitt, Postma, & De Haan, 2000, 2001; Schmitt, Spence & Driver, 2000; Ward, McDonald, & Lin, 2000).

Multisensory integration

Support for the idea that sound and light can also be integrated by the brain comes from, now classic, neurophysiological studies of multisensory integration (e.g., King & Palmer, 1985; Meredith, Nemitz, & Stein, 1987; Meredith & Stein, 1986; Stein & Meredith, 1993). The researchers observed that certain neurons in the superior colliculus (a mid-brain structure) not only responded to the onset of a light, but also to the onset of a sound. The presentation of sound and light within close spatial and temporal proximity evoked a much larger response relative to the sum of the activity that was measured when only a single sense was stimulated. When sound and light were not presented from approximately the same location no 'superadditive' response was observed in these so-called 'multisensory' neurons. Similarly, sound and light did not seem to be integrated when there was a relatively long delay between the onset of the sound and the light. Two rules or principles were formulated based on these observations: the spatial and temporal rule, respectively (Stein & Meredith, 1990; Stein & Stanford, 2008). Perfect spatial and temporal alignment is not required for multisensory integration to occur as long as the sound and light are presented within close spatial (Meredith & Stein, 1986; Stein & Meredith, 1993; Stein & Stanford, 2008) and temporal proximity (e.g., Colonius & Diederich, 2012; Meredith et al., 1987; Miller, Pluta, Stein, & Rowland, 2015; Vroomen & Keetels, 2010). A third principle, called 'the principle of inverse effectiveness', was formulated based on the observation that the relative increase in spike rate in multisensory neurons after multisensory stimulation was larger when the unimodal stimuli (sound or light alone) evoked only a weak response as compared to those unimodal stimuli that evoked a strong response in the neuron (Meredith & Stein, 1983; Holmes, 2007, 2009).

These neurophysiological principles have been useful in predicting the circumstances in which multisensory integration occurs in (human) behaviour as well (see for example Spence, 2007 for a review). The application of these principles, however, is not as straightforward as one may think. Especially the spatial rule and the principle of inverse effectiveness need not always to be fulfilled in order for integration to occur (see Holmes, 2007, 2009; see Spence, 2013 for a review). For example, in some cases no differences in multisensory integration were observed between spatially aligned and misaligned stimuli (Girard, Collignon, & Lepore, 2011; Girard, Pelland, Lepore, & Collignon, 2013; Zampini, Torresan, Spence, & Murray, 2007). For audiovisual multisensory facilitation the opposite pattern of inverse effectiveness was demonstrated in terms of RTs (i.e., strongly effective stimuli caused more facilitation relative to stimuli that were only weakly effective; Leone & McCourt, 2013). Even the temporal rule is more flexible than previously thought. It has been shown that what is perceived as being simultaneous can be changed through what is called 'multisensory temporal recalibration' based on prior multisensory experience (e.g., Machulla, Di Luca, Froehlich, & Ernst, 2012; Vroomen & Keetels, 2010). The idea that prior multisensory experience can alter what the brain 'sees' as belonging to the same event or source very much relates to the question of how multisensory neurons (and the principles for integration) develop during life. It seems that the development of multisensory neurons and the conditions under which integration occurs in these neurons is driven by the multisensory experience an organism has with the world (Wallace & Stein, 1997, 2001, 2007).

The observation that the principles of multisensory integration that are based on neurophysiological findings do not always directly relate to behavior suggests that the principles governing multisensory perception are more flexible and task- or situation-dependent in (human) behaviour. Keep in mind, though, that the original principles were based on observations on a much more basic level (i.e., single multisensory neurons), whereas these 'conflicting' findings are based on observations on a behavioural level (e.g., RTs). So, although the original (neurophysiological) principles may need some further refinement in order to apply to (human) multisensory behaviour, they still provide good guidelines for when multisensory integration tends to occur.

Interactions between attention and multisensory integration

Given that both spatial attention and multisensory integration can facilitate sensory information processing, the question arises as to whether and how these two processes interact. Whereas endogenously (voluntarily) attendending multisensory information enhances multisensory integration when spatial uncertainty is low (Talsma & Woldorff, 2005), the benefits of multisensory integration are larger for endogenously unattended locations when the spatial location of a target is uncertain (Zou, Müller, & Shi, 2012). Others have shown that under certain circumstances multisensory integration can occur pre-attentively (Bertelson, Vroomen, de Gelder, & Driver, 2000; Soto-Faraco, Navarra, & Alsius, 2004; Spence & Driver, 2000), but also that some attentional resources are required for multisensory integration to occur (Alsius, Navarra, Campbell, & Soto-Faraco, 2005; Alsius, Navarra, & Soto-Faraco, 2007). To make sense of all these findings it has been suggested that attention and multisensory integration can interact at multiple levels of processing. Depending on at which level multisensory integration takes place ('early' vs. 'late') and the strength of the stimuli, attention is able to enhance multisensory integration, or is even required for integration to occur (Koelewijn, Bronkhorst, & Theeuwes, 2010; Navarra, Alsius, Soto-Faraco, & Spence, 2010; Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010).

Most of the studies on the interaction between attention and multisensory integration have involved the manipulation of *endogenous* attention. Studies of the interaction between crossmodal *exogenous* spatial attention and multisensory integration are lacking. This interaction is especially interesting as both processes are essential for spatial orienting.

While there is little debate about whether crossmodal *endogenous* spatial attention and multisensory integration are separate processes, there is some discussion about whether crossmodal *exogenous* spatial attention and ('early') multisensory integration are different processes (McDonald, Teder-Sälejärvi, & Ward, 2001; Spence, 2010). This discussion may be caused by the fact that both crossmodal exogenous spatial attention and ('early') multisensory integration are bottom-up (stimulus-driven) processes whose benefits are most pronounced when sound and light are presented from approximately the same as compared the different spatial locations. As mentioned above, however, the facilitative effects of crossmodal exogenous spatial attention and multisensory integration have different temporal profiles, which can be used to tell them apart. The effects of crossmodal exogenous spatial attention are most pronounced when there is some delay between the onset of a sound and a light, while the effects of multisensory integration are often most pronounced with close temporal alignment (see chapter 2 for a more detailed discussion of their differences).

Depth: the forgotten dimension

Multisensory interactions have been extensively investigated in the last couple of decades. One aspect that has been receiving little to no attention in the literature about multisensory interactions, however, is how the distance between multisensory information and an observer affect multisensory integration. Almost all studies of multisensory interactions have presented stimuli at a fixed distance from the observer (often on a computer monitor, with LEDs and speakers, or a projection screen). Yet, neurophysiological and neuropsychological evidence indicates that information that is presented at different distances from the observer, or in different regions of space, is processed differently by the brain (Graziano, Reiss, & Gross, 1999; Halligan & Marshall, 1991; Previc, 1998). There seems to be a clear distinction between the region of space directly surrounding different body parts (peri-personal space) and the space that is currently out of reach (extra-personal space) in terms of multisensory interactions occurring in these regions of space. It is remarkable that the distance from which information is presented has been neglected in multisensory research, given that we perceive sounds and lights from varying distances all the time in our daily lives.

Relevance of the used methodologies

The goal of the current thesis was to gain a better understanding of when and how (crossmodal exogenous) spatial attention and multisensory integration facilitate spatial perception. More specifically, it was investigated how these two bottom-up processes interact and how the distance from which information is perceived affects the outcome of multisensory integration and (crossmodal) spatial attention.

The evidence for the conclusions drawn in the current thesis comes from the results of various behavioural paradigms in healthy (chapters 2-5, 7, 9) and clinical populations (stroke patients, chapter 8). We observed that the distance from which information is presented modulates (crossmodal) spatial attention and multisensory integration. The fact that we found distance-related modulations using various methods in different populations indicates how robust such spatial modulations are, and underlines the importance of studying multisensory interactions in different regions of space.

Furthermore, we did not only investigate differences in multisensory benefits relative to unimodal performance between conditions, but also investigated whether there were differences in the underlying causes of these benefits. Given that we were especially interested in interactions between the senses, it was important for us to be able to determine whether any multisensory benefits that we observed were due to multisensory integration or independent processing (i.e., statistical facilitation).

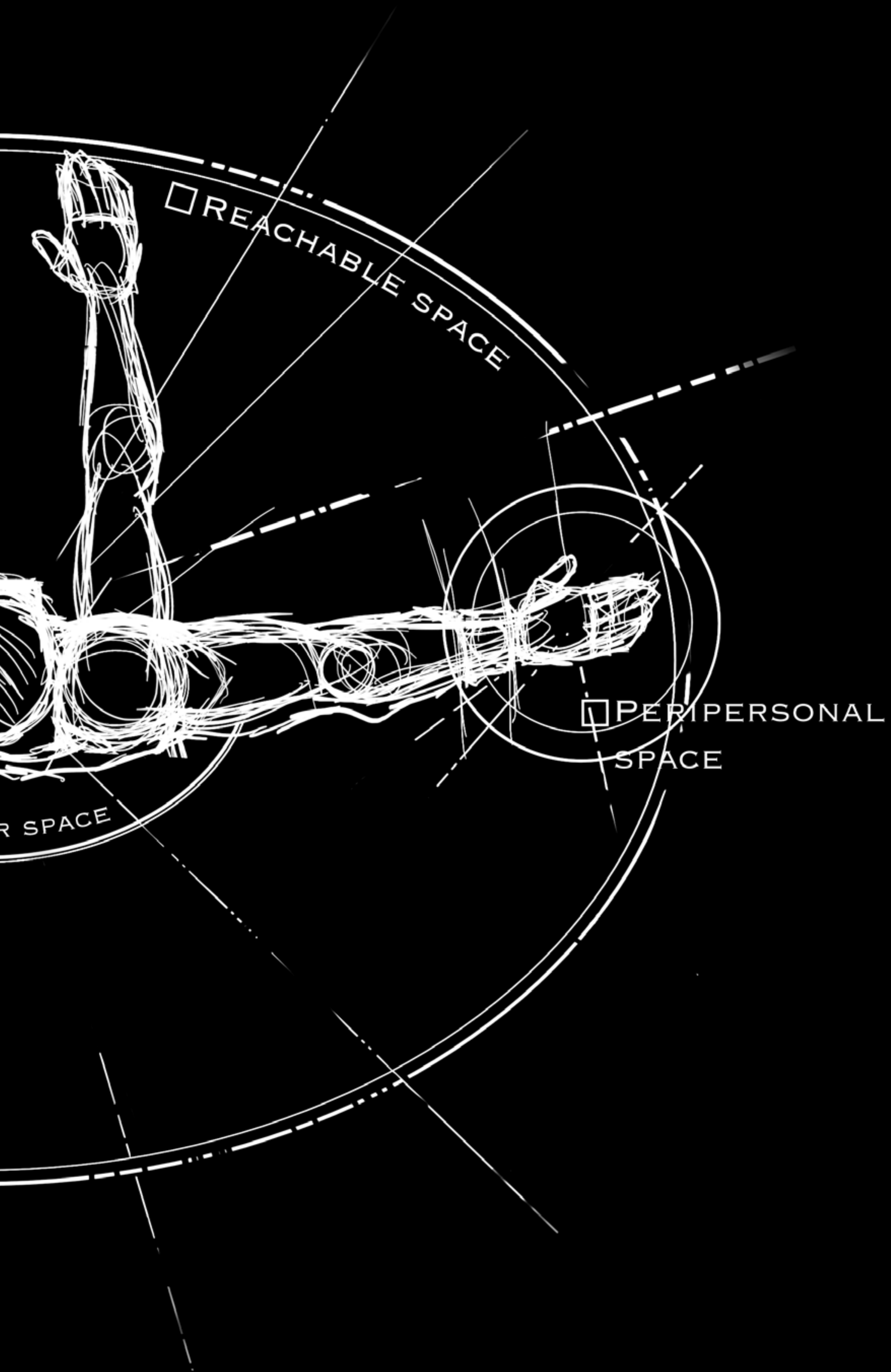
To do so, we compared performance in the multisensory condition to the optimal performance predicted by independent processing (a comparison with the race model; Miller, 1986; Ulrich, Miller, & Schröter, 2007). Whenever the observed performance in the multisensory condition was better than that predicted by independent processing (i.e., race model violation), we could conclude that an interaction between the senses must have occurred (i.e., multisensory integration). The fact that several effects that are not necessarily related to multisensory integration can also contribute to violations of the race model is a valid point of critique on this method (e.g., modality switch effects, Gondan, Lange, Rösler, & Röder, 2004). However, the influence of these effects can be measured and taken into account (see for example chapter 3, Experiment 1), and depend on the specific behavioral paradigm that is used. Strong support for the ability of race model violations to capture true multisensory integration comes from studies that have shown that race model violations are related to intra- (e.g., Molholm et al., 2006) and extracortical (e.g., Gondan, Niederhaus, Rösler, & Röder, 2005; Molholm et al., 2002) measures of multisensory integration.

Thus, this method allowed us to determine whether certain processes were able to modulate multisensory integration by comparing differences in the amount of race model violation between different conditions. Although the absolute benefits of multisensory stimulation can sometimes be similar for different conditions (e.g., the same absolute decrease in RT), the underlying process leading to this benefit can be different (optimal independent processing vs. multisensory integration; see chapter 3). This underlines the importance of comparing multisensory benefits with performance based on independent processing. Ignoring this type of analysis could lead to falls conclusions about whether or not certain factors are able to modulate multisensory integration.

Chapter outline

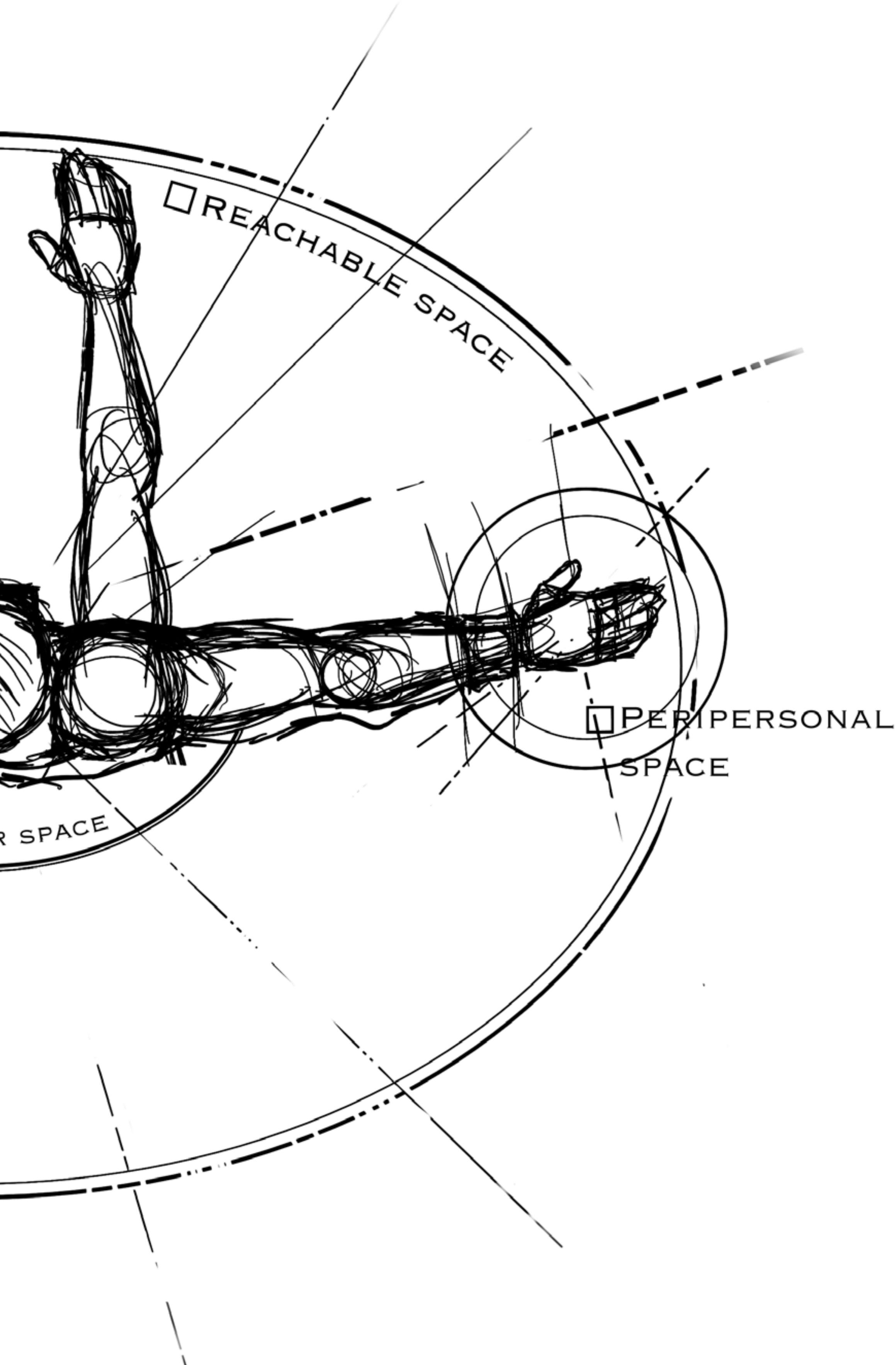
It should be evident by now that there is still a lot to gain in the understanding of crossmodal exogenous spatial attention and multisensory integration. The current thesis is divided into two parts. Part 1 contains four studies that are related to multisensory interactions in a single depth-plane. In chapter 2, we investigated at which stimulus intervals crossmodal exogenous spatial attention and multisensory integration contribute to multisensory response enhancement. Chapter 3 focuses on how exogenous spatial attention affects multisensory integration. The effects of IOR on multisensory integration are discussed in chapter 4. The last study described in the first part of this thesis is concerned with the question how eye-movements are affected by multisensory stimulation when multiple visual elements are presented together with a single sound (chapter 5).

The studies in Part 2 are all related to spatial attention and multisensory interactions in the depth-plane. We start off with providing an overview of how multisensory interactions between vision, audition, and touch are affected by the region of space in which information happens to be presented (chapter 6). We investigated whether crossmodal exogenous spatial attention is sensitive to the distance at which information is presented in chapter 7. Support for the idea that visual spatial attention can be selectively impaired in peripersonal and extrapersonal space after stroke comes from a study in chapter 8. In chapter 9 we investigated how the region of space in which information is presented and distance-related changes in stimulus-efficacy affect multisensory integration. In the last chapter we summarize the results, discuss their theoretical importance, and provide examples of their use in an applied context.


☐Reachable space
☐Peripersonal space
r space

Part I

Crossmodal exogenous spatial attention and multisensory integration


☐ Reachable space
☐ Peripersonal space
r space

Chapter 2

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

Van der Stoep, N., Spence, C., Nijboer, T. C. W., & Van der Stigchel, S.
(accepted pending minor revisions).
On the relative contributions of exogenous spatial attention and multisensory integration to multisensory response enhancement.
Acta Psychologica

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 to obtain 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, 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 visual targets. 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 observed MRE at SOAs of 100 and 200 ms was strictly attributable to crossmodal exogenous spatial attention. These results reveal that the commonly reported 'attentional cueing effects' at an SOA of 50 ms in crossmodal cueing tasks can partly be explained by MSI, in contrast to what is often assumed in the literature.

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; Spence & Driver, 2004; Ward, McDonald, & Lin, 2000). 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).

Both the crossmodal exogenous spatial attention and the multisensory integration accounts have in common the suggestion that the benefits of multisensory stimulation are most pronounced when unimodal components of a multisensory stimulus are spatially aligned (i.e., presented from the same spatial location, the spatial rule; Spence & Driver, 2004; Spence & McDonald, 2004; Leone & McCourt, 2013; though see Spence, 2013, for a review) 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, for example, McDonald, Teder-Sälejärvi, & Ward, 2001; and Spence, 2010, pp. 183-184), given the similarities between crossmodal exogenous spatial attention and multisensory integration. One way 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, Fister, Barnett, Nidiffer, & Wallace, 2012; though see, for example, King & Palmer, 1985) with the behavioral benefits decreasing more or less symmetrically when 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[2]. 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 such a 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). 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,

---

[2] For multisensory integration to occur, it seems especially important that the responses to an auditory and a visual stimulus in a multisensory neuron overlap (e.g., King & Palmer, 1985; Meredith, Nemitz, & Stein, 1987). Relatively small differences in temporal onset are allowed to 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.

whereas unimodal exogenous cues do not (see Spence & Santangelo, 2009, Spence, 2010, 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., orthogonal spatial cueing task, Driver & Spence, 1998; implicit spatial discrimination task, 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) modeling 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 so, 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., Ward, McDonald, & Lin, 2000) or multisensory integration (an RTE task; e.g., Miller, 1986; Stevenson, Krueger Fister, Barnett, Nidiffer, & Wallace, 2012; Laurienti, Burdette, Maldjian, & Wallace, 2006; 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).

Materials and methods

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 their participation 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 hours. 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.

Apparatus and stimuli

A custom-built audiovisual stimulus generator (also see 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 x 120 x 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-coloured LED colour: 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-height 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 of the signal, ~65dB(A)]. The use of tri-colored LEDs 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 cross was a blue light presented with a random duration 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).

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 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 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 with 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 Figure 1A.

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 locations (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. An overview of the timing of all trial types in each of the blocks is shown in Figure 1B. The response window was set to 1500 ms after target onset, followed by the start of the next trial.

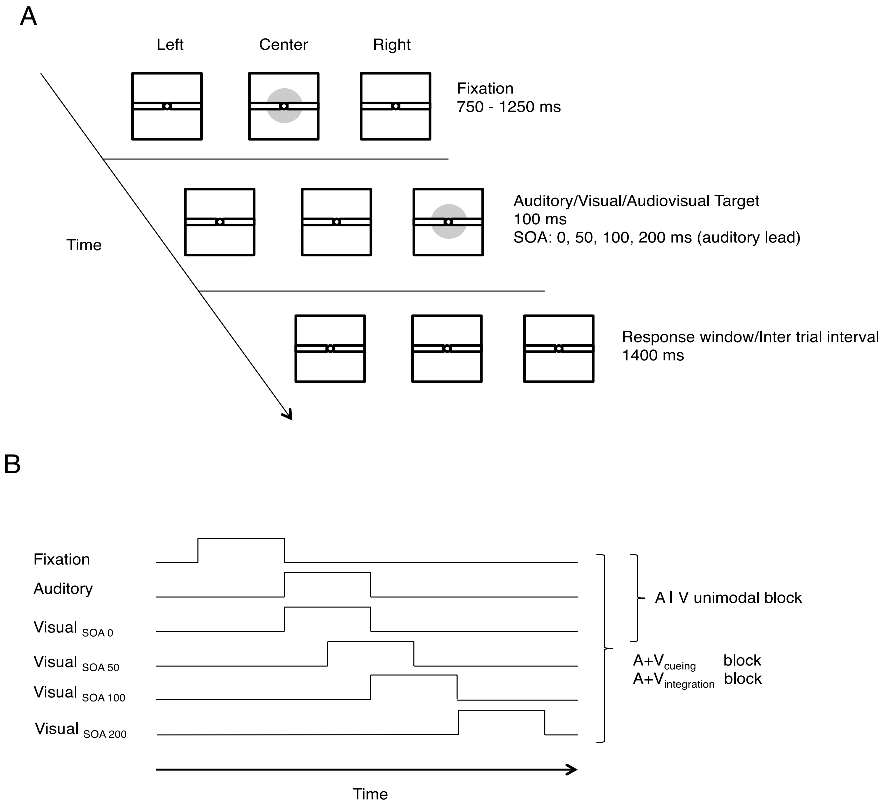


Figure 1. Panel A: Schematic representation of a trial. The grey filled circles indicate the onset of the fixation light and one of three possible locations of the onset of the visual target. Unimodal auditory and visual targets only appeared in the Unimodal block and bimodal targets only in the Integration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

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.

Data analysis

Reaction times (RTs) lower 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 trials) 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.

For the accuracy analysis, both the accuracy on Go and No-go trials were calculated. To analyze detection sensitivity, A was calculated (a non-parametric estimate of sensitivity was used because Hits and False alarms were often 0 and 1; Zhang & Mueller, 2005). The sensitivity was calculated separately for the Unimodal block, and for the spatially aligned and misaligned conditions for each SOA in the Cueing block. In the Integration block, the sensitivity was only calculated for the spatially aligned condition, as there were no spatially misaligned No-go trials in the integration block (which are required to calculate A).

To investigate the amount of speed-up in the multisensory condition 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:

$$\mathrm{MRE} = \min(\mathrm{median}(RT_A), \mathrm{median}(RT_V)) - \mathrm{median}(RT_{AV})$$

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 instructions of 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 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):

$$P(RT_{AV} < t) \leq P(RT_A < t) + P(RT_V + SOA < t)$$ [3]

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, 2007).

Results

Accuracy

Unimodal block

The average accuracy on Go trials in the unimodal task was very high (*M* Visual = .998, *SE* = .001; *M* Auditory = .998, *SE* = .001) and did not differ between A and V target trials ($t(23) = 0.000$, $p = 1.000$). The average accuracy was lower on No-go trials (*M* Visual = .942, *SE* = .022; *M* Auditory = .825, *SE* = .042). The accuracy for V No-go trials was higher compared to A No-go trials. [$t(23) = -3.077$, $p = .005$, $d = -.659$].

The detection sensitivity was slightly higher for V targets (*M* = .985, *SE* = .006) as compared to A targets (*M* = .955, *SE* = .011, $t(23) = -3.165$, $p = .004$, d = -.640). This difference in sensitivity is not problematic as the exact same A and V stimuli were consistently presented together for all manipulations of interest (i.e., the effects of instruction (integration vs. cueing), SOA, and spatial alignment).

Integration block

The repeated measures ANOVA's for accuracy on Go trials (Alignment x SOA), No-go trials (SOA), and sensitivity (SOA) did not indicate any main effects of and interactions

---

[3] The race model inequality can also be expressed as P(RTAV < t) ≤ P(RTA < t) + P(RTV + SOA < t) – P(RTA < t) x P(RTV + 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 the same.

between Alignment and SOA (*F*'s < .2.1, *p*'s > .1). The average accuracy across all SOAs was .993 (*SE* = .005) on Go trials and .863 (*SE* = .016) on No-go trials. The average sensitivity for spatially aligned trials was .962 (*SE* = .005).

Cueing block

A 3 x 4 repeated measures ANOVA with the factors Cue Type (Valid, Invalid, and Center cue) and SOA was used to analyze accuracies in the Cueing block. There was no significant main effect of Cue Type or SOA, and no interaction between Cue Type and SOA (*F*'s < 2.3, *p*'s > .1). The overall accuracy on Go trials was very high (*M* = .993, *SE* = .004).

The average accuracy on the No-go trials was slightly lower compared to the accuracy on Go trials (*M* = .898, *SE* = .016). A 2 x 4 repeated measure ANOVA revealed main effects of Cue Type [$F(1, 23) = 44.609$, $p < .001$, partial $\eta^2 = .660$] and SOA [$F(1.996, 45.901) = 7.034$, $p = .002$, $\varepsilon = .665$, partial $\eta^2 = .234$]. The accuracy was higher for the center cue (*M* = .954, *SE* = .012) than for the invalid cue No-go trials (*M* = .843, *SE* = .023) and the average accuracy on No-go trials decreased as the SOA increased. There was also a significant interaction between Cue Type and SOA [$F(2.236, 51.438) = 3.971$, $p = .021$, $\varepsilon = .745$, partial $\eta^2 = .147$] which could be explained by the fact that in the invalid cue condition, the accuracy on the No-go trials declined as the SOA increased, but remained the same across SOA in the center cue condition.

To analyze the detection sensitivity in the Cueing block a 2 x 4 repeated measures ANOVA with the factors Spatial Alignment (aligned, misaligned) and SOA was used. There was a main effect of spatial alignment on sensitivity [$F(1, 23) = 46.586$, $p < .001$, partial $\eta^2 = .699$]. Sensitivity was higher for spatially aligned (*M* = .986, *SE* = .005) than for spatially misaligned AV trials (*M* = .957, *SE* = .007). In addition, there was a significant main effect of SOA [$F(2.142, 49.277) = 5.990$, $p = .004$, $\varepsilon = .714$, partial $\eta^2 = .207$], indicating that sensitivity decreased as the SOA increased. The interaction between spatial alignment and SOA was also significant [$F(2.215, 50.953) = 3.677$, $p = .028$, $\varepsilon = .738$, partial $\eta^2 = .138$]. The decline in sensitivity as the SOA increased was present in the spatially misaligned, but not in the spatially aligned condition.

In sum, these accuracy analyses indicate that accuracy was slightly lower when participants had to withhold their response to a visual target when the auditory cue and the visual target were spatially and temporally misaligned (No-go trial Cueing block). Nonetheless, the overall accuracy was very high on both Go and No-go trials in the Unimodal, the Integration, and the Cueing block.

Response times

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 Figure 2).

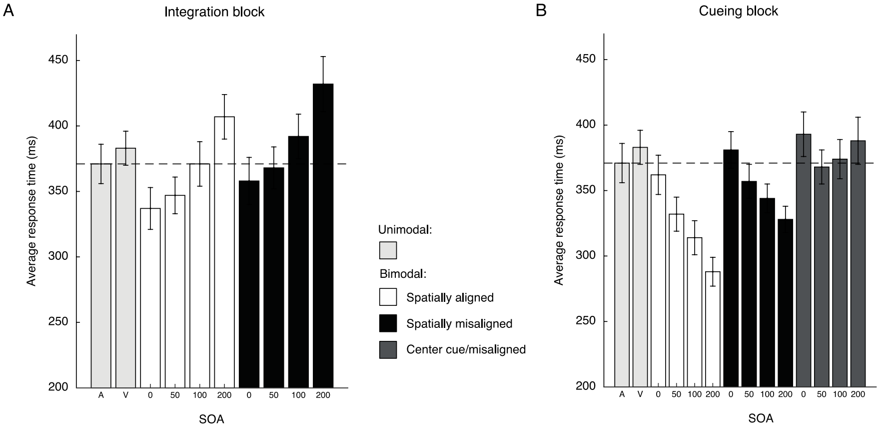


Figure 2. The average RTs in the unimodal conditions (A and V in panel A and B), and the multisensory conditions in the Integration (panel A) and in the Cueing block (panel B) for each SOA (0, 50, 100, 200 ms). The dashed line indicates the shortest average unimodal response time. Error bars represent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

Integration block

RTs in the Integration block were analyzed using a 2 x 4 repeated measure ANOVA with the factors Spatial Alignment (aligned, misaligned) and SOA. There was a 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 Figure 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 at the same location.

Cueing block

A 3 x 4 repeated measure ANOVA with the factors Cue Type (Valid, Invalid, and Center cue) and SOA was used to analyze RTs in 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 that appeared 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 Figure 2B.

RTs decreased as 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 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 SOA increased. The effects of spatial alignment were further explored in a separate analysis, as described below.

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 was observed in the Cueing block: responses were fastest when the A and V targets were spatially aligned but temporally misaligned.

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

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 Figure 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 $< .05$).

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 $> .05$). 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.

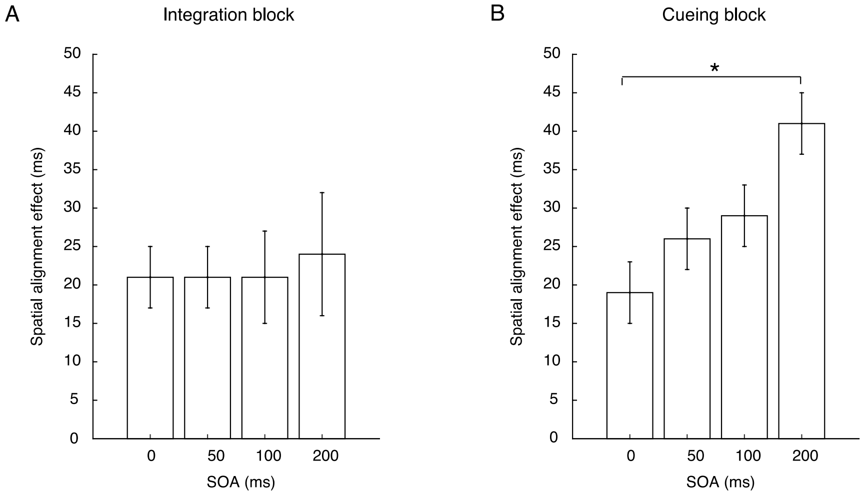


Figure 3. The effect of spatial alignment (i.e., RT AV aligned - RT AV misaligned) for each SOA in the Integration (panel A) and the Cueing block (panel B). Significant differences are indicated with an asterisk ($p < .05$, corrected for multiple comparisons).

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 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, Henik, & Rafal, 2005; Spence & McDonald, 2004).

Multisensory response enhancement

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

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 ($M = -2.167$, $SE = 10.820$) 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 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.

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 condition 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 Figure 3B).

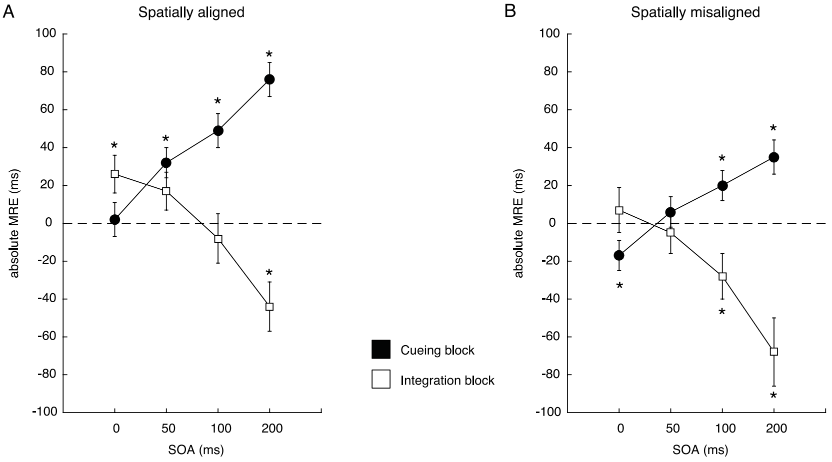


Figure 4. The average absolute multisensory response enhancement (positive values) and multisensory response interference (negative values) for each SOA in the Cueing and the Integration block in the spatially aligned (A) and misaligned (B) condition. Significant MRE (different from zero) is indicated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

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 ($M = 20$ 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, which is in line with the idea that multisensory integration often occurs within a certain temporal binding window (as discussed in Footnote 1, the most important aspect of this temporal rule seems to be the overlap of the unimodal neuronal activity in multisensory neurons; 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 was observed: that is, there was no MRE at 0 ms SOA, but multisensory response enhancement started to emerge and increase from 50 ms SOA up until 200 ms SOA. The 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.

Race model violation

Integration block

To investigate whether the observed multisensory response enhancement could be explained 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 Figure 5 for the average amount of race model violation in the spatially aligned and misaligned at the different SOAs in the Integration and Cueing block.

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 visual targets. Therefore, in the Cueing block, multisensory response enhancement could in part also be the result of alerting, response preparation (e.g., Los & Schut, 2008; see

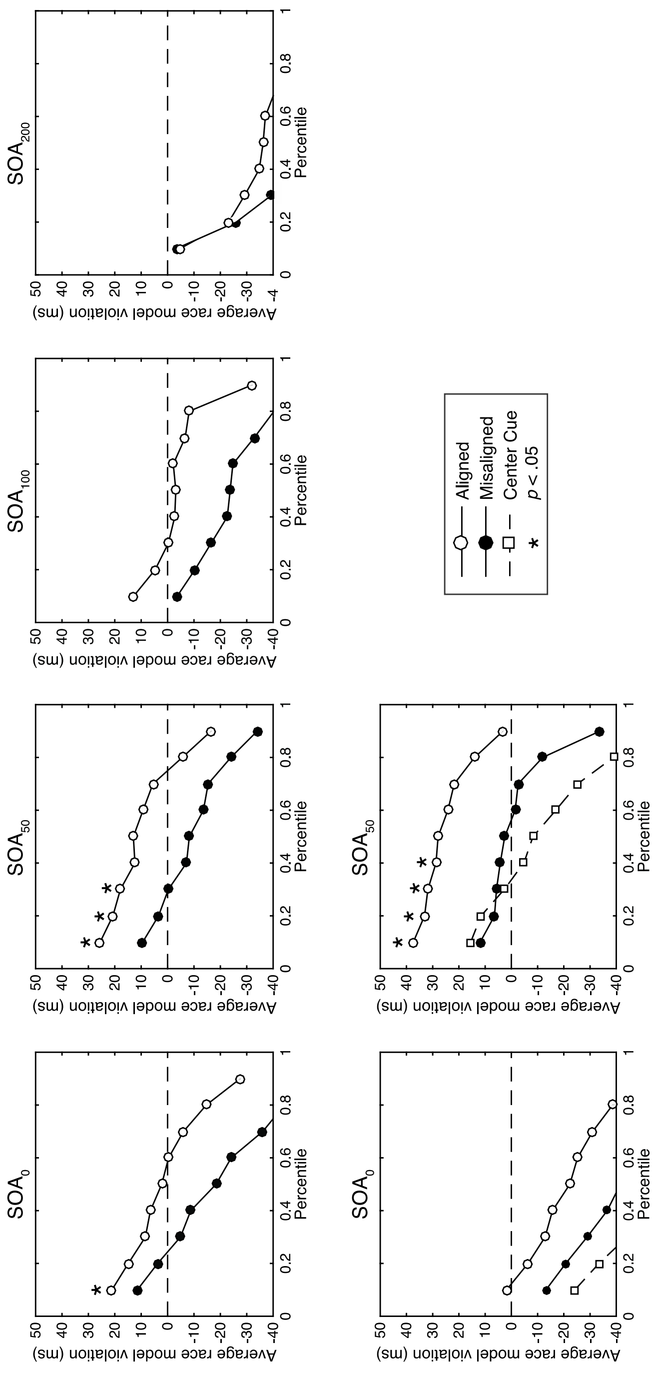


Figure 5. The average amount of race model violation in milliseconds across the 9 quantiles for the SOA_0, SOA_{50}, SOA_{100}, and SOA_{200} condition in the spatially aligned (white dots) and spatially misaligned condition (black dots) in the Integration block (top row) and the SOA_0, SOA_{50} condition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

also Los & Vanderburg, 2013), and/or crossmodal exogenous spatial attention (e.g., Spence & McDonald, 2004; also see Figure 3B and 5A). In the Cueing block, race model violations were only analyzed for the SOA conditions in which there were significant race model violation in the Integration block (the SOA_0 and SOA_{50} condition) because it is assumed that any differences in the amount of race model violation between SOAs in the Integration block could not be the result of additional alerting and preparation as participants could always respond to the first stimulus that was presented. In the Cueing block, the race model was not violated in the SOA_0 condition, but it was significantly violated from the 10th to the 40th percentile in the SOA_{50} condition (all *p*'s < .05).

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, Burdette, Maldjian, & Wallace, 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 at which SOAs integration and crossmodal exogenous spatial attention can 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, 2014; 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.

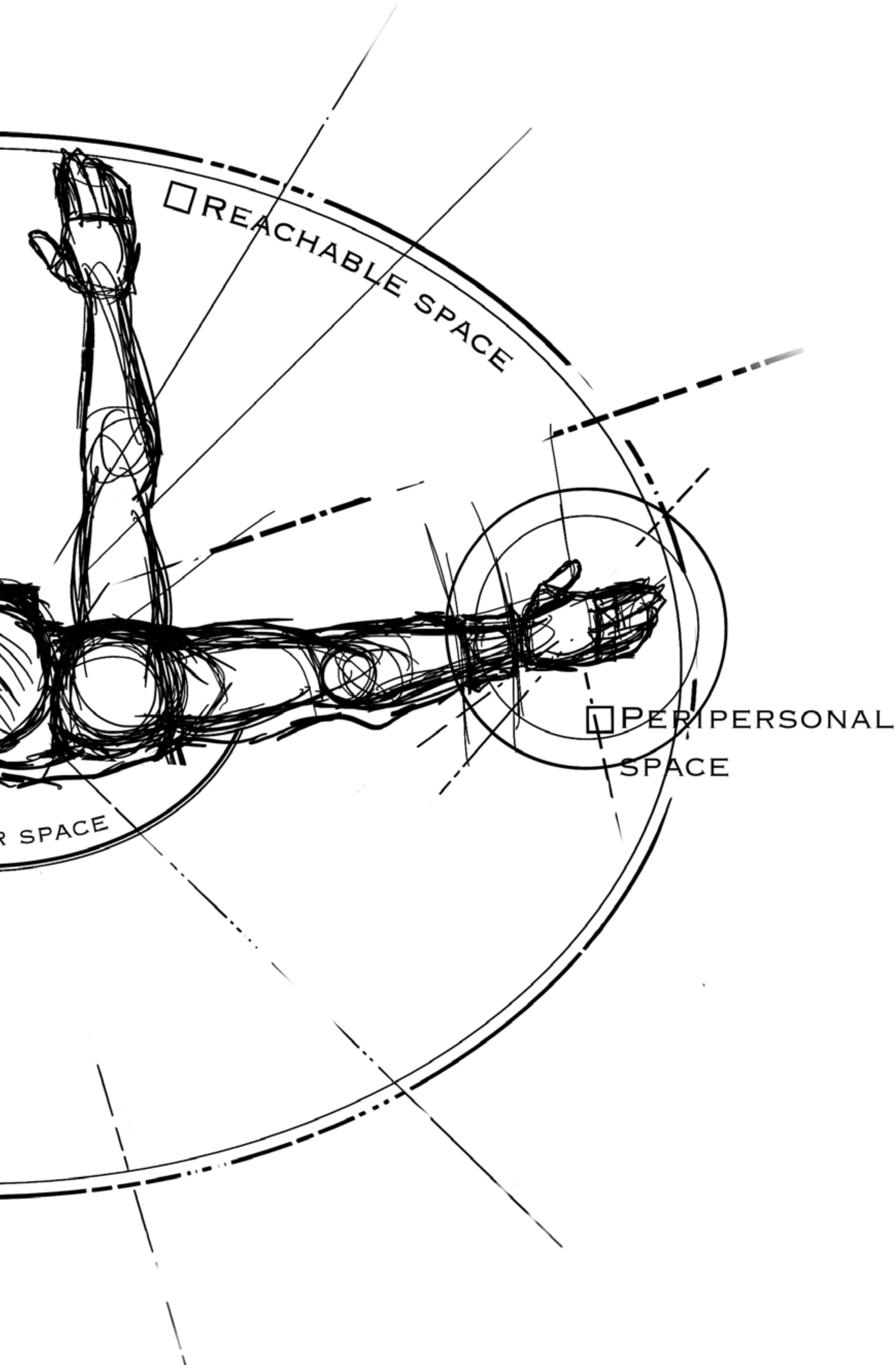
In some cases it is possible that MRE at 0 ms SOA can be accounted for in part by the non-spatial effects of alerting (Diederich & Colonius, 2008) or even in full by response preparation (Los & Van der Burg, 2013). For example, it has been shown that response preparation can speed up responses to multisensory stimuli even at 0 ms SOA, 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).

MRE in the Cueing block could potentially be the result (at least in part) of multisensory integration (see Figure 6). Therefore, it was especially interesting to take a closer look at the amount of MRE in the 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 Figure 3A), the spatial alignment effect increased as the SOA increased in the Cueing block (see Figure 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, Henik, & Rafal, 2005, and Driver & Spence, 1998, for example). It is therefore likely that crossmodal exogenous spatial attention also contributed to the observed MRE at 50 ms SOA in the Integration block. Given that there was no multisensory integration at the 100 and 200 ms SOAs in the Integration block, the observed MRE at the same two SOAs in the Cueing block cannot be explained by multisensory integration. 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 (as 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).

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 block 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. It has been shown that selectively attending to only one sense reduces multisensory integration of sensory information between that and another sensory modality (e.g., Hugenschmidt, Mozolic, & Laurienti, 2009; Talsma, Doty, & Woldorff, 2007; Mozolic, Hugenschmidt, Peiffer, & Laurienti, 2008). It could be that attention was divided across audition and vision in the Integration block and that vision was selectively attended in the Cueing block, thus reducing multisensory integration in the latter. It is possible that because of selective attention to the visual modality, visual stimulus processing was facilitated and auditory stimulus processing reduced (Spence & Parise, 2010). This difference in auditory and visual processing may have reduced multisensory integration when the stimuli were presented synchronously in the Cueing block. The fact that MRE was observed at the 50 ms SOA in the Cueing block seems to contradict this explanation. At the 50 ms SOA, though, 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 which 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 short 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 by 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 audition and vision. Our results indicate that at 50 ms SOA, the amount of MRE is likely the result of 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.


☐ Reachable space
☐ Peripersonal space
r space

Chapter 3

Exogenous spatial attention decreases audiovisual integration

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.

Abstract

Multisensory integration (MSI) and spatial attention are both mechanisms through which the processing of sensory information can be facilitated. Studies on the interaction between spatial attention and MSI have mainly focused on the interaction between endogenous spatial attention and MSI. Most of these studies have shown that *endogenously* attending a multisensory target enhances MSI. It is currently unclear, however, whether and how *exogenous* spatial attenttion and MSI interact. In the current study, we investigated the interaction between these two important bottom-up processes in two experiments. In Experiment 1 the target location was task-relevant, and in Experiment 2 the target location was task-irrelevant. Valid or invalid exogenous auditory cues were presented before the onset of unimodal auditory, unimodal visual, and audiovisual targets. We observed reliable cueing effects and multisensory response enhancement in both experiments. To examine whether audiovisual integration was influenced by exogenous spatial attention, the amount of race model violation was compared between exogenously attended and unattended targets. In both Experiment 1 and Experiment 2, a decrease in MSI was observed when audiovisual targets were exogenously attended, compared to when they were not. The interaction between exogenous attention and MSI was less pronounced in Experiment 2. Therefore, our results indicate that exogenous attention diminishes MSI when spatial orienting is relevant. The results are discussed in terms of models of multisensory integration and attention.

Introduction

Two processes that are involved in the interaction between information from different senses are multisensory integration (MSI) and crossmodal attention. Both MSI and (crossmodal) attention are able to facilitate the speed of detection, and the accuracy of localization and identification of targets (e.g., Leo, Bologinini, Passamonti, Stein, & Ladavas, 2008; Montagna, Pestilli, & Carrasco, 2009; Spence & Driver, 1997; Stevenson, Krueger Fister, Barnett, Nidiffer, & Wallace, 2012). To date, however, it is unclear under what circumstances and how these two processes interact. Although some studies have found that MSI occurs independent of whether attention has been allocated to the multisensory stimulus (e.g., Bertelson, Vroomen, de Gelder, & Driver, 2000; Bertelson, Pavani, Ladavas, Vroomen, & de Gelder, 2000; Soto-Faraco, Navarra, & Alsius, 2004; Vroomen, Bertelson, & de Gelder, 2001), other studies have shown that attention is able to modulate MSI (e.g., Alsius, Navarra, & Soto-Faraco, 2007; Fairhall & Macaluso, 2009; Talsma & Woldorff, 2005; Talsma, Doty, & Woldorff, 2007). To explain these different findings, it has been suggested that the influence of attention on MSI depends on several factors such as the type of task (e.g., detection vs. identification), the stimulus properties (e.g., salient vs. near threshold, simple vs. complex), and the attentional resources that are available (e.g., low attentional load vs. high attention load; exogenous vs. endogenous attention manipulation; for reviews see Koelewijn, Bronkhorst, & Theeuwes, 2010; Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010).

Interestingly, studies of the interaction between attention and MSI have mainly looked into the influence of endogenous attention on MSI. The results from these studies indicate that endogenous attention influences MSI, either by showing that it enhances MSI (e.g., Talsma & Woldorff, 2005; Fairhall & Macaluso, 2009; but see Bertelson, Vroomen, de Gelder, & Driver, 2000), that MSI is decreased when attentional resources are limited or depleted (Alsius, Navarra, Campbell, & Soto-Faraco, 2005; Alsius, Navarra, & Soto-Faraco, 2007), and that endogenous spatial attention spreads across the unimodal components of a multisensory stimulus (Busse et al., 2005). In contrast with these findings, Zou, Müller, and Shi (2012) observed larger benefits of multisensory stimulation on visual target detection in the endogenously unattended half of the stimulus display as compared to the endogenously attended half of the display (as evident from a larger pip-and-pop effect at the unattended side). A factor that might have contributed to these different findings is the degree of spatial uncertainty of the target in the task. For example, when the target location was not varied on a trial-by-trial basis, attention was shown to enhance multisensory integration for simple (e.g., Talsma & Woldorff, 2005) and complex stimuli (e.g., Alsius et al., 2005, 2007; Fairhall & Macaluso, 2009).

Differences in the specific task requirements and stimulus properties aside (e.g., see Navarra, Alsius, Soto-Faraco, & Spence, 2010, for a discussion), the majority of these studies have at least shown that MSI can be influenced by endogenous forms of attention. Much less is known about the effects of exogenous spatial attention on MSI. In fact, there is some debate about whether crossmodal exogenous spatial attention and MSI are actually different processes or the same process (e.g., Macaluso, Frith, & Driver, 2000; McDonald, Teder-Sälejärvi, & Ward, 2001; Spence, 2010, pp. 183-184). Some researchers have suggested that exogenous spatial attention and multisensory integration can be discriminated, for example, based on the time-course of their effects. The effects of exogenous spatial attention are often most pronounced with cue target onset asynchronies (CTOA) of 50-200 ms (e.g., Berger, Henik, & Rafal, 2005; Spence & Driver, 1994), whereas multisensory integration is typically most pronounced for stimuli that are presented in close temporal proximity (SOAs < ~100 ms, e.g., Leone & McCourt, 2013; Meredith, Nemitz, & Stein, 1987; Stevenson, Krueger Fister, Barnett, Nidiffer, & Wallace, 2012). Observations of a temporal binding window (TBW) in behavioral studies of temporal order judgment and simultaneity judgment are in line with such a window of integration (~100 visual lead and ~60 auditory lead, but note that the width of this window is task and stimulus dependent, see for example Hirsh & Sherrick, 1961; Keetels & Vroomen, 2005; Stevenson & Wallace, 2013; Vroomen & Keetels, 2010; Zampini, Guest, Shore, & Spence, 2005). Whereas a distinction between crossmodal exogenous attention and multisensory integration based on differences in temporal properties seems to hold well based on the behavioural findings in the literature, others have argued that this distinction is somewhat problematic in terms of the underlying neural interactions given that some researchers have reported observations of multisensory integration in multisensory neurons with SOAs larger than 100–200 ms (see McDonald, Teder-Sälejärvi, & Ward, 2001, for a discussion).

Whereas there is little research on whether exogenous spatial attention influences multisensory integration, there are several studies of whether the benefits of spatial attention shifts that are evoked by multisensory exogenous spatial cues are any different from unimodal exogenous cues (i.e., the effect of multisensory integration on exogenous spatial attention; e.g., Santangelo, Van der Lubbe, Belardinelli, & Postma, 2006; Santangelo, Ho, & Spence, 2008; Santangelo, Van der Lubbe, Belardinelli, and Postma 2008). Under low cognitive load, the size of the cueing effect (i.e., response times (RTs) validly cued targets < RTs invalidly cued targets) does not differ between multisensory and unimodal exogenous spatial cues. Under high cognitive load (i.e., while performing a secondary task), however, multisensory exogenous cues are the only cues that are able to evoke a cueing effect whereas unimodal cues do not evoke a significant cueing effect anymore (Santangelo, Ho, & Spence, 2008; see Spence & Santangelo, 2009, for a review). Other studies on multisensory processing and attention, which have mainly focused on temporal stimulus properties, have also observed benefits of multisensory stimulation, but in the detection of visual targets embedded in a complex visual environment (e.g., the freezing phenomenon, Vroomen & De Gelder, 2000; the pip and pop effect, Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008; Van der Burg, Talsma, Olivers, Hickey, & Theeuwes, 2011; see Ngo & Spence, 2010, for influences of both spatial and temporal alignment on multisensory enhancement of visual search). In the study by Van der Burg et al. (2011) the benefits of multisensory stimulation were most pronounced when the auditory cue and the visual target were presented simultaneously. These benefits could not be explained by a general alerting effect or by shifts of spatial attention (the sound was not lateralized, but presented to the left and right ear through headphones). These findings indicate that the observed benefits were probably due to the attention attracting effects of the integration of auditory and visual stimuli and are in line with a distinction between crossmodal attention and multisensory integration based on time differences. Although these studies provide insight into the influence of multisensory integration on possible shifts of exogenous spatial attention (and some also on whether MSI can occur pre-attentively, e.g., Soto-Faraco, Navarra, & Alsius, 2004; Spence & Driver, 2000; Vroomen, Bertelson, and De Gelder, 2001), they do not provide information on the influence of exogenous spatial attention on the integration of information from multiple senses at exogenously attended compared to exogenously unattended locations.

One study that did look into the influence of exogenous spatial attention on MSI was performed by Vroomen, Bertelson, and De Gelder (2001). In their study they investigated whether the ventriloquist effect (which is thought to be the result of MSI) was affected by the direction of exogenous spatial attention. They did not find such an influence of exogenous spatial attention, suggesting that exogenous spatial attention is not able to influence MSI. Yet, the conditions for observing an influence of exogenous spatial attention on MSI may not have been optimal in their study. Their main experiment consisted of the simultaneous presentation of four squares and a sound that had to be localized by participants ("was the sound coming from the left or the right?"). One of the squares was smaller than the other three, which caused it to act as a singleton and therefore automatically attracted attention. The exogenous cue (i.e., a singleton) was presented simultaneously with the auditory stimulus and was part of the bimodal stimulus. Their results showed that the ventriloquist effect did not depend on the direction of exogenous spatial attention. Several studies, however, have shown that it takes some time for exogenous spatial attention to develop its strongest effect (e.g., 100–300 ms, see Driver & Spence, 1998 for a review on exogenous and endogenous crossmodal spatial attention). Therefore, the onset of the integration process and the shift of exogenous spatial attention may have been temporally misaligned. This may have resulted in preattentive MSI (i.e., no influence of exogenous spatial attention on MSI), just because there was not enough time for exogenous attention to be shifted to the location of the cue.

In order to determine whether exogenous spatial attention is able to influence MSI, we investigated this interaction in a situation in which the exogenous cue was not only presented prior to the multisensory stimulus but was also not part of the multisensory target, using simple stimuli (sounds and light disks). This ensured that there was enough time for exogenous spatial attention to be allocated to the location of the multisensory stimulus. Additionally, as the stimulus that is causing an exogenous shift of spatial attention is different from the stimuli that need to be integrated, exogenous orienting of attention and MSI do not depend on the same stimulus, providing an opportunity for both processes to emerge and for exogenous spatial attention to influence MSI. Although in several studies it was shown that endogenous attention increases multisensory integration, there are also some studies that seem to suggest that multisensory integration is enhanced when the location of the multisensory stimulus is unattended. For example, multisensory integration is less affected by the depletion of attentional resources (e.g., Santangelo, Ho, & Spence, 2008) and multisensory benefits during visual search are larger at endogenously unattended regions of space (Zou, Müller, & Shi, 2012).

We hypothesized that if the interaction between exogenous attention and MSI depends on whether there was enough time for exogenous spatial attention to be allocated, then an interaction between exogenous spatial attention and MSI should be observed if the exogenous cue is presented slightly before the audiovisual stimulus. If this is indeed the case, two different results may be expected based on previous studies. On the one hand, exogenous attention may show an enhancement of MSI just as endogenous attention has been shown to enhance MSI (e.g.,Ngo & Spence, 2010; Talsma & Woldorff, 2005). That is, multisensory integration is enhanced at exogenously attended as compared to exogenously unattended target locations. On the other hand, one might expect multisensory integration to be more pronounced at exogenously unattended locations, as several studies have shown that multisensory integration is more pronounced at endogenously unattended locations (e.g., Zou, Müller, & Shi, 2012) and that multisensory cues are less affected by the depletion of attentional resources as compared to unimodal cues (e.g., Santangelo, Ho, & Spence, 2008, Spence & Santangelo, 2009).

Experiment 1

Materials and methods

Participants

Sixteen healthy participants were tested in the experiment (seven male, nine female, mean age = 25.70 years, *SD* = 3.08). This sample size was based on previous studies on exogenous spatial attention and on studies on the interaction between exogenous and endogenous attention and MSI in which sample sizes varied between 12 and 20 participants (e.g., Spence & Driver, 1997; Talsma & Woldorff, 2005; Vroomen, Bertelson, & De Gelder, 2001). Participants had normal or corrected-to-normal visual acuity, did not report any hearing problems, and received course credits for their participation. The experiment was conducted in accordance with the Declaration of Helsinki, and participants signed informed consent before the start of the experiment.

Apparatus

An Acer X1261P projector (60 Hz) was used to project visual stimuli on a black projection screen (50 × 75 cm). The screen was placed at 87 cm in front of the participant, whose head was placed in a chin-rest. Three speakers (Harman/Kardon HK206, Frequency response: 90–20,000 Hz) were used to present the auditory stimuli.

Stimuli, task, and procedure

The participants had to detect visual, auditory, and audiovisual targets to the left or the right of a fixation cross (black plus sign, 0.8° × 0.8°, 0.9 cd/m^2 as measured with a PhotoResearch SpectraScan PR 650 spectrometer, Weber contrast –1.0) with and without the prior presentation of an exogenous auditory cue. Target modality (visual, auditory, audiovisual) was randomized and we used the implicit cueing paradigm to make space relevant for the task to increase the possibility of finding a validity effect of the exogenous spatial cue (Ward, McDonald, & Lin, 2000). There were three possible cue and target locations: left peripheral, center, and right peripheral. Cues and targets were presented at the three locations each with equal probability: 33% left, 33% center, and 33% right. Participants were instructed to keep fixating on the fixation cross throughout the experiment and to press a button as quickly as possible when a target was presented to the left or the right of the central fixation cross (i.e., a Go trial), but not when a target was presented at the center location (i.e., a No-go trial). There was only one button to respond to the presence of a target, avoiding response-priming effects while the spatial location of the target still remained relevant for the task. Response-priming effects typically occur when participants have to indicate a side where a target is being presented using two buttons: one button for targets presented on the left, and another button for targets presented on the right. If cues and targets are presented from the left and the right side, then faster responses to validly cued targets as compared to invalidly cued targets can also be explained by the priming of the side that should be responded to and/or the hand that should be responded with. Here, we avoided response priming by using only one button to respond (e.g., Ward, McDonald, & Lin, 2000).

The experiment consisted of two blocks: one block without auditory exogenous cues and one block with auditory exogenous cues. Response times in the block without cues acted as a 'baseline' measure of audiovisual integration, and were always presented first. The second block was always the block in which cues were present. A direct comparison of MSI between the No Cue and Cue conditions would be problematic because of possible differences in training and/or fatigue, and the lack of an alerting signal and a temporal warning signal in the No Cue condition. The comparison of MSI in the No Cue block and the Cue block was not the focus of the current study, but rather a comparison of MSI between the Valid Cue and the Invalid Cue condition. We did incorporate the No Cue block in order to know whether any race model violations could be observed using the current stimulus parameters and paradigm. In the second block, cues were present during each trial and were presented with equal probability at the three locations (left, center, right).

Each block's visual, auditory, and audiovisual targets were randomly presented at one of three locations with equal probability either without (during the first block) or with the prior presentation of an auditory exogenous cue (during the second block). The first block (No Cue) consisted of 120 go trials (40 auditory, 40 visual, and 40 audiovisual go trials) and 60 No-go trials (20 auditory, 20 visual, and 20 audiovisual No-go trials). The second block (cue present) consisted of 360 go trials: 120 Valid Cue trials (cue presented at the same location as the target; 40 auditory, 40 visual, and 40 audiovisual targets), 120 Invalid Cue trials (cue presented at the opposite location as the target; 40 auditory, 40 visual, and 40 audiovisual targets), and 120 Center Cue trials (cue presented at the central location; 40 auditory, 40 visual, and 40 audiovisual targets). There were 180 No-go trials in the second block, containing 60 left cue trials, 60 center cue trials, and 60 right cue trials, with each Cue Type containing 20 auditory, 20 visual, and 20 audiovisual target stimuli that were presented at the center location.

Targets and cues were presented at eye-height at 13.6° degrees of visual angle to the left and the right of the fixation cross, and at the location of the fixation cross (directly in front of the participant). Visual targets consisted of a white filled circle (2.8° × 2.8°, 245 cd/m^2, Weber contrast 1.47) and were presented on a grey background (99.2 cd/m^2). Auditory targets consisted of a 100-ms white noise burst of ~70 dB(A) SPL (as measured with an audiometer from the location of the participant; 15-ms rise and fall of the signal). Audiovisual targets consisted of a combination of the auditory and visual stimulus and were always spatially and temporally aligned (timing was confirmed with an oscilloscope). Auditory cues consisted of a 75-ms, 2000-Hz sine wave of ~78 dB (15-ms rise and fall of the signal) that were presented with random cue target onset asynchronies between 200 and 250 ms (with steps of 16.7 ms). Each speaker was placed directly behind the projection screen and the center of each speaker was horizontally and vertically aligned with the location of projected visual targets. A schematic top view of the experimental setup is shown in the right panel of Figure 1. The setup was placed in the center of the room in order to keep the auditory reflections as similar as possible for sounds presented to the left and to the right of the fixation cross.

Each trial started with the presentation of a fixation cross with a random duration between 750 and 1250 ms. After this, the fixation cross disappeared and the auditory cue was presented for 75 ms, followed by the presentation of a 100 ms auditory, visual, or audiovisual target after a random cue target onset asynchrony between 200 and 250 ms. Participants were able to respond until the end of the inter-trial interval (ITI) of 1900 ms starting at target offset. The ITI consisted of the background only. A schematic overview of the procedure is shown in the left panel of Figure 1.

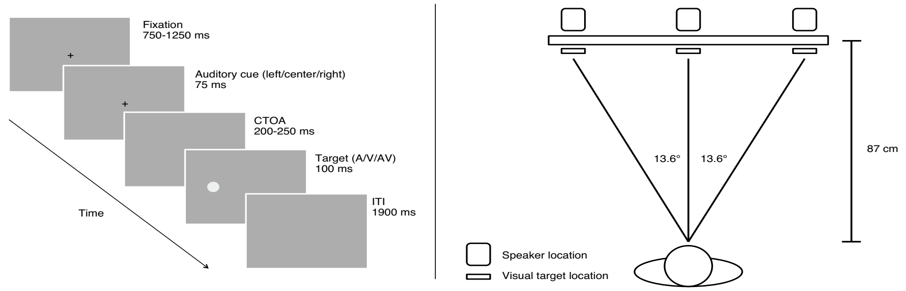


Figure 1. A schematic overview of the procedure (left panel) and a schematic top view of the setup (right panel).

Data analysis

Pre-processing

Incorrect trials and trials with response times shorter than 100 ms or longer than 1000 ms were removed from further analysis as they were assumed to be the result of either anticipation or not paying attention to the task, respectively. This led to the removal of 2.9% of the data. In the No Cue condition 0.2% of the Go trials (anticipations and slow responses) and 7.8% of the No-go trials (False Alarms) were removed, and 0.5% of the Go trials (anticipations and slow responses) and 7.9% of the No-go trials (False Alarms) in the cued conditions. One participant was removed from further analyses and replaced with a new participant because of low accuracy on catch trials in the No-go No Cue condition (60% accuracy). Median response times of each participant in each condition were used for further analyses.

For each cue type (No Cue, Valid, Invalid, and Center Cue), absolute Multisensory Response Enhancement (MRE) was calculated by subtracting the audiovisual median response time from the fastest unimodal median response time from the same cue type for each participant. The resulting values reflect the absolute amount of speed-up or slowing down in milliseconds in the audiovisual condition compared to the fastest unimodal median response time for each cue type and for each participant.

In order to test whether possible MRE in each condition could be explained by statistical facilitation (i.e., independent processing) or by MSI, we used the individual cumulative distributive functions (CDFs) of each Target Modality for each cue type to calculate the race model using the equality (Raab, 1962)[4]:

---

[4] This form of the race model equality is used in several recent studies on multisensory integration in which RTs to audiovisual stimuli are compared to RTs predicted by the race model (e.g., Molholm, et al., 2006; Stevenson, Krueger, Fister, Barnett, Nidiffer, & Wallace, 2012). The equality, however, can also be expressed as $P(RT_{\text{Race Model}}<t)$ =

$$P(RT_{\text{Race model}}) = P(RT_A < t) + P(RT_V < t) - P(RT_A < t) \times P(RT_V < t)$$

The race model provides the probability (P) of a RT that is less than a given time t in milliseconds, where t ranges from 100–1000 ms after stimulus onset. The race model is based on the combination of the unimodal auditory and unimodal visual CDFs. We compared the observed RTs of the audiovisual CDF of each participant of each cue type to its corresponding race model (e.g., Valid Cue audiovisual CDF vs. Valid Cue race model) at the 10th, 20th, to the 90th percentile of each CDF to test for race model violations (Miller, 1982; Ulrich, Miller, & Schöter, 2007; the resulting p-values were Bonferroni corrected, see Statistical analysis). Significant violations of the race model (i.e., $RT_{AV} < RT_{\text{Race model}}$) indicate multisensory interactions that exceed statistical facilitation.

There are multiple ways in which the race model can be compared to the audiovisual CDFs. One possibility is to compare differences in RT at each quantile, as we described above. Another way to compare the CDFs is to look at differences in probability at each RT for the full function (not just taking the values of certain quantiles; see Laurienti, Burdette, Maldjian, & Wallace, 2006, Figure 1B, for example). Using the latter option, subtracting the race model function from the audiovisual function results in a difference function showing exactly in which RT range the race model was violated.

Statistical analysis

Overall accuracy was very high regardless of cue type (see Results section). The accuracy for the different target modalities (A, V, AV) was compared between cue types for go-trials (No Cue, Valid, Invalid, Center Cue) and for No-go trials (No Cue, Cued, Uncued). We performed a 3 × 4 repeated measures ANOVA with the factors Target Modality (Auditory, Visual, Audiovisual) and Cue Type (No Cue, Valid, Invalid, Center Cue) for go trials, and for No-go trials a 3 × 3 repeated measures ANOVA with the factors Target Modality (A, V, AV) and Cue Type (No Cue, Cued, Uncued) was used to analyze accuracy. To further explore possible differences in detection performance for the different target modalities and cue types, A was calculated as a measure of sensitivity based on the hits and false alarms in each condition (Zang & Mueller, 2005; the non-parametric measure A was used as a measure for sensitivity rather than d-prime because the mean accuracy for most participants was either 1 or 0). A 3 × 3 repeated measures ANOVA with

---

$P(RT_A < t) + P(RT_V < t)$, leaving the second part out of the equation under the assumption that responses to signals on different channels compete for resources and may therefore be negatively correlated (see Miller, 1982). Given that both are being used in the literature, we also analyzed our data using the latter formula, which provided qualitatively the same results.

the factors Target Modality (A, V, AV) and Cue Type (No Cue, Cued, Uncued) was used to analyze possible differences in sensitivity. A distinction was made between Cued and Uncued trials instead of maintaining the Valid, Invalid, and Center Cue grouping for this analysis, because targets that were presented at the center could only be cued (Center Cue No-go trial) or uncued (No-go trial with a cue presented from the left or the right of fixation). The Go and No-go trials had to contain the same cue conditions to be able to calculate sensitivity. Therefore, in addition to a Valid Cue Go trial condition, the Invalid and Center Cue Go trials were included in an Uncued Go trial condition.

To analyze response times, a 3 × 4 repeated measures Analysis of Variance (ANOVA) was used with the factors Target Modality (Auditory, Visual, Audiovisual) and Cue Type (Valid, Invalid, Center, and No Cue). Planned pairwise comparisons between the levels of Cue Type were performed to investigate whether cueing effects were present for each Target Modality (e.g., Valid Cue RTs < Invalid Cue RTs for visual targets). We also used pairwise comparisons to investigate multisensory facilitation for each Cue Type (e.g., RT Valid Cue AV vs. RT Valid Cue V or A).

One-sample t-tests were used to test for the presence of Multisensory Response Enhancement for each Cue Type. A repeated measures ANOVA with the factor Cue Type (No Cue, Valid, Invalid, Center) was used to test for overall differences in MRE between the different Cue Types. Planned pairwise comparisons were used to look at the differences in MRE between pairs of Cues.

Race model violations (differences in ms for each quantile) were analyzed using paired samples t-tests for each quantile. The resulting *p*-values were Bonferroni corrected for the number of tests within a condition (N = 9 as there were nine quantiles) using the formula: $p_{corrected} = 1 - (1 - p)^n$ (Motulsky, 1995). The second type of race model violation was analyzed using one-sample t-tests on the difference in probability between the audiovisual and the race model CDF for each ms from 0 to 1000 ms and in each condition (*p*-values were not Bonferroni corrected, but only violations across more than 50 consecutive RTs were reported). If the race model is violated, this would indicate MSI. To test for differences in the amount of race model violation between cue types the median amount of race model violation across the nine percentile points of the CDF of each cue type of each participant was used in a repeated measures ANOVA with the factor Cue Type (No Cue, Valid, Invalid, and Center Cue), followed by planned pairwise comparisons.

The positive area under the difference curve (i.e., the difference in probability of the AV CDF and the race model CDF for the RT range of 0 to 1000 ms) was also used to investigate differences in race model violation between cue types. In order to extract the positive area under the curve for each participant, the difference curve between the AV CDF and the race model CDF was calculated for each participant. Next, all negative probabilities (no race model violation) were set to a value of zero and only the positive area under the curve was calculated for all participants. A repeated measures ANOVA with the factor Cue Type (No Cue, Valid, Invalid, Center) was used to test for differences in the positive area under the curve, followed by planned pairwise comparisons.

In each analysis the degrees of freedom were corrected using the Greenhouse-Geisser correction when necessary.

Results

Accuracy

Overall accuracy on Go trials was very high (average accuracy on Go trials across all cue types, $M = .996$, minimum accuracy = .980, maximum accuracy = 1.00). There was a main effect of Target Modality on Go trials [$F(1.063, 15.949) = 6.057$, $p = .024$, $\varepsilon = .532$, partial $\eta^2 = .288$]. Bonferroni corrected pairwise comparisons between target modalities revealed a significant difference between accuracy for Auditory ($M = 0.99$, $SE = .005$) and Audiovisual Go trials ($M = 1.00$, $SE = 0.00$, $t(15) = -2.043$, $p = .047$), but not between Visual ($M = 1.00$, $SE = .001$) and Audiovisual, nor Visual and Auditory go trials (all p's > .1). There was no main effect for Cue Type and no interaction between Target Modality and Cue Type (all p's > .1), indicating that there was no difference in the amount of anticipations (RTs <100 ms) and misses (no response or RTs >1000 ms) between the different cue conditions and target modalities.

For the No-go trials, there was a main effect of Target Modality [$F(2, 30) = 7.341$, $p = .003$, partial $\eta^2 = .329$]. Participants were significantly better in withholding their response on Visual No-go trials ($M = .960$, $SE = .008$) as compared to Auditory No-go trials ($M = .896$, $SE = .019$, $t(15) = -4.606$, $p = .001$). There were no significant differences between Auditory and Audiovisual No-go trials ($M = .916$, $SE = .017$, $t(15) = -0.979$, $p = .761$) and between Visual and Audiovisual No-go trials [$t(15) = -2.686$, $p = .050$]. There was no main effect of Cue Type and there was no interaction between Cue Type and Target Modality for accuracy on No-go trials (all p's > .05), indicating that the number of false alarms did not differ between the different cue conditions and target modalities.

The analysis of sensitivity (A) revealed a main effect of Target Modality [$F(2, 30) = 9.764$, $p = .001$, partial $\eta^2 = 394$]. Sensitivity for Visual targets was higher ($M = .990$, $SE = .002$) compared to Auditory ($M = .970$, $SE = .005$; $t(15) = -4.502$, $p = .001$), but comparable when compared to Audiovisual targets (*M* = .979, *SE* = .004, *t*(15) = 2.538, *p* = .067). There was no difference in sensitivity between Auditory and Audiovisual targets [*t*(15) = –1.879, *p* = .221]. The main effect of Cue Type [*F*(2, 30) = 1.962), *p* = .158, partial η^2 = .116] and the interaction between Target Modality and Cue Type were not significant [*F*(4, 60) = .414, *p* = .798, partial η^2 = .027], indicating that the sensitivity for detecting targets did not depend on the type of cue and the Target Modality.

Response times

A significant main effect of Cue Type was found (Valid, Invalid, Center, and No Cue; *F*(1.575, 23.618) = 50.973, *p* < .001, ε = .525, partial η^2 = .773). Figure 2 shows the mean RTs for each Target Modality for each cue type. Only the validity effects are indicated with an asterisk to keep the figure clear. Pairwise comparisons indicated that the presentation of a cue resulted in faster responses compared to the No Cue condition (*M* =375 ms, *SE* =13), regardless of whether it was a valid (*M* =297 ms, *SE* =9, *t*(15) = 9.422, *p* < .001), invalid (*M* =324 ms, *SE* =8, *t*(15) = 6.083, *p* < .001), or a central cue (*M* =318 ms, *SE* = 8, *t*(15) = 6.527, p < .001). The speed-up in RT as a result of the presence of a cue in the second block points to the presence of a general alerting effect and an effect of a temporal warning. More importantly, however, RTs in the Valid Cue condition were significantly shorter compared to the Invalid Cue [*t*(15) = -8.251, *p* < .001] and the Center Cue condition [*t*(15) = -4.443, *p* < .003]. There was no significant difference in RT between the Invalid and Center Cue condition [*t*(15) = 2.027, *p* = .315], indicating that RTs to targets following a Center Cue were much like RTs to targets following an Invalid Cue.[5]

Additionally, there was a main effect of Target Modality [*F*(1.390, 20.856) = 81.117, *p* < .001, ε = .695, partial η^2 = .844]. RTs on audiovisual target trials (*M* =297 ms, *SE* =10) were significantly shorter compared to RTs on visual (*M* =331, *SE* =8, *t*(15) =9.198, *p* < .001), and auditory target trials (*M* = 359 ms, *SE* =10, *t*(15) =14.685, *p* < .001), indicative of multisensory facilitation of response times. Responses to auditory targets (*M* = 359 ms, *SE* = 10) were significantly slower compared to visual targets (*M* = 331 ms, *SE* = 8, *t*(15) =4.362, *p* = .003). The observation that RTs to auditory targets were generally slower compared to RTs to visual targets can be explained by the fact that auditory localization is generally more difficult than visual localization (e.g., Frens & Van Opstal, 1995).

---

[5] The centrally presented cue is in fact an invalid cue, but presented at a smaller distance from the target location compared to the Invalid Cue condition. Still, it is difficult to say for sure whether audiovisual targets following center cues were diffusely attended or entirely unattended (given that it was an invalid cue).

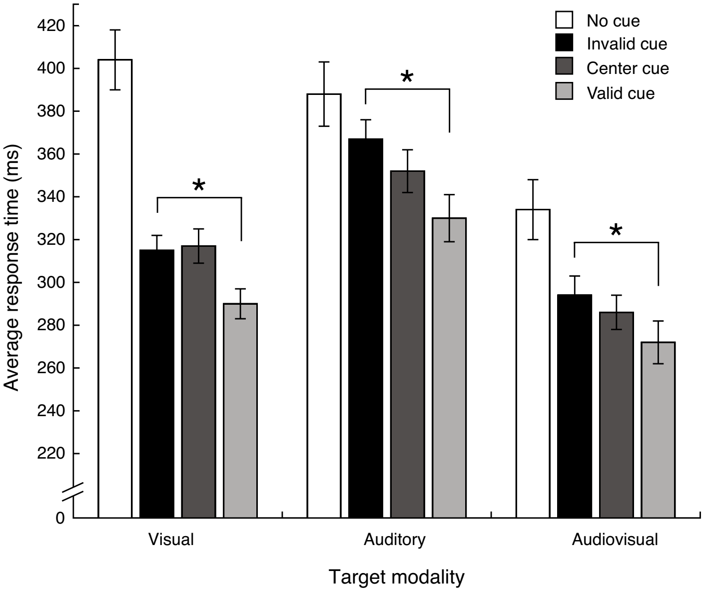


Figure 2. Mean response times to visual, auditory, and audiovisual targets for each cue type: No Cue (white), Invalid Cue (black), Center Cue (dark grey), and Valid Cue (light grey). Only the validity effects are indicated with an asterisk to maintain a clear view of the bar graph ($p < .05$). See the text for more statistics on the differences between target modalities and cue types.

The interaction between Cue Type and Target Modality was also significant [$F(3.588, 53.813) = 14.219$, $p < .001$, $\varepsilon = .598$, partial $\eta^2 = .487$]. This interaction could be explained by differences in the size of validity effects for different target modalities, by varying differences in RT between target modalities across the different cue types, or a combination of both.

To investigate the cause of the interaction, we first used pairwise comparisons to check whether cueing effects (CE; difference in RT between validly cued and invalidly cued targets) were present for each Target Modality and whether they were different in size. The difference in RT between Valid and Invalid cues was significant for auditory (mean CE = 36 ms, *M* valid = 330 ms, *SE* =11 vs. *M* invalid = 366 ms, *SE* = 9, $t(15) = -5.398$, $p < .001$), visual (mean CE = 25 ms, *M* valid = 290 ms, *SE* =7 vs. *M* invalid = 315 ms, *SE* = 7, $t(15) = -5.967$, $p < .001$), and audiovisual targets (mean CE = 22 ms, *M* valid = 272 ms, *SE* =10, vs. *M* = 294 ms, *SE* = 9, $t(15) = -6.267$, $p < .001$). There were no significant differences between target modalities in the size of the validity effect (mean validity effect = 27 ms, *SE* =3, all *p*'s > .22). Therefore, these results do not explain the interaction, but they do indicate that the exogenous auditory cue caused an exogenous shift of attention that facilitated responses to unimodal (in a crossmodal and in an intramodal way) and bimodal targets that were presented at the same location as the cue. The validity effects are clearly visible in Figure 2.

To further investigate the interaction between Cue Type and Target Modality, we compared differences between target modalities for each Cue Type using pairwise comparisons. In the No Cue condition, responses to audiovisual targets (*M* = 334 ms, *SE* = 14) were faster as compared to auditory targets [*M* = 388 ms, *SE* = 15, *t*(15) = 7.580, *p* < .001] as well as to visual targets (*M* = 404 ms, *SE* = 14, *t*(15) = 7.612, *p* < .001). The difference in RT between auditory and visual targets was not significant [*t*(15) = –1.329, *p* = .495].

In the Valid Cue condition, RTs on audiovisual target trials (*M* =272 ms, *SE* = 10) were shorter compared to RTs on auditory (*M* =330 ms, *SE* =11, *t*(15) = 10.629, *p* < .001) and visual target trials (*M* =290 ms, *SE* = 7, *t*(15) =3.469, *p* = .010]. Furthermore, responses to validly cued auditory targets were significantly slower compared to responses to validly cued visual targets [*t*(15) = 5.118, *p* < .001].

A similar pattern was observed for invalidly cued targets and centrally cued targets. In the Invalid Cue condition, RTs on audiovisual target trials (*M* = 294 ms, *SE* = 9) were shorter than on visual target trials (*M* = 315, *SE* = 7, *t*(15) = 4.175, *p* < .002) and auditory target trials (*M* =366, *SE* =9, *t*(15) = 11.194, *p* < .001). The difference in RT between invalidly cued visual and auditory targets was also significant [*t*(15) = 6.175, *p* < .001], with invalidly cued visual targets responded fastest to.

In the Center Cue condition, responses to audiovisual targets (*M* =286 ms, *SE* = 8) were faster than to auditory (*M* = 352 ms, *SE* =10, *t*(15) = 10.474, *p* < .001) and visual targets (*M* = 317, *SE* = 8, *t*(15) = 9.321, *p* < .001). Responses to visual targets were faster than to auditory targets in the Center Cue condition [*t*(15) = 5.822, *p* < .001].

In sum, these results indicate that whereas there was no difference in RT between auditory and visual targets in the No Cue condition, this was the case in all cued conditions, which explains the interaction.

Multisensory response enhancement

Significant absolute MRE was observed in all cue conditions as indicated by one-sample t-tests [No Cue: *M* = 44 ms, SE = 6, *t*(15) = 6.937, *p* < .001; Valid Cue: *M* = 16, *SE* = 5, *t*(15) = 3.316, *p* = .005; Invalid Cue: *M* = 20, *SE* = 5, *t*(15) = 4.175, *p* = .001; and Center Cue: *M* = 30, *SE* = 3, *t*(15) = 10.087, *p* < .001]. The mean absolute MRE for each cue type is shown in Figure 3.

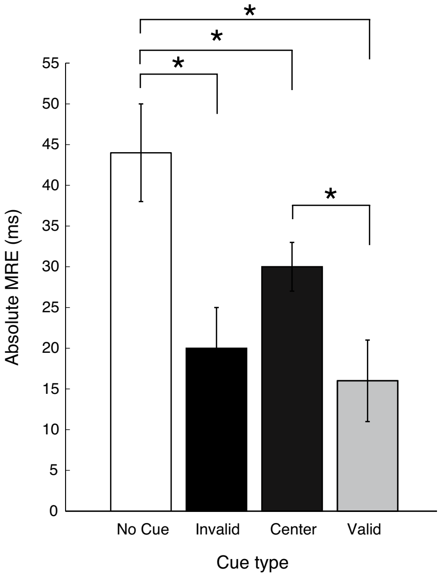


Figure 3. Mean absolute MRE in milliseconds for each cue type: No Cue (white), Invalid Cue (black), Center Cue (dark grey), Valid Cue (light grey). Significant differences are indicated with an asterisk ($p < .05$).

To test for difference in the amount of MRE, a repeated measures ANOVA was used. There was a main effect of Cue Type [$F(1.919, 28.791) = 7.854$, $p = .002$, $\varepsilon = .640$, partial $\eta^2 = .344$]. Planned pairwise comparisons indicated that MRE was significantly larger in the No Cue condition ($M = 44$ ms, $SE = 6$) compared to the Valid ($M = 16$ ms, $SE = 5$, $t(15) = 3.658$, $p = .002$), the Invalid ($M = 20$ ms, $SE = 5$, $t(15) = 3.059$, $p = .008$), and the Center Cue Condition ($M = 30$ ms, $SE = 3$, $t(15) = 2.158$, $p = .048$). The amount of MRE was larger in the Center compared to the Valid Cue condition [$t(15) = -2.307$, $p = .036$]. The differences in MRE between the Valid and the Invalid condition [$t(15) = -1.508$, $p = .152$] and the Invalid and Center condition [$t(15) = -1.780$, $p = .095$] were, however, not significant.

Center Cues are different from the No Cue and the other Cue conditions in that they were presented at the No-go location. Therefore, response inhibition to targets presented at the central location may have partly canceled the attention attracting effect of the Center Cue, resulting in a pattern of multisensory enhancement that is much more like the No Cue condition as compared to the other Cue Types (i.e., Valid and Invalid).

Race model violation

To investigate whether the race model could explain the speed-up in the audiovisual condition in each cue condition, we compared the audiovisual CDF and the race model CDF for each cue type for each quantile bin of each CDF. The results revealed significant race model violations for all cue types. Figure 4 (left panel) shows the average amount of race model violation in each of the cue conditions for each of the quantiles. A significant race model violation was observed for the 10th to the 70th percentile in the No Cue condition ($p < .05$), for the 10th to the 50th quantile in the Center Cue condition ($p < .05$), for the 10th to the 40th percentile in the Invalid Cue condition ($p < .05$), and for the 10th and the 20th percentile in the Valid Cue condition ($p < .05$). The race model could only be rejected for the fastest RTs in the Valid Cue condition, whereas this was true for a broader range of RTs in the Invalid, Center, and No Cue condition.

We observed a significant main effect of Cue Type on the average of the median race model violation [$F(3,45) = 5.601$, $p = .002$, partial $\eta^2 = .272$]. The average of the median amount of race model violation across quantiles in each Cue condition is shown in the right panel of Figure 4. Planned pairwise comparisons revealed that the amount of race model violation in the No Cue condition (M = 18 ms, $SE = 5$) was significantly larger compared to the Invalid Cue (M = 7 ms, $SE = 4$, $t(15) = 2.247$, $p = .040$), and the Valid Cue condition (M = 1 ms, $SE = 4$, $t(15) = 3.466$, $p = .003$), but not compared to the Center Cue condition (M = 11 ms, $SE = 3$, $t(15) = 1.447$, $p = .168$). More importantly, although there was no difference in absolute MRE between the Valid and the Invalid Cue condition, there was a difference in race model violation between the Valid and Invalid Cue condition [$t(15) = -2.634$, $p = .019$]. The difference between the Valid and Center Cue condition was also significant [$t(15) = -2.314$, $p = .035$], but there was no difference in the average of the median amount of race model violation between the Invalid and Center Cue condition [$t(15) = -1.122$, $p = .279$].

The audiovisual CDF and the race model CDF can be compared for comparable quantiles resulting in a difference in ms between two CDFs (as shown in Figure 4), but it can also be compared in terms of differences in probability for comparable RTs (see Figure 5; see also Laurienti, Burdette, Maldjian, & Wallace, 2006, for example). The finding that the race model was violated at a larger range of quantiles in the Invalid Cue condition as compared to the Valid Cue condition was also supported by the range of RTs in which the race model was violated in terms of probability enhancement (see Figure 5, left panel). In the No Cue condition the RT range in which the race model was positively violated was 142 ms long (from 244 to 385 ms). In the Valid Cue condition this window was 66 ms long (185–250 ms), which was smaller compared to the Invalid

Cue condition in which this window was 96 ms long and there was a shift to somewhat slower response times (214–309 ms). The size of the window of RTs in which the race model was violated in the Center Cue condition was similar to that in the No Cue condition, but shifted to faster RTs (window = 135 ms, from 192–326 ms).

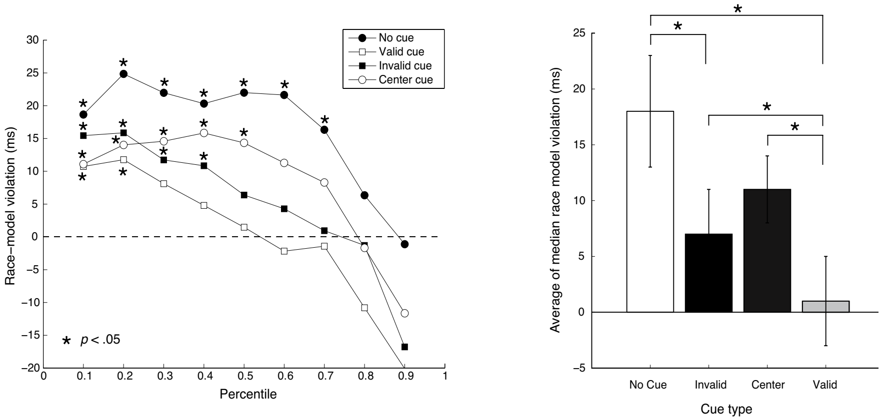


Figure 4. Left panel: Mean race model violations for each Cue condition. No Cue (filled black circle), Center Cue (open circle), Valid Cue (open square), and Invalid Cue (filled black square). Right panel: Average of the median race model violation across all percentiles for each cue condition. Significant differences are indicated with an asterisk (*p* < .05).

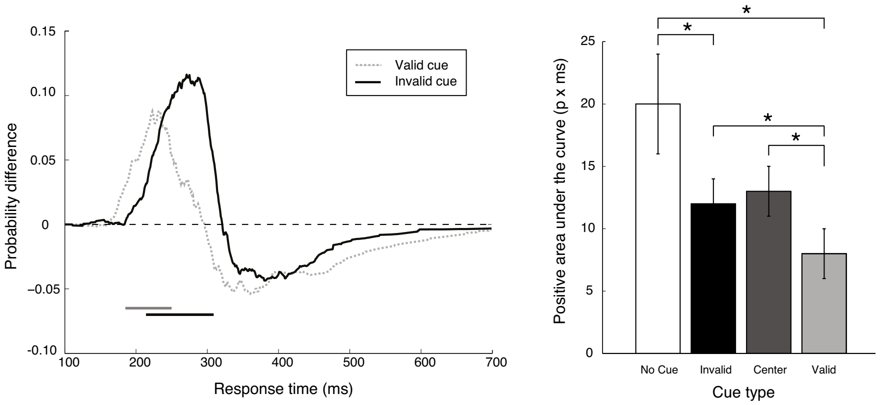


Figure 5. Left panel: Mean probability enhancement over the race model across the full range of RTs for the Valid Cue (dashed grey line), and the Invalid Cue condition (solid black line). Significant violations are indicated with a horizontal bar in the graph below the x-axis indicating the RT range in which the probabilities of consecutive time points are larger than predicted by the race model (*p*-values were uncorrected for multiple tests but had to be consecutive). Right panel: Average positive area under the curve in each of the cue conditions. Significant differences are indicated with an asterisk (*p* < .05)

The positive area under the curve was also compared between the different Cue conditions (see Figure 5, right panel). The repeated measures ANOVA revealed a main effect of Cue Type [$F(1.960, 29.396) = 5.242$, $p = .012$, $\varepsilon = .653$, partial $\eta^2 = .259$]. In line with the other measures of race model violation, the positive area under the curve was significantly larger in the Invalid Cue condition ($M = 12$ ms, $SE = 2.3$) compared to the Valid Cue condition ($M = 8$ ms, $SE = 2$, $t(15) = -2.356$, $p = .033$). The average positive area under the curve was significantly larger in the Center Cue ($M = 13$ ms, $SE = 2$) as compared to the Valid Cue condition [$t(15) = -2.173$, $p = .046$], but not compared to the Invalid Cue condition [$t(15) = -.635$, $p = .535$] and the No Cue condition ($M = 20$ ms, $SE = 4$, $t(15) = 1.917$, $p = .074$). The positive area under the curve in the No Cue condition was significantly different from the Valid [$t(15) = 3.337$, $p = .005$] and the Invalid Cue condition [$t(15) = 2.138$, $p = .049$].

Given that responses in the Valid Cue condition were generally faster as compared to the Invalid Cue condition, one might argue that race model violations are generally smaller in the Valid Cue condition as there may be less room for improvement in terms of RTs. In order to explore this possibility, correlations between RTs to audiovisual targets and the amount of race model violation in the Valid and Invalid Cue condition were calculated. There was a significant negative correlation between RTs and the median race model violation in both the Valid (Pearson $r = -.829$, $p < .001$) and the Invalid Cue condition (Pearson $r = -.821$, $p < .001$). These correlations indicate that the faster responses to audiovisual targets were, the larger the race model violation was at both the validly and the invalidly cued target locations, which is in contrast with the explanation that race model violations decrease with increasing RTs to audiovisual targets. Negative correlations were also observed between unimodal RTs and race model violation in the Valid and Invalid Cue condition (all correlations were negative, but not all correlations were significant, see Table 1). Furthermore, there was no significant difference in absolute MRE between the Valid and Invalid Cue condition, indicating that the differences in race model violation were not the result of differences in the amount of absolute MRE that was possible in the Valid and Invalid Cue condition.

Modality switch effects

Several different factors may contribute to the observed amount of race model violation, indicating that one must be careful with interpreting race model violations as a pure measure of multisensory integration. One of the effects that can contribute to the observed amount of race model violation is the modality Switch effect (MSE; Gondan, Lange, Rösler, & Röder, 2004; Otto & Mamassian, 2012, Spence, Nicholls,

Table 1 *Pearson correlations between response times to unimodal targets and the corresponding median amount of race model violation for the Valid and Invalid Cue condition.*

	Median race model violation Valid Cue	Median race model violation Invalid Cue
RT Audiovisual Valid	-.829*	-.729*
RT Audiovisual Invalid	-.760*	-.821*
RT Visual Valid	-.475	-.380
RT Visual Invalid	-.452	-.478
RT Auditory Valid	-.732*	-.687*
RT Auditory Invalid	-.813*	-.659*

* p < .01

& Driver, 2001). In a typical redundant target effect paradigm (e.g., Miller, 1986), the modality of the target is randomized across trials. This randomization results, however, in differences in RTs between trials in which a Target Modality switch occurred (relative to the previous trial) as compared to No Switch trials. Modality switches have been shown to contribute to the amount of race model violation that was observed (e.g., Gondan, Lange, Rösler, & Röder, 2004; Otto and Mamassian, 2012). As the race model is based on the distribution of RTs for unimodal auditory and visual targets, building a race model using the CDFs of faster RTs (e.g., No Switch trials) will result in a faster RT prediction of the race model as compared to slower RTs (e.g., Switch trials). In order to investigate whether the MSE partly explained the amount of race model violation in our paradigm, we also analyzed the MSE for unimodal auditory visual target trials using a N-1 trial history analysis with respect to Target Modality. Switch trials were defined as trials in which a unimodal auditory or a unimodal visual target was preceded by a unimodal target of the different modality on the previous trial (e.g., a unimodal visual target on the current trial, and a unimodal auditory target on the previous trial). All other trials were considered No Switch trials (although AV to V and AV to A trials contained both a Switch and a No Switch). The number of Switch trials in the current experiment was, however, rather low, and smaller than the number of unimodal No Switch trials as indicated by a main effect of Switch Type [$F(1, 15) = 541.710$, $p < .001$, partial $\eta^2 = .973$, mean number of Switch trials = 8, $SE = .258$, mean number of No Switch trials = 17, $SE = .260$, see Figure 6, left panel]. There was no main effect of Target Modality [$F(1, 15) = .273$, $p = .609$, partial $\eta^2 = .018$], but there was a significant interaction between Target Modality and Switch Type [$F(1, 15) = 11.739$, $p = .004$, partial $\eta^2 = .439$]. Although not significant after correction, this interaction seemed to be driven by a slightly large number of auditory No Switch trials as compared to visual No Switch trials [$t(15) = 2.766$, $p = .055$]. There was no significant difference between the number auditory and visual Switch trials [$t(15) = -2.192$, $p = .168$].

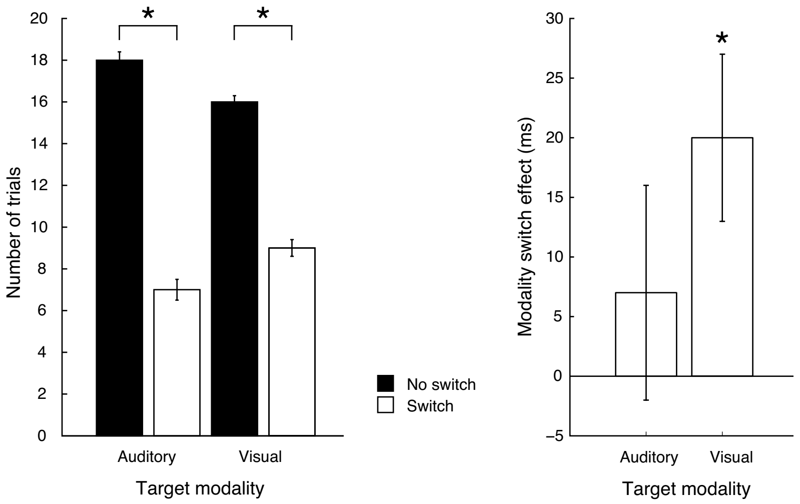


Figure 6. Left panel: Mean number of Switch and No Switch trials for unimodal auditory and visual targets in the No Cue condition. An asterisk indicates a significant difference between conditions ($p < .05$). Right panel: The modality switch effect for auditory and visual targets. The asterisk indicates a significant difference from zero ($p < .05$).

The MSE was not analyzed for the Cue block because an auditory cue was always presented between consecutive targets. Thus, there was never a modality switch between the auditory exogenous cue and the auditory target (intramodal cueing), whereas there was always a modality switch for visual target trials (crossmodal cueing). In trials in which an audiovisual target was present, both a Switch and a No Switch occurred (i.e., both intra- and crossmodal cueing). We therefore decided that a trial history analysis would not make sense in the Cue block, because modality switches were fixed for each Target Modality and the same for each cue condition.

The repeated measures ANOVA with the factors Target Modality (A, V) and Switch Type (Switch, No Switch) revealed no effects of Target Modality [$F(1, 15) = .000$, $p = 1.000$, partial $\eta^2 = .000$], Switch Type [$F(1,15) = 4.130$, $p = .060$, partial $\eta^2 = .216$], nor an interaction between Target Modality and Switch Type [$F(1, 15) = 1.689$, $p = .213$, partial $\eta^2 = .101$, see Figure 6, right panel]. We also directly analyzed the size of the MSE (the difference in RT between Switch and No Switch trials) for unimodal auditory and unimodal visual targets using one-sample t-tests, and observed an MSE for unimodal visual targets [$t(15) = 2.794$, $p = .014$], but not for auditory targets [$t(15) = .703$, $p = .493$].

The race model was calculated separately using unimodal Switch and unimodal No Switch trials to further explore the contribution of MSE to race model violations in the

No Cue condition. The audiovisual CDF was then compared to both the Switch and the No Switch race models. For two participants there were not enough unimodal auditory Switch trials (N = 4 for both participants) to calculate the race model so the race model violation analysis was performed on data from the remaining fourteen participants. On average the No Switch race model was a bit faster as compared to the Switch race model. There were fewer percentiles at which the race model was significantly violated during No Switch trials (significant violations at the 10th and 70th percentile) as compared to Switch trials (significant violations at the 10th, 20th, 30th, and 70th percentile, see Figure 7, left panel). The average of the median amount of race model violation for the Switch and the No Switch condition was still significant as indicated by one sample t-tests (No Switch: M =16 ms, SE =6 ms, $t(13)$ =2.238, p = .043; Switch: M = 26 ms, SE = 7 ms, $t(13)$ =3.659, p = .003). There was no significant difference in the average of the median amount of race model violation between the Switch and the No Switch condition [$t(13)$ = –1.978, p = .069, see Figure 7, right panel]. Thus, although MSE did seem to contribute a little to the amount of race model violation, the observed difference was not significant. These results must be interpreted with care because the number of No Switch and especially the number of Switch trials was very low (see Figure 6).

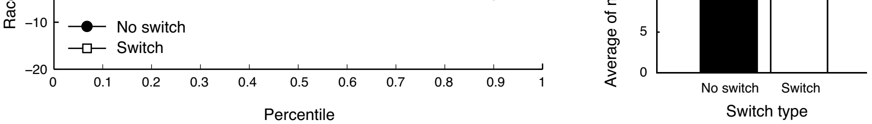


Figure 7. Left panel: The average race model violation in the Switch and No-switch trials in the No Cue condition of Experiment 1. Significant differences are indicated with an asterisk (p < .05, Bonferroni corrected). Right panel: The average of the median race model violation for Switch and No Switch trials. Asterisks indicate a significant difference from zero (p < .05)

Discussion

The goal of the present study was to investigate whether exogenous spatial attention was able to change the outcome of multisensory integration. In order to do so, valid, invalid, and center exogenous auditory cues were presented before unimodal and bimodal audiovisual targets. The results of Experiment 1 revealed significantly larger race model violations at exogenously unattended compared to exogenously attended locations. These results suggest that multisensory integration is decreased at exogenously attended locations as compared to exogenously unattended locations. A simple explanation for this effect could be that response times cannot be much faster at attended locations and therefore results in less benefit of multisensory stimulation. There are two arguments that can be made against this explanation. First, there were significant negative correlations between response times to unimodal auditory/visual/audiovisual targets and the corresponding amount of race model violations in both the Valid and the Invalid Cue condition. The amount of race model violation increased as responses became faster, both in the Valid and in the Invalid Cue condition. This indicates that the amount of race model violation did not systematically decrease as the absolute RTs decreased (i.e., faster responses did not results in less race model violation). Furthermore, the absolute amount of MRE did not significantly differ between the validly and the invalidly cued targets, indicating that in terms of absolute speed up in the multisensory condition compared to the fastest unimodal condition, there was no difference between attended and unattended locations. Crucially, most of the multisensory speed up at exogenously attended locations could, however, be explained by statistical facilitation, whereas this was not the case at exogenously unattended locations.

One could argue that the difference in sensitivity between visual targets and auditory targets makes the interpretation of the results difficult. However, audiovisual targets always consisted of the same unimodal components and the critical manipulation of interest was a possible modulation of MRE by exogenous spatial attention. There were no differences in sensitivity between cue conditions and no interaction between cue condition and Target Modality. Therefore, we do not think that the observed differences in sensitivity are problematic because these differences were equal for the Valid and the Invalid Cue condition.

Race model violations can sometimes be explained by taking trial-history effects into account. In particular modality switch effects can contribute to the amount of race model violation (e.g., Otto & Mamassian, 2012). As others have shown, this is not always the case (Gondan, Lange, Rösler, & Röder, 2004). In the No Cue condition of Experiment 1, there was a slight decrease in the amount of race model violation, which could be explained by MSE. However, this decrease was not significant and the race model was still violated when the AV CDF was compared to a race model that consisted of only unimodal No Switch trials. Trial history effects were not analyzed in the cued conditions as the exogenous auditory cue always caused a modality switch for visual targets within a trial, and never for auditory targets within a trial. For audiovisual targets, both a modality Switch and a No Modality Switch occurred after the presentation of the cue.

Additionally, we observed spatial cueing effects for all target modalities. If modality switch effects contributed more to the speed up of responses to different targets than exogenous spatial attention, then validly cued unimodal auditory targets should be responded to faster as compared to unimodal visual targets, because auditory processing benefits both from within modality priming and exogenous spatial attention. This cannot be concluded from the data, as responses to visual targets were always faster compared to responses to auditory targets in the cued conditions despite the modality switch that was always present within a trial for visual targets. Importantly, there was no difference in RTs for unimodal visual and auditory targets in the No Cue condition. We therefore argue that the current observation of larger race model violations for invalidly cued targets as compared to validly cued targets reflects differences in multisensory integration between exogenously attended versus unattended target locations.

The amount of multisensory integration has been shown to differ between localization and detection paradigms (e.g., Hecht, Reiner, & Karni, 2008). Whereas in localization paradigms participants have to process the spatial location of targets before responding, spatial information is taskirrelevant in detections paradigms. The results from previous studies suggest that multisensory integration of simple lights and sounds is especially helpful when spatial orienting is taskrelevant (e.g., Hecht, Reiner, & Karni, 2008). Exogenous spatial attention may therefore have a larger influence on multisensory integration in tasks in which spatial localization is important. When spatial localization is task-irrelevant, the influence of exogenous spatial attention on multisensory integration may decrease substantially as the task can be completed without relying on spatial information. In order to test possible task dependencies of the effect of exogenous spatial attention on multisensory integration, we conducted another experiment in which a Go/No-go detection task was used. Although the task was still to detect auditory, visual, and audiovisual targets using one button, this time targets were only presented from the left and the right side of the fixation cross but not at the central location. No-go (catch) trials did not consist of the presentation of the target at the fixation cross, but of the absence of a target. This way, participants did not have to localize the target before deciding to respond to the presence of a target in Experiment 2, which was the case in Experiment 1.

Experiment 2

Materials and methods

Participants

Twenty-four participants were tested in Experiment 2 (six male, 18 female, mean age = 25.5 years, *SD* = 3.5). All participants had normal or corrected-to-normal visual acuity and did not report any hearing problems. The participants received either course credits or a monetary reward for their participation. The experiment was conducted in accordance with the Declaration of Helsinki, and participants signed informed consent before the start of the experiment.

Apparatus

Visual and auditory stimulus presentation was controlled through a custom-built audiovisual stimulus generator that was connected to a PC running MATLAB. Visual stimuli consisted of tri-colored Light Emitting Diodes (LEDs, Forge Europa, bulb size: 5 mm, viewing angle: 65°, LED colour: red, green, and blue) of which the intensity and color could be adjusted independently. Auditory stimuli were presented over loudspeakers (e-audio black 4" Full Range Mini Box Speaker, dimensions: 120 × 120 × 132 mm, frequency response: 80–20,000 Hz). With the use of the audiovisual stimulus generator, stimuli were presented with high temporal precision (i.e., an accuracy of 1 ms in terms of their onset and offset during a trial, timing was confirmed using an oscilloscope). Participants were instructed to place their chin in a chin-rest to keep the distance between the stimuli and the participants the same. Three loudspeakers and three LEDs were used to present the stimuli at a distance of 64 cm from the participant. Each LED was attached to the center of a speaker ensuring precise spatial alignment. The left and the right speakers with LEDs were placed at 33° to the left and the right of the central speaker and LED.

Stimuli, task, and procedure

As in Experiment 1, participants had to detect auditory, visual, and audiovisual targets. This time, targets were presented only at the left and the right location, and not at the central location. Whereas Experiment 1 was a Go/No-go localization detection task, Experiment 2 was a simple detection task. Exogenous auditory cues were present in each trial and were presented from the left or right location. Targets could be presented at the same or a different location (i.e., from the opposite side of the central fixation location) as the cue. Thus, targets could be either validly or invalidly cued. Each trial started with the onset of a central blue LED (266 cd/m^2) with a random duration between 750 and 1250 ms. After the fixation offset, a 100-ms exogenous auditory cue was presented from the left or right location. The cue was a 75-ms 800 Hz pure tone [~66 dB(A)]. The target consisted of the illumination of a green LED (2072 cd/m^2) for 100 ms, the presentation of a white noise sound [~66 dB(A)] for 100 ms, or the combination of the green LED and the white noise sound (SOA =0 ms; Target Modality and target location were randomized). The cue-target stimulus onset asynchrony varied randomly between 200 and 250 ms and the response window was set to 2000 ms after target onset after which the next trial started. There were 80 auditory targets (40 validly cued, 40 invalidly cued), 80 visual targets (40 validly cued, 40 invalidly cued), and 80 audiovisual targets (40 validly cued, 40 invalidly cued). Sixty catch trials (20%) were implemented in which no target was presented after the presentation of the cue to ensure that participants were paying attention to the task and responded only to the targets. The participants were instructed to press a button on a custom-made response box as quickly as possible after the detection of the green target LED, the white noise sound, or their combination, and to withhold their response when no target appeared after the auditory cue.

Data preprocessing

The same data preprocessing steps as in Experiment 1 were used.

Statistical analyses

The same analyses as in Experiment 1 were used with the exception that the factor Cue Type contained only two levels in Experiment 2 (Valid, Invalid).

Results

Accuracy

Overall, accuracy was high on Go (*M* = .974, *SE* = .007) and No-go trials (*M* = .916, *SE* = .100).

The repeated measures ANOVA of accuracy on Go trials revealed a main effect of Cue Type [*F*(1, 23) = 5.793, *p* = .025, partial η^2 = .201], but no main effect of Target Modality [*F*(1.585, 35.466) = 1.077, *p* = .338, ε = .793, partial η^2 = .045] and no interaction between Target Modality and Cue Type [*F*(2, 46) = 1.103, *p* = .341, partial η^2 = .046]. Accuracy on Go trials in the Invalid Cue condition was generally higher (*M* = .978, *SE* = .007) as compared to the Valid Cue condition (*M* = .969, *SE* = .007).

The analysis of sensitivity also indicated a main effect of Cue Type [*F*(1, 23) =5.121, *p* = .033, partial η^2 = .182]. There was no main effect of Target Modality [*F*(1.545, 35.537) =1.201, *p* = .303, partial η^2 = .050] and no interaction between Target Modality and Cue Type [*F*(2, 46) =1.355, *p* = .268, partial η^2 = .056]. The sensitivity was higher in the Invalid Cue condition (mean A = .971, *SE* = .005) as compared to the Valid Cue condition (mean A = .968, *SE* = .005), but the overall accuracy was so high that we want to be careful in interpreting this difference and do not want to draw strong conclusions from this result.

Response times

A main effect of Target Modality was observed [*F*(1.332, 30.627) = 60.376, *p* < .001, ε = .666, partial η^2 = .724]. RTs for audiovisual targets (*M* =268 ms, *SE* =11 ms) were significantly shorter as compared to auditory (*M* =333 ms, *SE* =15 ms, *t*(23) =9.531, *p* < .001) and visual targets (*M* =329 ms, *SE* =10 ms, *t*(23) = 15.347, *p* < .001). RTs for unimodal auditory and unimodal visual targets did not differ [*t*(23) = .454, *p* = .654].

The main effect of Cue Type was also significant [*F*(1, 23) = 42.162, *p* < .001, partial η^2 = .647], indicating that responses to validly cued targets were generally faster (*M* = 303 ms, *SE* = 12 ms) as compared to invalidly cued targets (*M* = 317 ms, *SE* = 12 ms, see Figure 8, left panel). There was no interaction between Target Modality and Cue Type [*F*(2, 46) = 2.349, *p* = .107, partial η^2 = .093].

A closer inspection of the validity effect (VE) for each Target Modality using one-sample t-tests revealed that there was a significant VE for visual (*M* = 24 ms, *SE* = 5 ms, *t*(23) = –5.020, *p* < .001) and audiovisual targets (*M* = 12 ms, *SE* = 3 ms, *t*(23) = –3.537, *p* = .002), but not for auditory targets (*M* = 8 ms, *SE* = 6 ms, *t*(23) = –1.311, *p* = .203).

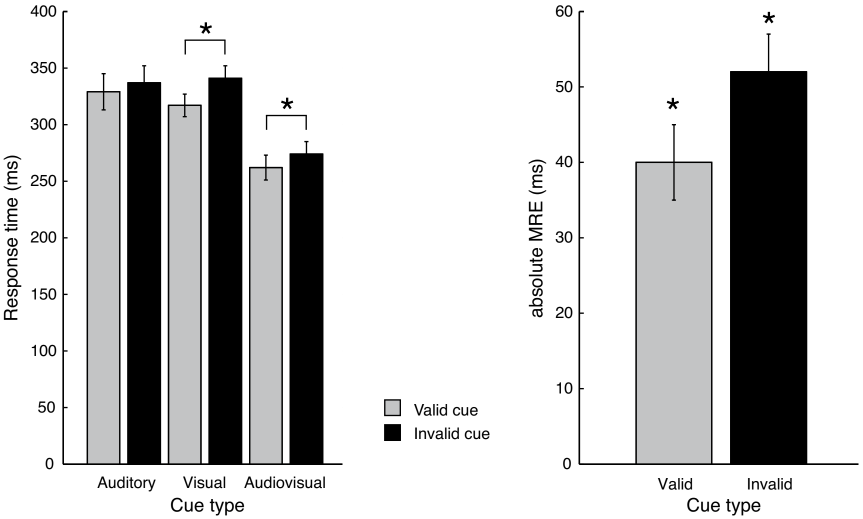


Figure 8. Left panel: The average of median response times for each Target Modality and Cue Type. Significant validity effects are indicated with an asterisk (*p* < .05). Right panel: The absolute amount of MRE in the Valid and Invalid Cue condition. Significant absolute MRE is indicated with an asterisk (*p* < .05).

Multisensory response enhancement

One-sample tests revealed that there was significant absolute MRE in both the valid (mean MRE = 40 ms, *SE* = 5 ms) and the invalid condition (mean MRE = 52 ms, *SE* = 5 ms, all *p*'s < .001, see Figure 8, right panel). The amount of absolute MRE was not significantly different between the Valid and the Invalid Cue condition (but close to significance, *t*(23) = −2.019, *p* = .055).

Race model violation

Significant race model violations were observed only in the Invalid Cue condition, not in the Valid Cue condition. The race model was significantly violated at the 40th and 50th percentiles (*p* < .05, 2-tailed, Bonferroni corrected, see Figure 9, left panel). There was no difference in the average of the median amount of race model violation between the Valid (*M* = 6 ms, *SE* = 4) and the Invalid Cue conditions (*M* = 7 ms, *SE* = 4, *t*(23) = −.294, *p* = .771, see Figure 9, right panel).

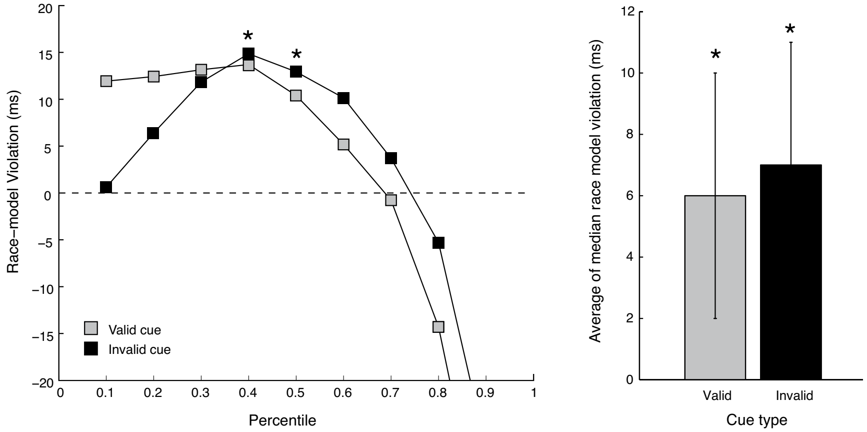


Figure 9. Left panel: The average race model violations for the Valid Cue (filled grey squares) and the Invalid Cue (filled black squares) condition. Right panel: Average of the median race model violation across all quantiles for the Valid and Invalid Cue condition. Significant race model violations are indicated with an asterisk ($p < .05$).

A more detailed look at the range of RTs in which the race model was violated at consecutive RTs is provided in Figure 10 (left panel). This was the case in the 193–323 ms range (window width = 131 ms) in the Valid Cue condition. As in Experiment 1, this range was larger and shifted to later RTs in the Invalid Cue condition in which consecutive significant race model violations were observed in the RT range 207– 437 ms (window width = 231 ms). The RT range in which consecutive race model violations occurred was 100 ms larger in the Invalid Cue condition as compared to the Valid Cue condition. The average positive area under the curve was significantly different from the race model both in the Valid ($M = 19$, $SE = 3$, $t(23) = 7.047$, $p < .001$) and the Invalid Cue condition ($M = 21$, $SE = 15$, $t(23) = 7.792$, $p < .001$), as shown by one-sample t-tests. There was no difference in the average positive area under the curve between the Valid and Invalid Cue condition [$t(23) = -.368$, $p = .716$, see Figure 10, right panel].

Discussion

In Experiment 2 we investigated the influence of exogenous spatial attention on multisensory integration in a Go/No-go detection paradigm in order to investigate whether the interaction between exogenous spatial attention and MSI depended on spatial localization. The observed effect of exogenous spatial attention on MSI was not as pronounced in Experiment 2 as in Experiment 1. There were differences in the range

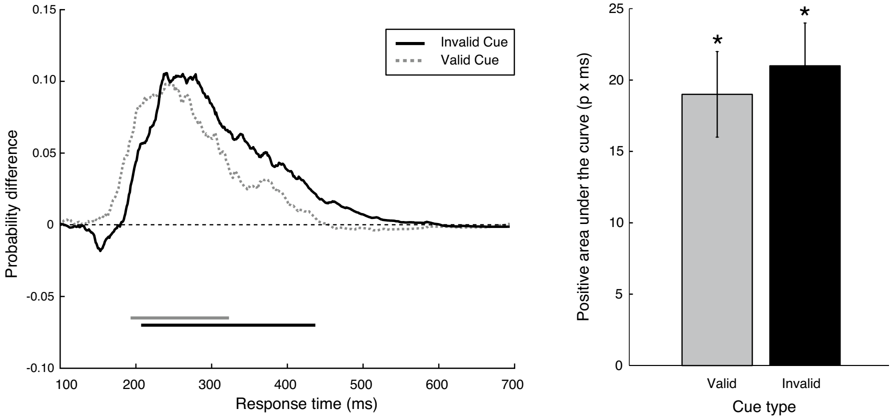


Figure 10. Left panel: Mean probability enhancement over the race model across the full range of RTs for the Valid Cue (solid black line), and the Invalid Cue condition (dashed grey line). Significant violations are indicated with a bar in the graph depicting the RT range at which the probabilities of consecutive time points are larger than predicted by the race model (*p*-values were uncorrected for multiple tests but had to be consecutive). Right panel: The average positive area under the curve in Valid and Invalid Cue condition. Significant differences are indicated with an asterisk ($p < .05$).

in which the race model was violated, but this was not reflected in all measures. The larger amount of MSI in the Invalid Cue condition was mainly visible in the absence of any race model violation in the Valid Cue condition (in terms of the absolute difference in ms between the race model CDF and the audiovisual CDF for different quantiles) and the broader range of RTs in which the race model was violated in terms of probability enhancement in the Invalid Cue condition. The RT range in which the race model was violated was 100 ms larger for targets appearing at exogenously unattended locations as compared to exogenously attended locations. In sum, although the spatial location of stimuli was task-irrelevant in Experiment 2, we still observed some effects of the exogenous cue on MSI. These results could be taken to suggest that the interaction between exogenous spatial attention and MSI is affected by whether spatial localization is task-relevant. This may not be very surprising considering that exogenous spatial attention is inherently spatial and its effects on other processes may therefore be especially pronounced in spatial tasks.

General discussion

The goal of the present study was to investigate whether exogenous spatial attention could affect multisensory integration of simple lights and sounds. This study is the first to show that exogenous spatial attention can influence MSI, given that enough time is provided for the exogenous cue to attract attention to the target location and that the cue is not part of the stimuli that need to be integrated. Although the absolute amount of MRE was the same when the target location was attended compared to when it was unattended, MSI was decreased at the attended location as compared to the unattended locations. This indicates that exogenous spatial attention speeds up the processing of multisensory stimuli (as indicated by cueing effects for unimodal and multisensory targets), but also decreases the amount of MSI. Besides the effects of exogenous spatial attention on MSI, the exogenous cue also caused an alerting effect and acted as a temporal warning, which contributed to a general decrease in MSI both at attended and unattended locations as indicated by a decrease in race model violation in the cued conditions as compared to the condition in which no cue was present. A temporal warning can increase the expectation of the appearance of a target and decrease MSI compared to a situation in which there is no temporal warning, or, in other words, a lower expectation of an upcoming target (e.g., Reches, Netser, and Gutfreund, 2010).

Several frameworks of multisensory integration have been proposed to account for the interactions (or lack of an interaction) between attention and multisensory integration (Koelewijn et al., 2010). Whereas the 'early integration' framework states that MSI occurs at an early pre-attentive stage, the 'late integration' framework states that integration takes place a late stage, leaving room for attention to affect the unimodal sensory input. Additionally, the parallel integration framework (Calvert & Thesen, 2004) states that MSI can occur at both 'early' and 'late' stages of sensory processing depending on spatial, temporal, and/or content properties of the stimuli. Several studies have provided examples of each of these stages and together provide support for the parallel integration framework (see Koelewijn et al., 2010, for a review). By presenting the exogenous spatial cue a little ahead in time of the multisensory target, exogenous spatial attention was allowed time to shift to the location of the target and decrease MSI. If we assume that the integration of spatially and temporally aligned simple lights and sounds is indeed the result of 'early' MSI, then we should perhaps reconsider the idea that 'early' integration always occurs in a pre-attentive way (as suggested by Koelewijn et al., 2010).

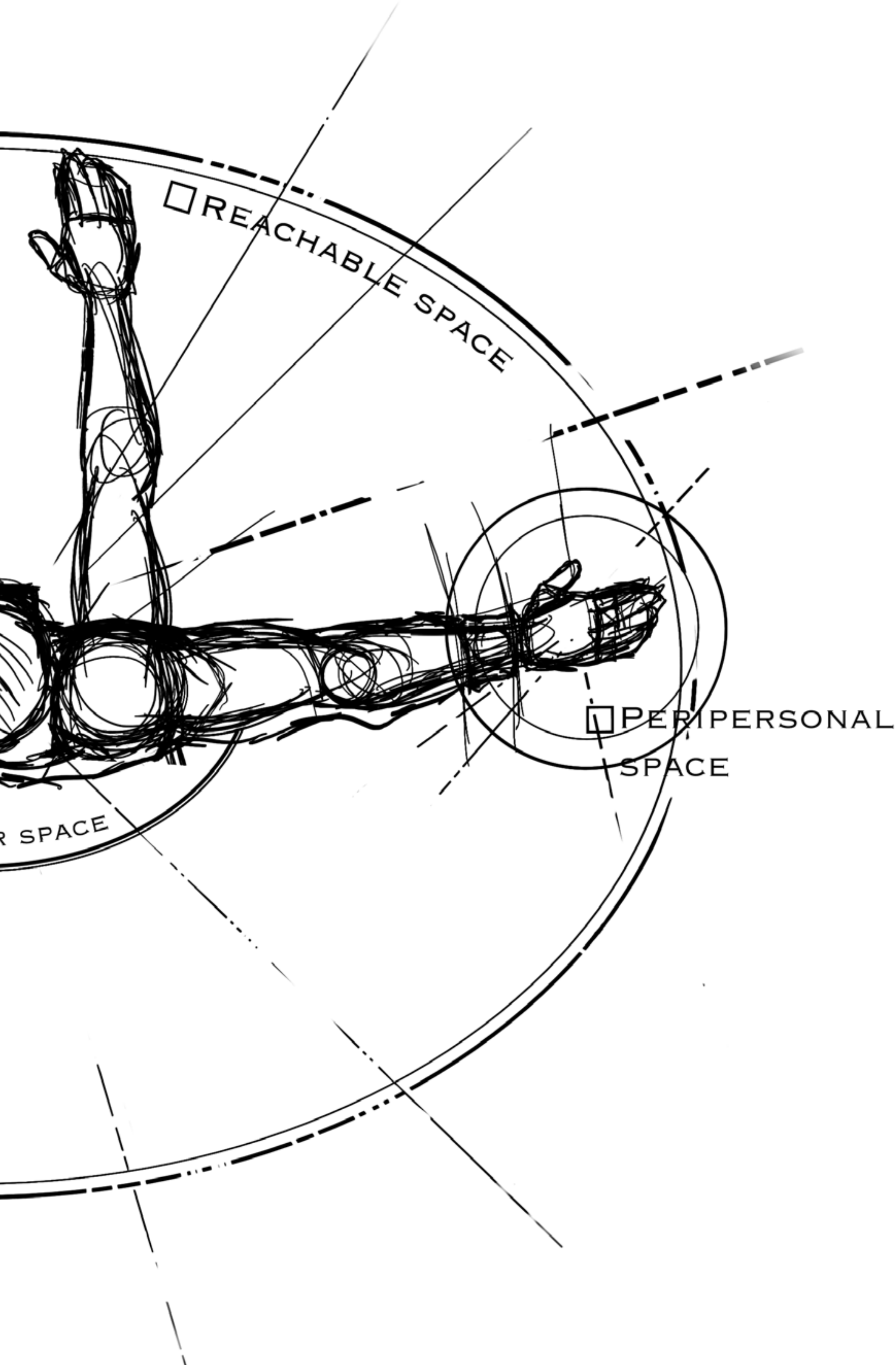
The effects of exogenous attention on MSI that we observed are opposite to those found in most studies in which the influence of endogenous attention on MSI was investigated (e.g. Talsma & Woldorff, 2005; Fairhall & Macaluso, 2009). Whereas previous studies have provided evidence for the idea that endogenously attending the target location enhances MSI of simple stimuli, we observed that exogenously attending the location of the target decreased MSI. Our findings are in line with the results from Zou, Müller, and Shi (2012) who found that MSI was enhanced in the endogenously unattended half of the stimulus field (as shown by the strength of the pip and pop effect). A key factor in whether attention is able to decrease multisensory integration at attended locations as compared to unattended locations might be the need to localize targets in space. When targets are always presented at the same location, there is no need to pay attention to multiple spatial locations. When the target location is uncertain (in both Experiment 1 and Experiment 2 of our study and the study by Zou, Müller, and Shi, 2012), MSI is decreased at the attended location as compared to the unattended location, while the opposite seems to be true for fixed target locations (e.g., Talsma & Woldorff, 2005; Fairhall & Macaluso, 2009).

The spatial uncertainty of target locations seems to be linked to the need for spatial orienting, with a higher uncertainty of the target location leading to a higher need for spatial orienting. This was also clear from the results of Experiment 2, in which targets did not have to be localized in order to perform the task, making spatial orienting task-irrelevant. As in Experiment 1, a small decrease in MSI was observed at exogenously attended locations as compared to exogenously unattended locations, but the effect was less pronounced (as seen in the lack of, for example, a significant difference in average race model violation). Even though the spatial location of cues and targets was task-irrelevant in Experiment 2, the location of the cue and the target was still uncertain (i.e., the target randomly appeared to the left or the right of the fixation cross, or did not appear), which may have increased the chances of spatial orienting and finding similar but less pronounced effects of exogenous spatial cues as in Experiment 1. The idea that some spatial orienting still occurred is supported by the observation of a cueing effect for visual and audiovisual targets, but the absence of a spatial cueing effect for auditory targets in Experiment 2 indicates that the cueing effect was less robust as in Experiment 1. If spatial uncertainty is indeed a key factor in the interaction between exogenous spatial attention and MSI, then MSI might be decreased at attended target locations because spatial orienting that is evoked by the cue and spatial orientation caused by the multisensory target (and therefore multisensory integration) are redundant (i.e., multisensory integration is not as helpful in spatial localization when attention is already at the right location). The opposite is true for exogenously unattended or invalidly cued targets.

A further possible explanation for the decrease in MSI at exogenously attended locations as compared to exogenously unattended locations could be related to the changes in perceptual sensitivity caused by exogenous cues. Exogenous spatial attention is able to increase perceptual sensitivity at attended locations as compared to unattended locations (e.g. Carrasco, Penpeci-Talgar, & Eckstein, 2000; see Carrasco, 2011, for a review). Research has shown that the effects of exogenous attention on perceptual sensitivity can be explained by a combination of response gain and contrast gain models (Ling & Carrasco, 2006). Whereas contrast gain is effectively similar to an increase of stimulus contrast at the attended location, response gain can be interpreted as an increase in stimulus intensity at the attended location. Exogenously attending a multisensory target could thus lead to an increase in perceived intensity and contrast. The principle of inverse effectiveness states that the benefit of multisensory integration (e.g., in terms of RT) is larger for weaker stimuli (e.g., less intense) than for stronger stimuli (e.g., more intense). One could argue that because perceptual sensitivity is higher at exogenously attended locations, MSI is decreased at exogenous attended locations through inverse effectiveness (but see Leone & McCourt, 2013; Holmes, 2007, 2009).

The current results may also be interesting from a more applied point of view. Exogenously attended multisensory targets were responded to the fastest in an absolute sense, compared to unattended multisensory targets. Thus, regardless of whether or not integration occurred when the target location was attended compared to when it was unattended, participants benefited the most from the combination of an exogenous cue and a multisensory stimulus, as shown by their RTs. This may be especially relevant in designing multisensory interfaces when the focus is on the speed of responding.

As far as we know, this is the first study to report that exogenous spatial attention influences audiovisual integration. Exogenous attention was able to speed up responses to multisensory stimuli, but at the same time also decreased the amount of MSI. These findings are in contrast with the idea that 'early' integration cannot be affected by exogenous spatial attention. Last, there may be an important role for spatial orienting and spatial uncertainty in this interaction, as indicated by a more pronounced interaction when the location of stimuli was task-relevant compared to when the location of the stimuli was task-irrelevant.


☐ Reachable space
☐ Peripersonal space
r space

Chapter 4

Inhibition of return affects audiovisual integration

Van der Stoep, N., Van der Stigchel, S., Nijboer, T. C. W., & Spence, C. (submitted).
Inhibition of return affects audiovisual integration.

Abstract

Multisensory integration (MSI) and exogenous spatial attention are both known to speed-up responses to perceptual events. Recently, it has been shown that although responses to exogenously attended multisensory targets are faster than to unattended targets, audiovisual integration at the exogenously attended location is actually reduced. This effect was observed for short cue-target intervals (200-250 ms). At longer intervals, the initial benefits of exogenous shifts of spatial attention at the cued location are often replaced by response time costs (a phenomenon 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 at cue-target intervals of 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. Thus, it would appear that IOR affects audiovisual integration by delaying visual information processing at cued as compared to uncued locations. This delay increased the difference in auditory and visual processing times and reduced multisensory response enhancement to the level of statistical facilitation at cued locations.

Introduction

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 and we also do not perceive everything that stimulates our sensory organs at each moment. Two mechanisms that help in combining and processing all this information are multisensory integration (MSI) and crossmodal exogenous spatial attention (Calvert & Spence, 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 facilitate MSI. First, MSI is often most pronounced when information from the different senses is presented from approximately the same spatial location (i.e., 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 by no means essential, just as long as sensory inputs arrive within a certain temporal binding window (i.e., discharge trains for visual and auditory inputs should overlap in a multisensory neuron; 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; Meredith & Stein, 1983).

Crossmodal exogenous spatial attention can 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 at an another, unattended, location (e.g., (Spence & Driver, 2004; Spence, Nicholls, Gillespie, & Driver, 1998; Ward, McDonald, & Lin, 2000). The 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) in those tasks requiring simple detection. 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 Spence, Nicholls, Gillespie, & Driver, 1998; Ward, McDonald, & Lin, 2000; Spence & Driver, 1998a, 1998b; Spence, Lloyd, McGlone, Nicholls, & Driver, 2000). It has been suggested that IOR facilitates efficient visual search as it inhibits the exploration of previously visited 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 (Hilchey, Hashish, MacLean, Satel, Ivanoff, & Klein, 2014; Hilchey, & Klein, 2014; Kingstone & Pratt, 1999; Taylor & Klein, 2000).

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. In that study, auditory exogenous spatial cues were presented before the onset of unimodal auditory (A), unimodal visual (V), and audiovisual (AV) targets. As expected, responses to exogenously attended AV targets were faster compared to unattended AV targets, but MSI was *reduced* at exogenously attended locations as compared to unattended locations. One explanation offered for this 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 responding to, information at cued locations at longer cue-target intervals, especially in simple speeded detection tasks. As IOR is most often investigated 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 cuing paradigm (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 the return of attention and/or responses to cued target locations are inhibited at *longer* cue-target intervals, then MSI might be enhanced at inhibited/cued locations as compared to uncued locations (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 large differences in auditory and visual processing times. In the latter case, visual, but not auditory, processing is inhibited, thus making auditory processing more dominant. Previous studies have shown that sensory dominance can affect the outcome of MSI (Ernst & Banks, 2002; Mozolic, Hugenschmidt, Peiffer, & Laruienti, 2008; Spence, Shore, & Klein, 2001).

The aim of the current study was to see how *multisensory* target processing would be affected by IOR. By calculating the amount of multisensory response enhancement and race model inequality (RMI) violation (Raab, 1962; Miller, 1986l; Ulrich, Miller, Schröter, 2007) for validly and invalidly cued 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, submitted) in a single session lasting approximately 1.5 hours in total. The order in which the studies were conducted was counterbalanced across participants. 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 participation 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 audiovisual stimulus generator was used to present the auditory and visual stimuli with an accuracy of 1 ms in terms of their onset and offset during a trial. Auditory stimuli were presented by means of loudspeakers (e-audio black 4” Full Range Mini Box Speaker, dimensions: 120 x 120 x 132 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. A verbal instruction was presented from the central loudspeaker while the central LED emitted blue light. The participants were asked whether they 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 during the actual experiment after making sure that the participants correctly performed 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 investigated. Three participants were removed from further analysis because their accuracy was equal to or fell below 50% in at least one condition. 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) was 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:

$$rMRE = [\min(\text{median}(RT_A), \text{median}(RT_V)) - \text{median}(RT_{AV}))]/[\min(\text{median}(RT_A), \text{median}(RT_V))] \times 100\%$$

To investigate whether a speed-up in the multisensory condition could be explained by statistical facilitation or by MSI, the audiovisual cumulative distributive function (CDF) was compared with the sum of the unimodal CDFs for each Cue Type at the 10th, 20th, 30th up to the 90th percentile (Miller, 1986). 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. Therefore, we did not go into detail on the results of the center cue condition, but we do present the results of this condition for the purpose of transparency.

Results

Accuracy

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

Response times

A 3 x 3 repeated measures ANOVA was used to analyze the RT data on the Go trials with the factors Cue Type (Valid, Invalid, and Center Cue) and Target Modality (Auditory, Visual, Audiovisual). A main effect of Target Modality was observed [$F(2, 40) = 99.473$, $p < .001$, partial $\eta^2 = .833$]. 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 [$F(2, 40) = .123$, $p = .884$, partial $\eta^2 = .006$]. The average of median RTs for A, V, and AV targets in the valid, invalid, and center cue condition are shown in Figure 1A.

There was a significant interaction between Cue Type and Target Modality [$F(2.576, 51.511) = 6.776$, $p = .001$, $\varepsilon = .644$, partial $\eta^2 = .253$], 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). The difference between validly and centrally cued visual targets (M = 476 ms, SE = 18) was not significant [$t(20) = -.712$, $p = .485$], but the difference between invalidly and centrally cued visual targets was [$t(20) = -2.597$, $p = .017$, d = -.266]. The latter two comparisons indicate that valid and central cues elicited a similar inhibitory aftereffect. 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). The difference between the invalid cue and center cue condition (M = 417 ms, SE = 17) was significant [$t(20) = 2.341$, $p = .030$, d = .317]. RTs in the valid cue auditory target condition did not differ from RTs in the central cue auditory target condition [$t(20) = .454$, $p = .655$].

No differences in RTs were observed between validly cued (*M* = 393 ms, *SE* = 19), invalidly cued (*M* = 382 ms, *SE* = 18), and centrally cued (*M* = 394, *SE* = 17) *audiovisual* targets (-1.5 < *t*'s < 1.5, *p*'s > .150). In sum, then, IOR was observed for visual targets, but not for auditory or audiovisual targets.

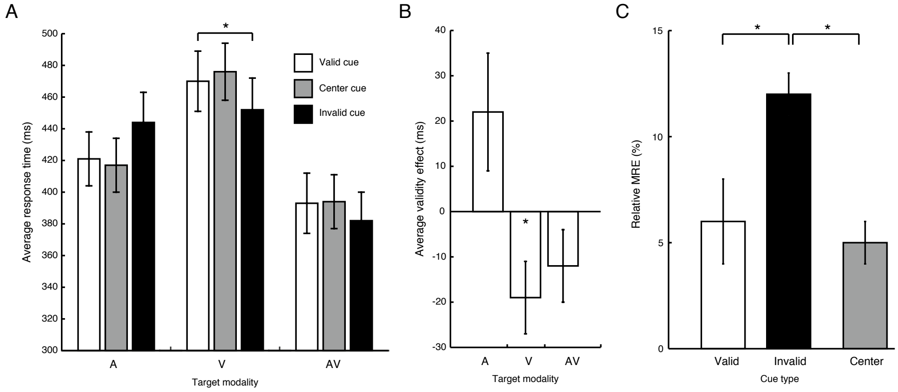


Figure 1. A: The average of median RTs for each Target Modality and Cue Type. Only significant validity effects are indicated with an asterisk (*p* < .05). B: The size and direction of the validity effect for each Target Modality. Positive values indicate that responses to validly cued targets were faster 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).

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. Differences in rMRE between Cue Types were analyzed using a repeated measures ANOVA. This revealed a significant main effect of Cue Type [*F*(2, 40) = 4.507, *p* = .017, partial η^2 = .184]. 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). The difference between the Invalid and the Center Cue condition was also significant (*M* = 5%, *SE* = 1, *t*(20) = 3.121, *p* = .005, d = 1.005), but the difference between the Valid and Center Cue condition was not [*t*(20) = .164, *p* = .872]. These results indicate that the amount of speed-up attributable to multisensory stimulation was significantly larger for invalidly as compared to validly 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.

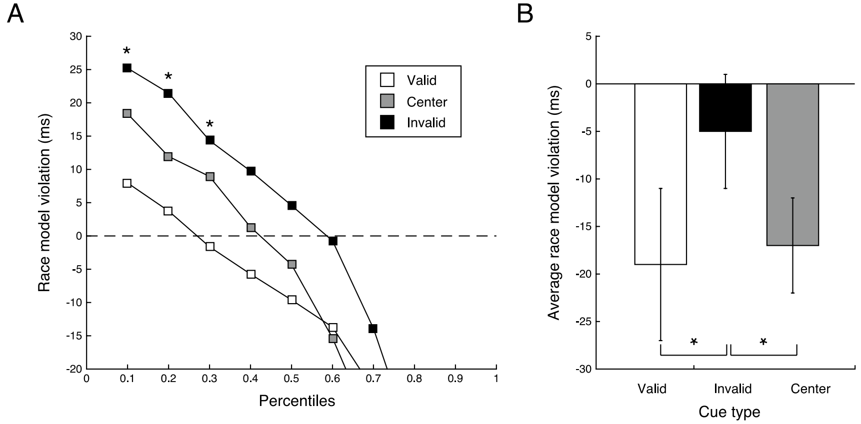


Figure 2. A: RMI violations in the Valid (white), Invalid (black), and Center Cue condition (grey). Significant positive violations of the RMI are indicated with an asterisk (*p* < .05). B: The average RMI violation across all percentiles for each Cue Type. Significant differences in RMI violation are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

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). In the Center Cue condition, the RMI was violated at the 10th and the 20th percentile (*t*'s > 3, *p*'s < .05). The RMI was not violated in the Valid Cue condition (*t*'s < 1.8, *p*'s > .4).

A repeated measures ANOVA with the factors Cue Type (Valid, Invalid, Center) and Percentile (10th to 90th) was conducted to quantify the differences in RMI violation. This analysis revealed a main effect of Cue Type [$F(2, 40) = 4.043$, $p = .025$, partial $\eta^2 = .168$]. The average amount of violation was significantly larger in the Invalid Cue condition ($M = -5$, $SE = 6$) than in the Valid Cue condition ($M = -19$, $SE = 8$, $t(20) = -2.748$, $p = .012$, d = -.406) and the Center Cue condition ($M = -17$, $SE = 5$, $t(20) = 2.235$, $p = .037$, d = .491, see Figure 2B). There was no significant difference in average RMI violation between the Center Cue and the Valid Cue condition [$t(20) = -.345$, $p = .734$]. There was also a significant main effect of Percentile [$F(1.983, 39.660) = 49.428$, $p < .001$, $\varepsilon = .145$, partial $\eta^2 = .712$], 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 [$F(3.457, 69.137) = 1.206$, $p = .315$, $\varepsilon = .216$, partial $\eta^2 = .057$].

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

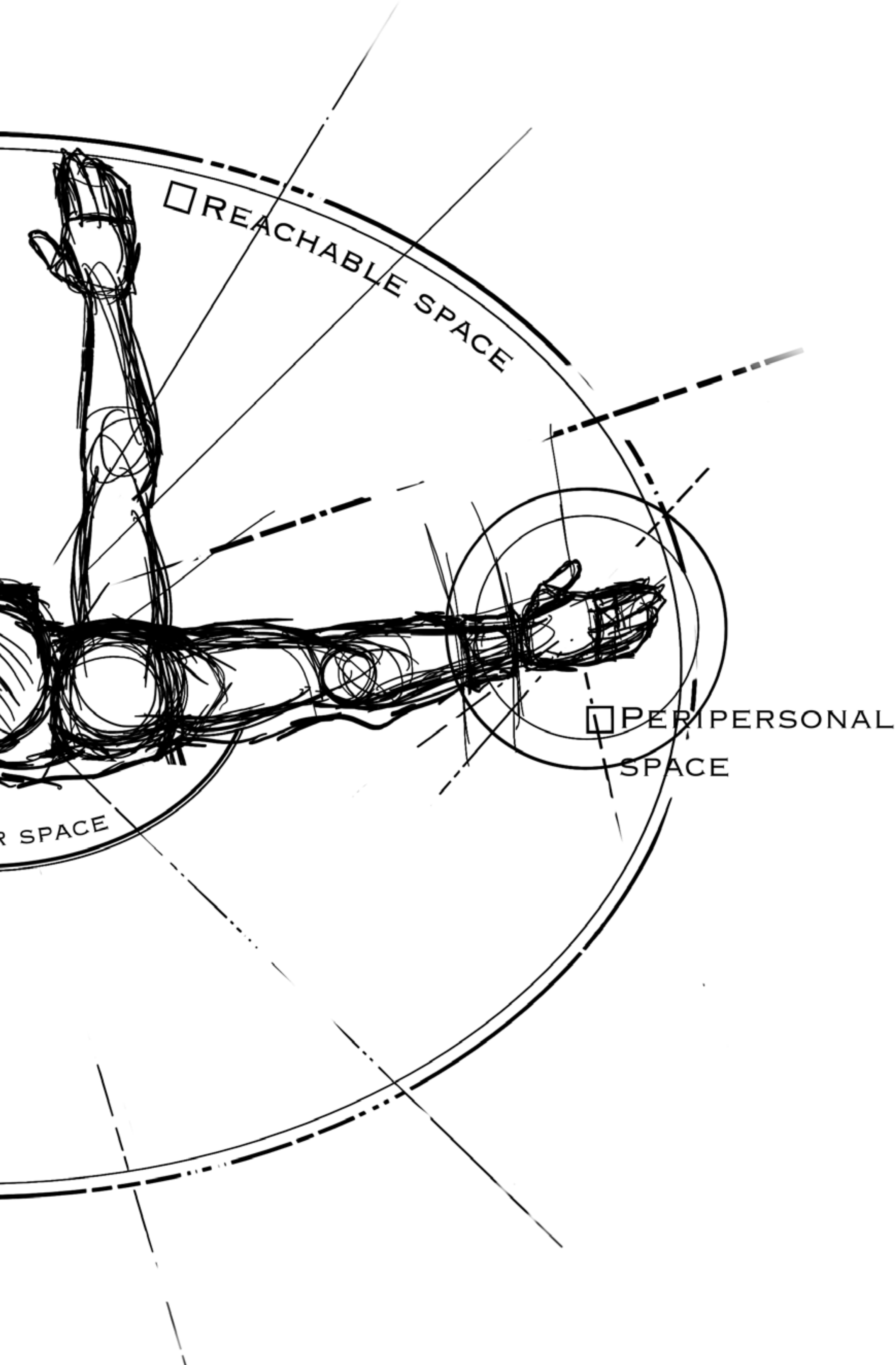
As can be seen from Figure 1A, the difference in the average of median RTs between A and V targets was larger for validly cued than for invalidly cued targets. It has previously been suggested that equal performance in different sensory modalities is important for MSI and often leads to the largest benefits 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 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 all cue types (Valid, Invalid, Center), there was a significant correlation between the difference in unimodal RTs and the amount of rMRE ($r = -.639$, $p < .001$). In addition, the difference in unimodal RTs was also significantly related to the average amount of RMI violation ($r = -.423$, $p = .001$). These correlations indicate that those participants whose difference between unimodal RTs was smaller showed a larger amount of rMRE and RMI violation. This finding is in line with the idea that equal performance for different sensory modalities leads to larger multisensory benefits of MSI (in this case faster RTs) as compared to unequal performance. Thus, IOR may modulate MSI by changing differences in unimodal processing times as it differentially affects processing in the auditory and visual sensory modality.

Discussion

The present study investigated whether IOR affects the integration of auditory and visual 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 delayed visual information processing at cued as compared to uncued locations.

The current findings seem to be in 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 reduced for exogenously attended compared to unattended locations. Therefore, one could argue that MSI should decrease at uncued relative to cued locations due to IOR in the current study, as the processing of, or responding to, the cued instead of the uncued location is impaired or slowed down. Yet, these seemingly contradicting findings can be explained by the fact that IOR was only observed for unimodal visual targets in the current study (in line with Ward, McDonald, & Lin, 2000). IOR for visual targets and the lack of IOR for auditory targets increased the difference in auditory and visual processing times at cued as compared to uncued locations. The relatively large difference in auditory and visual processing times at cued as compared to uncued locations likely decreased MSI. This idea is in line with the observation that the benefits of multisensory stimulation are most pronounced when performance in the different modalities is similar compared to when one modality is dominant (e.g., in terms of processing speed or estimation accuracy; 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 negligible. The strong correlation between the difference in unimodal RTs and the amount of rMRE and RMI violation that was observed in the current study is in line with this idea.

To conclude, the present study indicates that IOR (evoked by a visual exogenous cue) decreases audiovisual integration at cued locations. This modulation of audiovisual integration through IOR is likely driven by an increase in the difference between unimodal processing times at cued as compared to uncued locations.


☐ Reachable space
☐ Peripersonal space
r space

Chapter 5

Non-lateralized auditory input enhances averaged vectors in the oculomotor system

Van der Stoep, N., Nijboer, T.C.W., & Van der Stigchel, S. (2012).
Non-lateralized auditory input enhances
averaged vectors in the oculomotor system.
Experimental Brain Research, 221(4), 377-384.

Abstract

The decision about which location should be the goal of the next eye movement is known to be determined by the interaction between auditory and visual input. This interaction can be explained by the vector theory that states that each element (either visual or auditory) in a scene evokes a vector in the oculomotor system. These vectors determine the direction in which the eye movement is initiated. Because auditory input is lateralized and localizable in most studies, it is currently unclear how non-lateralized auditory input interacts with the vectors evoked by visual input. In the current study, we investigated the influence of a non-lateralized auditory non-target on saccade accuracy (saccade angle deviation from the target) and latency in a single-target condition in Experiment 1 and a double-target condition in Experiment 2. The visual targets in Experiment 2 were positioned in such a way that saccades on average landed in between the two targets (i.e., a global effect). There was no effect of the auditory input on saccade accuracy in the single-target condition, but auditory input did influence saccade accuracy in the double-target condition. In both experiments, saccade latency increased when auditory input accompanied the visual target(s). Together, these findings show that non-lateralized auditory input enhances all vectors evoked by visual input. The results will be discussed in terms of their possible neural substrates.

Introduction

Whereas people are accurate in making eye movements to a single visual target, eye movements are generally less accurate when executed to a single auditory target in the absence of visual information (Frens & Van Opstal, 1995; Zambarbieri, Schmid, Magenes, & Prablanc, 1982). Furthermore, saccade latencies are known to be shorter to an auditory stimulus than to a visual stimulus (Hughes, Reuter-Lorenz, Nozawa, & Fendrich, 1994). When one visual and one auditory stimulus are presented at the same location and at the same time, these signals are integrated and saccade reaction times are even shorter than to an auditory stimulus (i.e., intersensory facilitation, Frens, Van Opstal, & Van der Willigen, 1995; Hughes, Reuter-Lorenz, Nozawa, & Fendrich, 1994). The aim of the present study was to investigate the effect of a non-lateralized auditory stimulus on initial saccade direction in the presence of one or two peripheral visual targets.

When two visual targets are presented in close proximity, the endpoint of an eye movement toward these elements is generally not positioned on one of the two elements but rather on a location in between the two elements (i.e., the "global effect"; Findlay, 1982; Van der Stigchel & Nijboer, 2011). Generally, this effect is stronger for shorter saccade latencies and smaller distances between the two elements (Ottes, Gisbergen, & Eggermont, 1985). The global effect has been explained by the center of gravity account (Pitts & McCulloch, 1947). Pitts and McCulloch (1947) suggested that a saccade is initiated to the "center of gravity of the distribution of brightness" (p. 137). However, this explanation could not account for the finding that the global effect is only apparent when two stimuli are in close proximity. A more recent theory by Tipper et al. (1997) could account for these results. Whereas this theory was initially used to explain reaching behavior, it also became an explanation for the way in which eye movements are initiated (for a review, see Van der Stigchel, 2010). The theory proposes that elements in a visual scene are processed in parallel during initial perceptual analysis and evoke activity in neurons coding for a specific direction. Together, these neurons form a neuronal population and represent a vector that codes for the movement toward the location of an element in the visual scene. When two elements in a visual scene are in close proximity, the activity of neural populations that is evoked by these elements overlaps. Consequently, the vectors are averaged, which results in the initiation of saccades to the average direction (i.e., a global effect). Although a robust finding, saccade averaging does not occur in every trial, and a subset of eye movements is still directed to one of the two visual elements (Van der Stigchel & Nijboer, 2011). When two visual elements are further apart, the activity of neural populations does not overlap and the resulting vectors are not averaged. This results in saccades being initiated to one of the two elements (i.e., no global effect). When two visual elements are in close proximity, but one of the elements evokes a stronger vector (because, for example, this element is brighter), the direction of the initiated saccade will deviate on average more toward the element associated with the stronger vector (i.e., less strong global effect).

One of the brain areas that is thought to be responsible for saccade programming and is assumed to be involved in representing these population vectors is the superior colliculus (SC), a retinotopically organized saccade map area in the midbrain (Lee, Rohrer, & Sparks, 1988). The SC responds not only to stimuli presented in the visual modality but also to stimuli presented in other modalities (Sparks & Nelson, 1987). Because this important area in the oculomotor system appears to respond to stimuli of various modalities, previous studies have investigated the influence of multisensory input on eye movements. For example, Doyle and Walker (2002) showed that visual, auditory, and tactile distractors influenced the trajectory of saccades to a single visual target. Furthermore, in a study by Frens et al. (1995), a single visual stimulus and a single auditory stimulus were presented simultaneously. They observed a cross-modal global effect when a low-intensity visual stimulus and an auditory stimulus were presented simultaneously in close proximity. These results are in line with the findings by Lueck et al. (1990) who showed that the amplitude of a saccade to a horizontal auditory target on the left or right of fixation could be increased or decreased with a visual distractor in the same hemifield. These results suggest that stimuli from different modalities each evoke their own vector in the oculomotor system and are combined before an eye movement is initiated.

Although the influence of lateralized stimuli of different modalities is known, it is unclear how vectors are integrated when both lateralized visual and non-lateralized auditory information is presented. To this end, we measured eye movements to unimodal (visual) and bimodal (visual and auditory) stimuli under single- and dual-target conditions. The auditory stimulus was presented through headphones to the left and the right ears simultaneously. This allowed us to investigate how eye movements to visual stimuli are influenced by a non-lateralized auditory stimulus. We hypothesized that the auditory input in our experiments would not evoke an individual vector in the oculomotor system due to its non-lateralized nature. In contrast, we expected that the auditory input would be coupled with the vectors evoked by the lateralized visual targets, enhancing the activity of all vectors present in the oculomotor system. In Experiment 1, we measured saccade angle deviation and saccade latency in response to a single visual target with and without an additional non-lateralized auditory non-target. In terms of the population coding theory, we hypothesized that when a visual target is presented together with a non-lateralized auditory non-target, the vector elicited by the visual stimulus will be strengthened by the auditory input. A stronger vector would imply a higher probability to initiate saccades in the direction of the vector. However, because saccades to single visual targets are already quite accurate, no effect on saccade endpoints was expected. In Experiment 2, we investigated the influence of the non-lateralized auditory input on saccade angle deviation and latency when two visual targets are presented at the same time. The two visual stimuli were presented in such a configuration that we expected them to elicit a global effect. Under dual-target conditions, we expected a stronger global effect (i.e., a higher proportion of saccades being initiated to the average vector) in the presence of auditory information than when no auditory information is present. Based on Tipper's vector theory, we assume that auditory information enhances both vectors and the average vector will therefore be stronger compared to when no sound is presented (even though the average vector has the same direction in both conditions). As mentioned, saccade averaging does not occur on every trial and only occurs when saccade vectors are averaged. Therefore, the stronger averaged vector due to the presence of the auditory information will increase the probability that saccades are initiated to the average vector and will result in more trials in which the global effect is observed. Another possibility is that the auditory stimulus will automatically be associated with one of the two visual targets. This would lead to only one of the vectors being strengthened and consequently results in more saccades being initiated toward one of the two targets (i.e., a weaker global effect).

Experiment 1

Materials and methods

Participants

Six participants (mean age = 24.00, *SD* = 4.14 years; 3 male) were tested in this experiment for either money or course credits and had normal or corrected-to-normal visual acuity. All participants reported that they were able to hear the auditory stimulus. The experiment was performed in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants.

Apparatus

To record eye movements, a desktop-mounted Eyelink 1,000 system (SR Research Ltd., Canada) was used with a temporal resolution of 1,000 Hz and a spatial resolution of 0.01°. Each participant's head was stabilized with a chin rest. For all participants, the left eye was monitored. The monitor was positioned at a distance of 70 cm from the participants.

Stimuli, task, and procedure

Participants were tested in a dimly lit, sound-attenuated room. The experiment was presented on a 22-inch monitor (100 Hz, resolution: 1,024 9 768 px). The experiment consisted of three blocks, each containing 128 trials adding up to a total of 384 trials. Before each trial, a drift check was performed. Next, a fixation cross was shown with a duration of either 750, 1,000, or 1,250 ms (for each participant a pre-generated random list was used). The fixation cross was a black plus sign (0.9° 9 0.9°, 0.14 cd/m^2) positioned at the center of the screen. After the offset of the fixation cross, a black target circle (diameter = 0.9°, line width = 0.15°, 0.14 cd/m^2) was presented for 1,000 ms. Targets were positioned at an imaginary circle around the fixation point at the center of the screen (radius = 9.6°). Target locations were at 35° and 55° (upper right quadrant), at 125° and 145° (lower right quadrant), at 215° and 235° (lower left quadrant), and 305° and 325° (upper left quadrant) with 0° being the top position. In each trial, the target randomly appeared at one of the eight target locations. All stimuli were presented on a gray background (5.6 cd/m^2). On half of the trials, the visual target was accompanied by a 440-Hz square wave (100 ms, 70 dB SPL) presented through headphones (stereo) synchronous with the onset of the visual target. Unimodal and bimodal trials were randomized. After the target disappeared, only the gray background was presented for 200 ms before going to the next trial. Figure 1 shows the procedure for Experiment 1. Participants were instructed to fixate on the fixation cross and, as soon as it disappeared, to look as fast as possible to the element appearing on the screen. The duration of the experiment was approximately forty minutes.

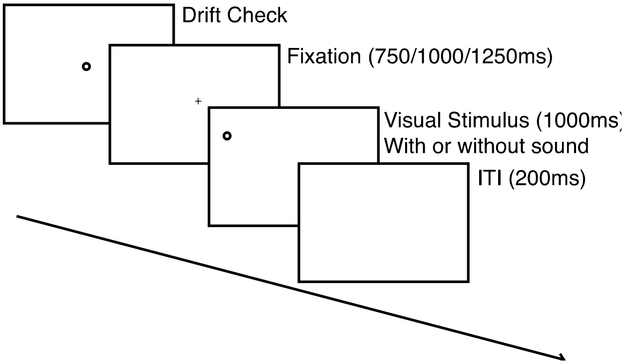


Figure 1. A schematic presentation of the procedure of Experiment 1.

Data analysis

Pre-processing

Saccade latency was defined as the interval between target onset and the initiation of the first saccadic eye movement. For each trial, the first saccade executed after target onset was analyzed. Saccade landing positions for all four quadrants on the screen were recalculated to the upper left quadrant. Saccades were filtered on amplitude (min. 4°, max. 30°) and onset latency (min. 60 ms, max. 500 ms). Only saccades that started within a radius of 1.8° (two times the diameter of the fixation cross) from the center of the fixation cross were included.

Saccade angle deviations were calculated as the absolute difference in saccade angle between the saccade landing position and the position in between the two possible target stimuli in a quadrant (10° from both targets), from saccade starting position. This way, saccades landing on the target closest to the horizontal axis had a relative angle deviation of -10° (absolute 10°), whereas saccades landing on the target closest to the vertical axis had a relative angle deviation of 10°. A schematic representation of saccade angle deviations is shown in Figure 2. Saccade angle deviations larger than two and a half standard deviations from the mean in a condition were removed from further analysis as they were considered outliers. Average saccade angle deviations for each condition were based on absolute saccade angle deviations. Absolute saccade angle deviations were averaged over all targets locations. After applying all filters, 6% of the trials in the sound condition and 5% of the trials in the no-sound condition were removed from further analysis.

Statistical analysis

A two-tailed paired-samples t-test on saccade latencies and saccade angle deviations between the sound and the no-sound conditions was done. To test whether the variance in absolute saccade angle deviations differed between the sound and the no-sound conditions for each participant, a paired-samples t-test on the average variances in saccade angle deviation between the sound and the no-sound conditions was performed.

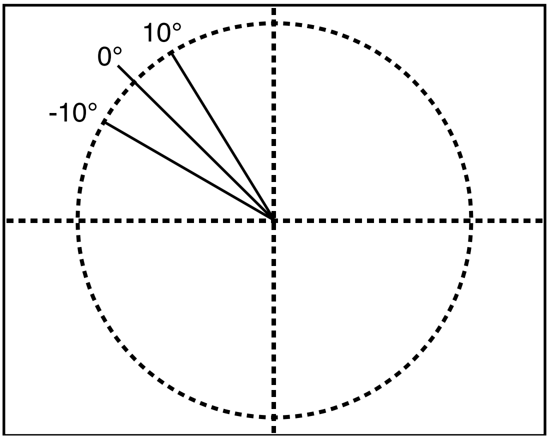


Figure 2. A schematic representation of the saccade angle deviation measure that is used in Experiment 1 and 2.

Results

Saccade latencies

There was a significant difference in saccade latency between the sound and the no-sound conditions [$t(5) = 6.230$, $p < .005$]. The average saccade latency was shorter for the no-sound condition ($M = 229$ ms, $SD = 6.8$ ms) than for the sound condition ($M = 240$ ms, $SD = 3.7$ ms).

Saccade angle deviation

There was no significant difference in average absolute saccade angle deviation from the target between the sound and the no-sound conditions [$t(5) = .097$, $p > .90$]. The average saccade angle deviation was 10.30° ($SD = .52°$) for the sound condition and 10.29° ($SD = .45°$) for the no-sound condition. The distributions of relative saccade angle deviations for the two possible target locations in a quadrant in the sound and the no-sound conditions are shown in Figure 3.

There was no significant difference in absolute saccade angle deviation variance between the sound (23.897) and the no-sound conditions (24.226) [*t*(5) = .142, *p* = .893]. This means that the auditory input did not cause saccades to single targets to be more accurate.

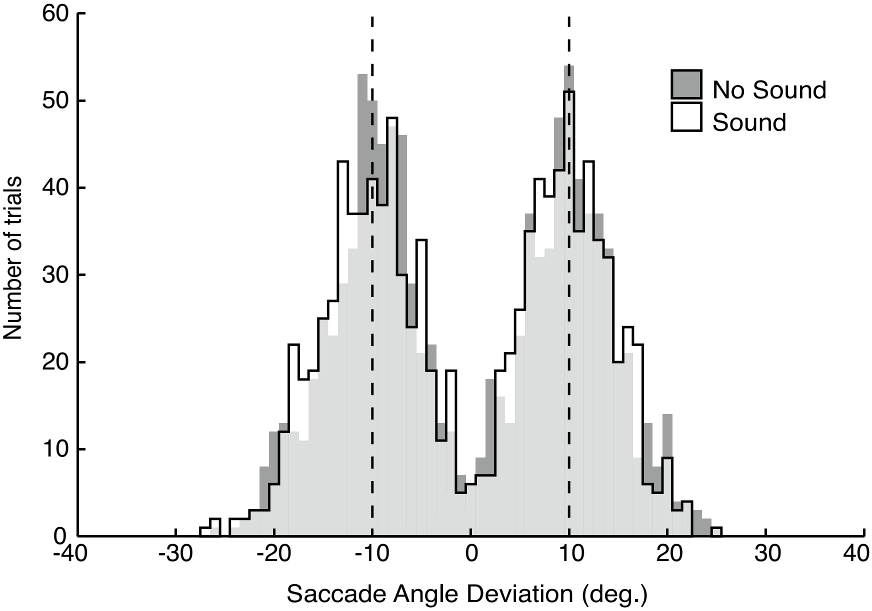


Figure 3. The distribution of relative saccade angle deviations for the no-sound (grey) and the sound condition (transparent white with black outline) of Experiment 1. The dashed lines indicate the two possible target locations.

Discussion

The non-lateralized auditory non-target did not influence primary saccade accuracy to a single visual target as measured by absolute saccade angle deviation. This can be explained in terms of Tipper's vector theory. The visual target would elicit a vector in the oculomotor system in both the unimodal and the bimodal conditions. In the bimodal condition, the auditory input enhanced the vector, increasing the probability that a saccade is initiated in the direction of the average vector and possibly saccade accuracy.

Although this effect was not observed, this can be explained by a ceiling effect of saccade accuracy to a single visual target. The auditory input did influence saccade latency to single visual targets. Saccade latencies to bimodal targets were significantly longer than to unimodal visual targets. As mentioned previously, intersensory facilitation has been frequently reported during bimodal stimulation (e.g., visual and auditory), resulting in shorter response times or shorter saccade latencies compared to unimodal stimulation (in this case, visual only) (Colonius & Diederich, 2004; Frens, Van Opstal, & Van der Willegen, 1995; Gielen, Schmidt, & Van den Heuvel, 1983). Intersensory facilitation, however, is subject to several rules. Perhaps the most important of these rules are the spatial and the temporal rules (Colonius & Arndt, 2001). These rules suggest that intersensory facilitation is stronger when an auditory stimulus and a visual stimulus are closer to each other in space and time. In Experiment 1, the visual and auditory stimuli were aligned in time. However, it is likely that the auditory stimulus was not associated with one specific location, as it was presented through headphones. This might have resulted in an absence of intersensory facilitation.

To further investigate the influence of non-lateralized auditory input on initial saccade direction, we presented two visual targets instead of one in Experiment 2. Because saccade accuracy is diminished when two visual targets are presented in close proximity (a global effect/saccade averaging), there is more room for change in saccade accuracy when the visual stimulus is accompanied by auditory input. An enhancement of the vectors to the two visual targets would result in a stronger average vector and a higher probability of saccades landing in the average direction of the visual targets and thus a stronger global effect. Because auditory input did not change the variance in saccade angle deviation in a single-target condition, a stronger global effect in the sound condition can be explained by a stronger average vector, but not by differences in saccade angle deviation variance between the sound and the no-sound condition.

Experiment 2

Materials and methods

Participants

Eleven participants (mean age = 27.8, *SD* = 4.9 years; 7 males) were tested in this experiment for either money or course credits. All participants had normal or corrected-to-normal visual acuity and reported that they were able to hear the auditory stimulus. The experiment was performed in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants.

Apparatus

The same apparatus as described under Experiment 1 was used.

Stimuli, task, and procedure

The same experimental setup as in Experiment 1 was used. The only difference between Experiments 1 and 2 is that instead of one visual target, two visual targets were presented in each trial. The 8 stimulus locations were the same as in Experiment 1, but the two visual targets were always presented in the same quadrant of the screen.

Data analysis

Pre-processing

The same filters and conditions as in Experiment 1 were used to select saccades. In this experiment, a global effect could occur, so we expected saccades to land on average in between the two target locations (between -10° and 10°). After applying all filters, 5% of the trials in the sound condition and 6% of the trials in the no-sound condition were removed from further analysis.

Statistical analysis Because the global effect tends to be stronger at shorter saccade latencies, we did a repeated measures analysis of variance (ANOVA) with latency bin (three bins) and modality (unimodal vs. bimodal) as independent variables and saccade angle deviation as dependent variable. This allowed us to investigate not only a main effect of modality but also whether there was an interaction between saccade accuracy and saccade latency. In addition, a two-tailed paired-samples t-test between the sound and the no-sound conditions was performed on saccade latencies.

Results

Saccade angle deviation

Mauchly's test of sphericity indicated that the assumption of sphericity had been violated for bin and the interaction of condition and bin. Therefore, we used Greenhouse–Geisser-corrected values for the results of the repeated measures ANOVA of these variables. The repeated measures ANOVA showed a significant main effect of condition [$F(1,10) = 6.272$, $p < .05$], but not of bin [$F(1.006,10.064) = .083$, $p > .05$, $\varepsilon = .503$]. Furthermore, the interaction between bin and condition was not significant [$F(1.024,10.243) = .000$, $p > .9$, $\varepsilon = .512$]. Average saccade angle deviations from the center between the two visual stimuli were smaller in the sound condition ($M = 6.27°$, $SD = 1.81°$) than in the no-sound condition ($M = 6.63°$, $SD = 2.14°$), indicating that the global effect was stronger in the sound than in the no-sound condition. The main effect of bin did not reach significance, which shows that saccade angle deviations were on average not larger or smaller for different saccade latencies. The interaction between condition and latency did not reach significance, indicating that the auditory stimulus did not influence saccade angle deviation differently in the sound than in the no-sound condition. The average saccade angle deviation for each bin for the sound and the no-sound conditions is shown in Figure 4. The distribution of relative saccade angle deviations is shown in Figure 5 and shows a larger amount of saccades made to the average direction of the targets in the sound condition compared to the no-sound condition. A perfect averaging of vectors would result in an absolute average saccade deviation angle of 0°, while this would be 10° if saccades always landed on one of the two targets.

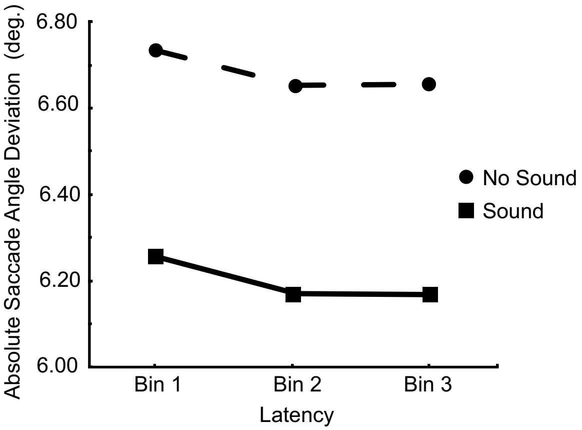


Figure 4. Mean saccade angle deviations for the three latency bins in the sound (straight line) and the no-sound condition (dashed line).

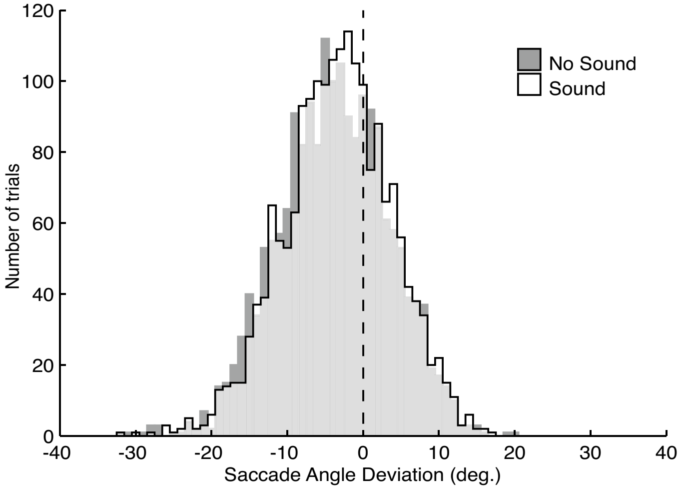


Figure 5. The distribution of relative saccade angle deviations for target one and two in the no-sound condition (grey) and the sound condition (transparent white with black outline) of Experiment 2. The dashed line indicates the perfect average of the two target locations. The distribution of the sound condition has more values around a saccade angle deviation of zero than the no-sound condition, which results in a stronger global effect.

Saccade latencies

A two-tailed paired-samples t-test again revealed a significant difference in saccade latencies between the sound and the no-sound conditions [$t(10) = 8.811$, $p < .001$]. Saccade latencies were shorter in the no-sound condition ($M = 223$ ms, $SD = 16$ ms) compared to those in the sound condition ($M = 240$ ms, $SD = 19$ ms).

Discussion

A global effect was found in both the sound and the no-sound condition. Moreover, as we expected, the global effect was significantly stronger in the sound condition (i.e., closer to the center between the two targets). This supports the idea that non-lateralized auditory input enhances vectors that arise from visual input. An enhancement of the vectors does not result in a different average vector, but it does enhance the probability that a saccade is initiated to the average vector. This is also visible in the distribution of saccade angle deviations shown in Figure 5. More saccades landed close to the center between the two targets in the sound compared to the no-sound condition.

It could be argued that the global effect in the sound condition is stronger because of a broader distribution of saccade endpoints in the sound condition, irrespective of the presentation of the two elements. This would then result in a larger number of saccade endpoints landing in between the two targets in the sound condition. The average of saccade angle deviations in the sound condition would then be closer to 0° (a stronger global effect) than in the no-sound condition. However, the distribution of relative saccade angle deviations in Figure 5 does not support this explanation, because it shows a clear unimodal distribution of relative saccade angle deviations. In addition, there was no difference in variance of absolute saccade angle deviations between the sound and the no-sound conditions in Experiment 1; therefore, this cannot account for the stronger global effect in the sound condition.

It has been shown that the global effect is generally stronger at shorter saccade latencies (Findlay, 1982; Van der Stigchel & Nijboer 2011). However, latency cannot account for the current findings as we found significantly longer saccade latencies and a significantly stronger global effect in the sound condition compared to the no-sound condition.

General discussion

In the current study, we showed that a non-lateralized auditory input enhances vectors in the oculomotor system that arise from visual input in a double-target condition. Additionally, we did not find an enhancement of the visual vectors by the auditory input in a single-target condition, which can be explained by a ceiling effect of saccade accuracy to single visual targets. In Experiment 2, we did find an enhancement of the visual vectors by the auditory input in the double-target condition, as indicated by a stronger global effect during bimodal stimulation compared to unimodal stimulation. Both results can be explained in terms of the vector theory (Tipper, Howard, & Jackson, 1997). The theory states that a visual stimulus evokes a neural population that produces a vector. Subsequently, an eye movement is initiated in the direction of the vector. When two visual elements are close enough in space, the neural populations overlap and the vectors are averaged, which results in the initiation of a saccade into the average direction. We argue that the vectors that are evoked by the visual input in Experiments 1 and 2 were strengthened by the auditory input. An enhancement of the vectors resulted in a stronger global effect during bimodal stimulation in Experiment 2, because the average vector was also stronger compared to the no-sound condition.

This stronger vector implies a higher probability that a saccade is initiated in the direction of the vector.

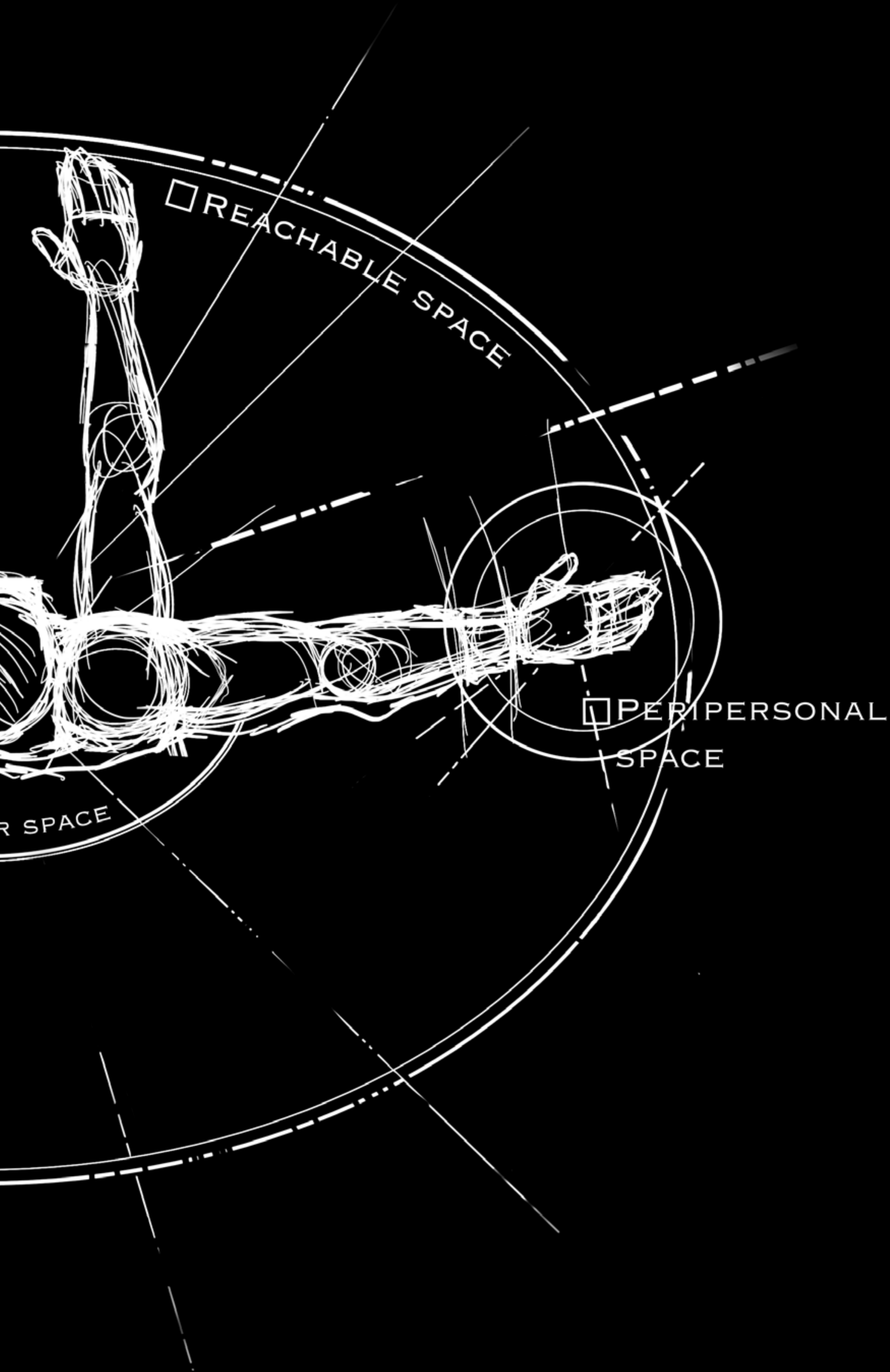
The shorter saccade latency in the no-sound condition compared to the sound condition in Experiments 1 and 2 can be explained in several ways. First, as has been mentioned previously, the spatial and temporal rules are specifically important for intersensory facilitation to occur (Frens, Van Opstal, & Van der Willigen, 1995). Our experimental setup did adhere to the temporal rule in the sound condition, but not the spatial rule. This spatial incongruity could have caused longer saccade latencies. Although it might explain why the saccade latencies in the sound condition are not shorter compared to the no-sound condition, it does not explain why the saccade latencies in the sound condition are longer than those in the no-sound condition. The difference in saccade latencies between the sound and the no-sound conditions in both experiments may be explained by the fact that the auditory stimulus was presented through headphones to both ears simultaneously with the same loudness, and therefore, the auditory stimulus might be associated with the entire visual scene. This way, the auditory stimulus might have acted as an auditory fixation stimulus, prolonging the visual fixation stimulus in the auditory modality after the offset of the visual fixation stimulus. A study by Taylor et al. (1999) showed that an auditory fixation stimulus is, just as a visual stimulus, able to produce longer latencies in an overlap condition compared to a gap condition. This might explain why the saccade latencies in the sound condition are slower compared to those in the no-sound condition.

The lack of an interaction between latency and saccade angle deviation appears to contradict previous findings (e.g., Ottes, Van Gisbergen, & Eggermont, 1985), which showed a decline in the global effect for longer latencies. In these paradigms, however, there was a strong top-down component, because it was clear which of the two visual elements was the target and which was the distractor. As it is known that top-down selection is stronger for longer latencies (Van Zoest, Donk, & Theeuwes, 2004), this explains why the global effect generally decreases later in the selection process. However, in our study, we limited top-down selection by giving a general task instruction ("Look to the elements that appear on screen as fast as possible"). Therefore, the two stimuli had the same status in the oculomotor selection process. For that reason, we assume that selection in our study is mostly

based on bottom-up information. The fact that we did not find an interaction between saccade angle deviation and saccade latency can then be explained by a lack of top-down selection, resulting in a global effect for all latencies These results are in line with a previous study on saccade averaging without a strong top-down selection (Van der Stigchel, Heeman, & Nijboer, 2012).

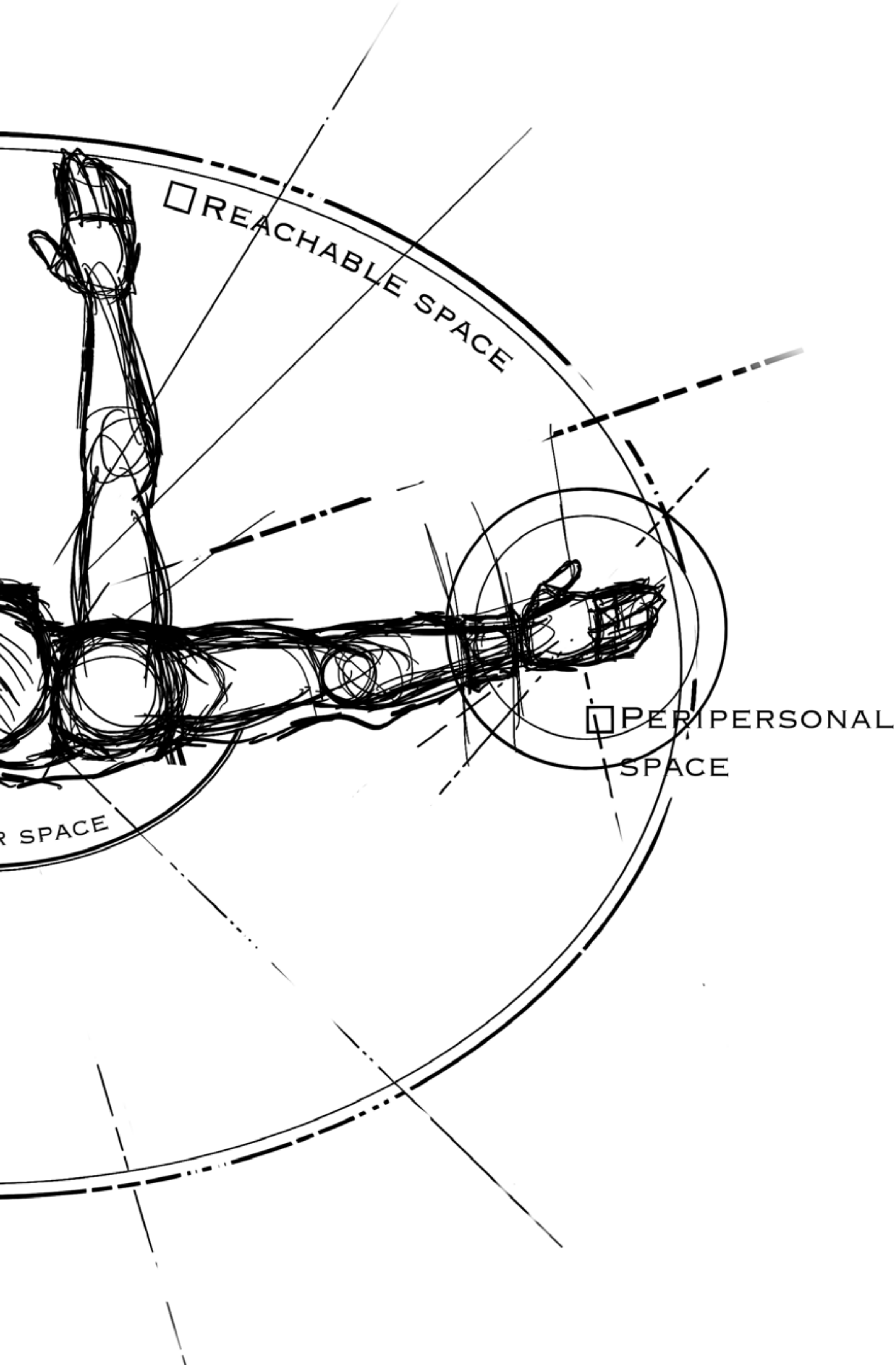
With respect to the neural correlates of these results, it is likely that the SC is responsible for the observed findings. King and Palmer (1985) showed that the response of a specific type of neurons in the superficial layers of the SC that responded exclusively to visual stimuli was enhanced by an auditory stimulus. In the deeper layers, however, they found neurons responding to both visual and auditory stimuli, which showed an enhanced or depressed response when visual and auditory stimuli were presented together. These multimodal neural interactions are bound to the spatial and temporal rules and might underlie the intersensory facilitation effect found in many studies (e.g., Gielen, Schmidt, & Van den Heuvel, 1983; Hughes, Reuter-Lorenz, Nozawa, & Fendrich, 1994; Frens et al. 1995; Colonius & Arndt, 2001; Colonius & Diederich, 2004). As visual and auditory receptive fields of unisensory and multisensory cells in the SC are topographically arranged, we hypothesize that the visual input in our study evoked activity in those cells in the SC whose receptive fields corresponded to the specific locations of the targets in the visual field. However, as the auditory input was non-lateralized and presented through headphones, the auditory input might have evoked activity in a wider range of cells, together covering a large part of the auditory field with their auditory receptive fields. Consequently, two types of cells might have had an enhanced response: auditory-enhanced visual cells (not responding to auditory input in isolation) and multisensory cells (responding to both visual and auditory input in isolation). Yet, the activity in these types of cells is not directly related to the production of eye movements. Eye movements are generated in saccade-related neurons in the deeper layers of the SC, and only a small amount of neurons in the superficial layers of the SC that respond to visual input directly cause activity in saccade-related neurons in the deeper layers of the SC (Mays & Sparks, 1980). Thus, activity in neurons sensitive to visual and auditory input does not necessarily lead to saccades being initiated to the corresponding locations in space. Lee et al. (1988) investigated population coding of saccadic eye movements by neurons in the deeper layers of the SC and provided support for the idea that saccades are initiated in the direction of the average activity of a population of saccade-related neurons. The generation of saccades is influenced by both bottom-up and top-down input (Van Zoest, Donk, & Theeuwes, 2004), but we minimized top-down influences in the current study with our specific task instruction. Therefore, we hypothesize that the bottom-up activity in the described auditory-enhanced and multisensory neurons greatly influenced the direction in which eye movements were initiated. An increased activity in these neurons during bimodal input compared to unimodal visual input could be the basis of the larger theoretical vectors and of the larger average vector in the sound condition in Experiment 2.

To summarize, the results of Experiments 2 indicated that when primary saccades are less accurate, a non-lateralized sound enhances the vectors evoked by the visual stimuli and thus the average vector, resulting in a higher proportion of saccades landing in between the two targets. While we were not able to show that vectors evoked by single visual targets were enhanced by auditory input, this might be explained by the fact that saccades to single visual targets are quite accurate. The global effect might be a useful phenomenon in investigating complex interactions between multiple stimuli from different modalities because the global effect itself is the result of the competition between multiple stimuli.


☐Reachable space
☐Peripersonal space
r space

Part II

Multisensory interactions in the depth plane


☐ Reachable space
☐ Peripersonal space
r space

Chapter 6

Multisensory interactions in the depth plane in front and rear space: A review

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.

Abstract

In this review, we evaluate the neurophysiological, neuropsychological, and psychophysical evidence relevant to the claim that multisensory information is processed differently depending on the region of space in which it happens to be presented. We discuss how the majority of studies of multisensory interactions in the depth plane that have been conducted to date have focused on visuotactile and audiotactile interactions in frontal peripersonal space and underline the importance of such multisensory interactions in defining peripersonal space. Based on our review of studies of multisensory interactions in depth, we question the extent to which peri- and extra-personal space (both frontal and rear) are characterized by differences in multisensory interactions (as evidenced by multisensory stimuli producing a different behavioral outcome as compared to unisensory stimulation). In addition to providing an overview of studies of multisensory interactions in different regions of space, our goal in writing this review has been to demonstrate that the various kinds of multisensory interactions that have been documented may follow very similar organizing principles. Multisensory interactions in depth that involve tactile stimuli are constrained by the fact that such stimuli typically need to contact the skin surface. Therefore, depth-related preferences of multisensory interactions involving touch can largely be explained in terms of their spatial alignment in depth and their alignment with the body. As yet, no such depth-related asymmetry has been observed in the case of audiovisual interactions. We therefore suggest that the spatial boundary of peripersonal space and the enhanced audiotactile and visuotactile interactions that occur in peripersonal space can be explained in terms of the particular spatial alignment of stimuli from different modalities with the body and that they likely reflect the result of prior multisensory experience.

Introduction

Traditionally, researchers have tended to study the spatial senses (e.g., vision, audition, and touch/proprioception) in isolation from one another[6]. That said, the last few decades have seen something of an explosion of interest in the topic of multisensory perception (see Figure 1). Much of this interest has been inspired by neurophysiological studies documenting the existence of neurons in animals such as macaques and cats that are responsive to stimuli from more than one sensory modality (e.g., Bruce, Desimone, & Gross, 1981; Meredith, Nemitz, & Stein, 1987; Meredith & Stein, 1986; see Stein & Meredith, 1993 and Stein & Stanford, 2008, for reviews). What is more, on closer inspection, many of these neurons have been found to have interesting (that is, non-linear) response properties.

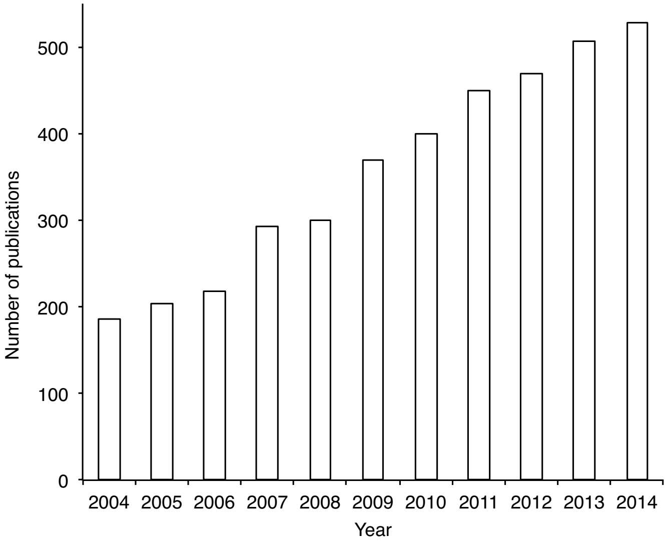


Figure 1. Number of articles indexed in PUBMED over the last decade that have included the keywords 'multisensory', 'crossmodal', or 'cross-modal' in the title or abstract (see also Murray, Spence, & Harris, 2013; Stein et al., 2010).

In many cases, the relative and/or absolute spatial location from which the stimuli in the different sensory modalities were presented has proven to be important in terms of determining the kinds of multisensory interactions (and neuronal response

---

[6] Largely ignoring the chemical senses (of smell and taste) altogether, as we will do here.

properties) that have been reported. So, for example, neurophysiological research has demonstrated that in those situations in which the auditory and visual receptive fields (RFs) of a bimodal neuron overlap, multisensory response enhancement primarily occurs when the auditory and visual stimuli are spatially aligned. When a pair of stimuli is spatially misaligned (as when visual and auditory stimuli are presented from different azimuthal positions), and, for example, the visual stimulus is presented just outside of the visual RF of a bimodal neuron while the auditory stimulus is presented within the auditory RF of the bimodal neuron, multisensory response depression typically ensues (Stein & Meredith, 1990). The relation between the azimuthal alignment of stimuli and the strength of any multisensory interactions that are documented is commonly referred to as "the spatial rule" (e.g., Holmes & Spence, 2005).

In humans, however, the available evidence suggests that this rule is very much task-dependent. That is, the spatial rule is more often observed to modulate performance in those tasks that are in some sense spatial as compared to those tasks in which the spatial location of the stimuli is entirely task-irrelevant to the task being performed (see Spence, 2013, for a review). In other words, the principles of multisensory integration that have often been observed in neurophysiological studies in (typically anaesthetized) animals cannot always necessarily readily be observed in behavioral studies in awake humans.

Varying the distance in depth between multisensory stimuli and the observer has also been shown to modulate the responsiveness of at least certain bimodal neurons. So, for example, some (pericutaneous) neurons in the macaque only appear to respond to somatosensory stimuli delivered to the body surface and to visual stimuli presented from a location that lies within reach, but not to the very same visual stimuli when presented beyond the animal's direct reach (e.g., Graziano & Gross, 1994; Rizzolatti, Scandolara, Matelli, & Gentilucci, 1981). A similar distance-dependent boundary has also been observed in the responsiveness of trimodal neurons with auditory stimuli that were presented from close to, vs. further away from, the animal's head (Graziano, Reiss, & Gross, 1999). Such results therefore suggest that the spatial alignment of stimuli presented in different sensory modalities in terms of their depth may be just as important as their alignment in azimuthal space when it comes to evoking a response from this type of neuron.

The scientific data would indeed appear to suggest that different regions of space are coded differently by the brain (Previc, 1990, 1998), but this does not seem to be reflected in the way in which we subjectively experience the world around us, namely as a seamless whole. Given this rather curious disconnect, it would seem sensible to try and gain a further understanding of multisensory perception in depth. The importance of investigating multisensory interactions in different regions of space becomes all the more apparent when one considers the enormous amounts of multisensory information that we receive from different locations (e.g., distances beyond the reach of peripersonal space) and which we perceive on a daily basis. We may not think about it, but during the daily drive to work, for example, the most crucial sensory information necessary to drive safely comes from frontal extrapersonal space. Although we also receive sensory information from peripersonal space (e.g., think only of the dashboard lights and alerts, tactile stimulation from the driving seat, steering wheel, and feedback from the gas, break, and clutch pedal), sensory information from frontal and rear extrapersonal space (the latter seen via the rearview mirror, or else perhaps heard) is crucial in terms of our ability to navigate successfully through the environment[7] (see Previc, 2000, for an example of applying knowledge about 3-D spatial information processing to the design of aircraft controls; see Spence & Ho, 2008, and Ho & Spence, 2009, for the application of knowledge of multisensory processing to the design of warning signals in the context of driving). It is currently unclear, however, under which circumstances sensory information from the different senses interact in terms of their spatial (mis)alignment in depth and/or any differences in their lateral position.

Although the investigation of the multisensory interactions taking place in depth has received a growing amount of research attention in recent years, the majority of studies have tended to look at multisensory interactions in two-dimensional (2-D) space (that is, experimenters have mostly varied only the azimuth and, on occasion, the elevation of the stimuli, while keeping their distance from the observer fixed; e.g., Frens, Van Opstal, & Van der Willegen, 1995; Stevenson et al., 2012; Ten Brink, Nijboer, Van der Stoep, & Van der Stigchel, 2014). In fact, in one oft-cited edited volume on the topic of crossmodal space and crossmodal attention (Spence & Driver, 2004), variations in depth rarely get mentioned at all. On those occasions where the authors do talk about variations in depth, it is mainly in the context of the coding of peripersonal space (e.g., in the context of tool-use, and distance-dependent extinction).

In this review, we evaluate the growing body of cognitive neuroscience research that has documented the nature, and peculiarities, associated with multisensory interactions in depth in front and rear space. Below, we review studies of both crossmodal spatial

---

[7] Given that some visual RFs in (stationary) monkeys have been observed to extend in depth as the speed of an approaching visual stimulus increased (Fogassi et al., 1996), one could argue that an extension of RFs in depth may also depend on the speed of movement with which humans move through their environment. This might result in an increase in the extent of peripersonal space and the observed multisensory interactions in this region (see Section 3).

attention and of multisensory integration[8] (see Spence & Driver, 2004). We compare and contrast the results of those studies that have presented their experimental stimuli in both peripersonal and extrapersonal frontal space, as well as in those more recently-discovered regions, referred to as near (peripersonal), and far (extrapersonal) rear space (see Occelli, Spence, & Zampini, 2011, for a review). According to the definition adopted here, peripersonal space can be defined as the space directly surrounding different parts of the body as defined by receptive fields of bimodal and trimodal neurons (Rizzolatti, Scandolara, Matelli, & Gentilucci, 1981), and extrapersonal space as being the space that is further away from the body (Previc, 1998). In most of the studies that have investigated multisensory interactions involving the tactile modality, peripersonal space is referred to as the space that is reachable (e.g., Occelli, Spence, & Zampini, 2011). In such studies, tactile stimuli are often delivered to the hands while the other parts of the skin surface are largely ignored (although a small number of studies have presented tactile stimuli to, for example, the earlobes, cheeks, abdomen, feet, or shin, see for example Schicke, Bauer, & Röder, 2009; Wan, Spence, Mu, Zhou, & Ho, 2014; and Ho & Spence, 2014; see Gallace & Spence, 2014, for a review). It is important to note that similar body-related regions of space have been observed for other parts of the body than the hands, such as the head, trunk, and arms (e.g., Fogassi et al., 1996; Graziano & Gross, 1994). When it comes to rear peripersonal space, we refer to the region lying directly behind the head and shoulders of the observer and extending out to a distance of approximately 20– 30 cm from the body. Beyond this lies the region known as rear extrapersonal space (see Figure 2 for a bird's-eye view of the putatively functionally distinct regions of space that will be discussed in this review). Furthermore, we will also cover the differences and similarities in multisensory interactions that have now been observed between stimuli that are presented in either the same or different regions of space.

We start by discussing those studies that have investigated visuotactile and audiotactile interactions in frontal peripersonal space and the flexibility of the extent of the peripersonal boundary. Given that tactile perception is fixed to the body, most

---

[8] Although the difference between these two phenomena is undoubtedly a topic of keen scientific debate (see, for example, McDonald, Teder-Sälejärvi, Ward, & Lin, 2001 and Spence, 2010, pp. 183–184), differences in the timing of the stimuli presented to different sensory modalities could potentially be used to differentiate between these two processes. So, for example, the most pronounced exogenous spatial cuing effects have typically been demonstrated with cue-target onset asynchronies (SOAs) of between 50 and 200 ms, whereas multisensory integration is often most pro- nounced with close temporal proximity (e.g., centered roughly on physical syn- chrony). Thus, multisensory interactions occurring with stimulus intervals of 50—100 ms, say, could therefore easily be explained in terms of both multisensory in- tegration and crossmodal exogenous shifts of spatial attention (for more on the interaction between exogenous attention and multisensory integration see, for example, Vroomen, Bertelson, & de Gelder, 2001b; Santangelo & Spence, 2007; Spence & Santangelo, 2009; Van der Stoep, Van der Stigchel, & Nijboer, 2015). While we most certainly agree that it is important to try to disentangle these empirical phenomena, we feel that discussing studies of both multisensory integration and exogenous crossmodal attention provides relevant insights in terms of the understanding of multisensory interactions and the boundaries in depth in front and rear space.

observations of stronger audio- or visuo-tactile interactions in peripersonal space can be explained by the particular spatial alignment of unimodal stimuli in depth (i.e., alignment with the body in the case of multisensory interactions involving touch). No such asymmetry in depth would be expected for audiovisual interactions. Therefore, in order to gain a fuller understanding of the multisensory interactions that occur in the different regions of space, we also discuss those audiovisual interactions that have been documented in frontal peripersonal and extrapersonal space. Lastly, we also discuss those multisensory interactions that have been reported in rear space. We conclude by summarizing our observations and suggesting a number of key principles that may help to define and understand multisensory interactions in depth in front and rear space.

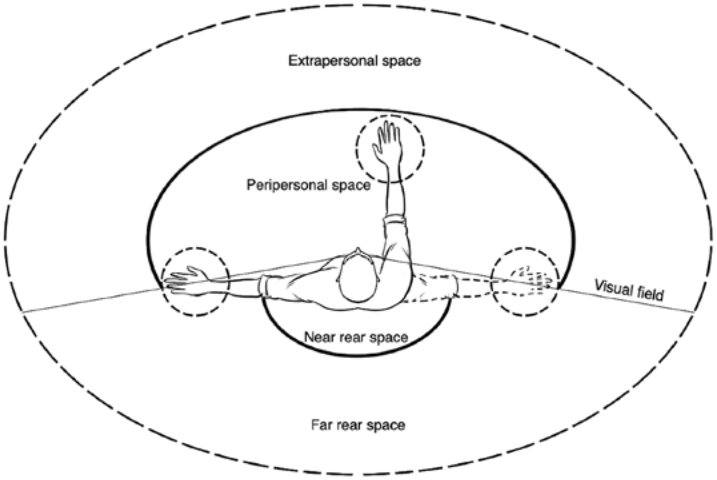


Figure 2. Bird's-eye view of the different regions of space discussed in this review. The dashed circles around the hands represent just one of the various body-part related regions of multisensory frontal peripersonal space that have been documented in monkeys (e.g., Rizzolatti, Scandolara, Matelli, Gentilucci, 1981).

2. Visuotactile and audiotactile interactions in frontal peripersonal space

The special role of the space directly adjacent to the body in terms of those multisensory interactions involving the stimulation of the skin (especially visuotactile and audiotactile interactions) would seem to have been inspired in large part by the results of those neurophysiological studies documenting the existence of bimodal and trimodal neurons in macaques that specifically respond to stimuli presented close to the body, but not to those stimuli that happen to be presented further from the body (see, for example, Canzoneri, Magosso, & Serino, 2012; Farnè & Làdavas, 2002; Graziano & Gross, 1994; Graziano, Reiss, & Gross, 1999; Rizzolatti, Scandolara, Matelli, & Gentilucci, 1981). As for the functions of such enhanced multisensory interactions in this region of space, the integration of visual/auditory, proprioceptive, and somatosensory stimuli appear to contribute to the efficient guidance of actions and defensive behaviors in response to those stimuli/events that are observed in peripersonal space (see Holmes & Spence, 2004, for a review; De Paepe, Crombez, Spence, & LeGrain, 2014).

2.1. Neuropsychological studies of visuotactile and audiotactile interactions in peripersonal space

Support for the importance of audiotactile and visuotactile multisensory interactions in defining peripersonal space has come from neurophysiological studies. Neuropsychological evidence appears to support the notion that space is divided into several separable (or distinct) regions. For example, deficits in the orienting of visuospatial attention in peripersonal space have been observed after brain damage, while visuospatial orienting in extrapersonal space remains intact (or vice versa, e.g., Aimola et al., 2012; Halligan & Marshall, 1991; Van der Stoep et al., 2013). The results of studies of audiotactile extinction also point to a similar conclusion; Namely, that differences in audiotactile extinction have been observed in right brain-damaged patients when auditory stimuli are presented from either close to, or further away from, the patient in either frontal or rear space (Farnè & Làdavas, 2002). Audiotactile extinction was significantly more pronounced under those conditions in which the auditory stimuli were presented closer to the patient4 (that is, from peripersonal space; see, for example, Làdavas et al. (1998), for similar results with visuotactile extinction). Additionally, differences in the magnitude of crossmodal extinction between near and far space were found to be larger for those audiotactile stimuli that were presented in rear space as compared to those that were presented in frontal space. Taken together, then, these results would appear to suggest that both audiotactile and visuotactile interactions are especially pronounced in the space directly surrounding the body.

2.2. Psychophysiological studies of visuotactile and audiotactile interactions in peripersonal space in healthy individuals

Given the clear distinction between peripersonal and extrapersonal space documented in both neurophysiological and neuropsychological studies, it would seem plausible to suggest that such a border might also influence the kinds of crossmodal and multisensory behavioral effects that are seen in healthy individuals under the appropriate experimental conditions. Relevant to this notion, Canzoneri et al. (2012) presented dynamic auditory stimuli that simulated the approach or receding of a sound source using a pair of loudspeakers: one sound source was positioned in peripersonal space (close to the hand), while the other was placed in extrapersonal space (~100 cm from the participant). The tactile targets in this study were presented to the participant's right hand at various times relative to the approaching or receding sound.

Canzoneri et al.'s (2012) results demonstrated the existence of a boundary in space. In particular, the participants in this study responded most rapidly to the tactile targets when the simulated sound location was situated close to their hand as compared to when it appeared to be further away. Similar results have also been reported when it comes to those multisensory interactions taking place between vision and touch (Gray & Tan, 2002), and when measuring motor evoked potentials in response to an approaching visual stimulus (Makin, Holmes, Brozzoli, Rossetti, & Farnè, 2009)[9]. It would therefore seem that both dynamic auditory and visual stimuli that give the impression of approaching an observer's body are especially efficient in terms of enhancing an observer's responses to tactile stimulation presented to the body surface (Makin, Holmes, & Zohary, 2007).

3. The flexibility of the extent of frontal peripersonal space

The border between peripersonal and extrapersonal space that can be defined by the strength of visuotactile and audiotactile interactions appears to be modified depending on the position of the body in space. So, for example, consider a study of audiotactile interactions with tactile stimulation of the hand. Responses are generally faster when the hand is placed near a sound source as compared to when the stimulated hand is moved away from it (e.g., Serino, Canzoneri, & Avenanti, 2011; but see Zampini, Torresan, Spence, & Murray, 2007; Kennett & Driver, 2014). Such results are in line with the idea that unimodal RFs of bimodal neurons can shift with the position of, for example, the hand, the arm, or the rotation of the head (e.g., Fogassi et al., 1996; Graziano & Gross, 1994).

To date, many studies have demonstrated that peripersonal space can be extended or projected simply by modulating the area that an individual can reach by means

---

[9] Whereas in these latter studies dynamic stimuli approached or receded from a static observer, the improvement of visual movement detection by congruent auditory motion signals during simulated self-motion would appear to suggest that multisensory (audiovisual) interactions can also be observed during self-motion (Calabro, Soto-Faraco, & Vaina, 2011).

of, say, the use of a tool (e.g., Holmes, Calvert, & Spence, 2004; see Spence, 2011, for a review). The results of a large number of such tool-use studies now indicate that visuotactile spatial interactions can also be observed at a distance that would normally be considered to fall outside the bounds of peripersonal space following tool-use (e.g., Bassolino, Serino, Ubaldi, & Làdavas, 2010; Canzoneri et al., 2013; Farnè, Bonifazi, & Làdavas, 2005; Holmes, Calvert, & Spence, 2007; Longo & Lourenco, 2006). The flexible properties of peripersonal space have also been demonstrated in neuropsychological research. So, for example, those patients who exhibit visuospatial neglect for stimuli presented in peripersonal space may well display a similar deficit in extrapersonal space following the use of a stick to perform a particular task, but not, interestingly, when using a laser pointer (e.g., Berti & Frassinetti, 2000; see also Pegna et al., 2001; Maravita, Husain, Clarke, & Driver, 2001). The observations of changes in distance-based multisensory interactions after tool-use fit well with the observation of RF changes of bimodal neurons in the macaque (Iriki, Tanaka, & Iwamura, 1996). It was shown that the visual RFs of bimodal visuotactile neurons in the macaque cover a larger region of space surrounding the body after tool-use as compared to before tool-use (but see Holmes & Spence, 2004, for a critical note).

Furthermore, it has now been demonstrated that social interactions can also modulate the boundary of peripersonal space (e.g., Heed, Habets, Sebanz, & Knoblich, 2010; Teneggi, Canzoneri, Di Pellegrino, & Serino, 2013). So, for example, in one study, Teneggi et al. reported that the speed with which their participants were able to detect a tactile stimulus presented to their right cheek was dependent on the perceived location of a simulated approaching auditory stimulus. When the auditory stimulus was perceived as being closer to the cheek, the participants' vocal responses to tactile stimulation of the right cheek were faster as compared to when the auditory stimulus was perceived as being further away (using an experimental paradigm similar to that used by Canzoneri et al., 2012). The critical region within which approaching auditory stimuli facilitated the participants' detection of tactile stimuli (when compared to auditory stimuli perceived at a more distant location) was decreased in the presence of an actor who was standing close to the far loudspeaker. With the actor present, the boundary of peripersonal space appeared to move closer to the participant as compared to when a mannequin (i.e., a human-like doll) was placed close to the far loudspeaker instead.

In another experiment, Teneggi et al. went on to show that depending on whether or not the actor was cooperating with the participant in an economic game, the border of peripersonal space either extended to incorporate the space surrounding the actor (in the cooperative condition) or not (in the non-cooperative condition). Figure 3 shows the shift in the time point of a simulated approaching auditory stimulus at which audiotactile interactions start to decrease RTs to tactile targets.

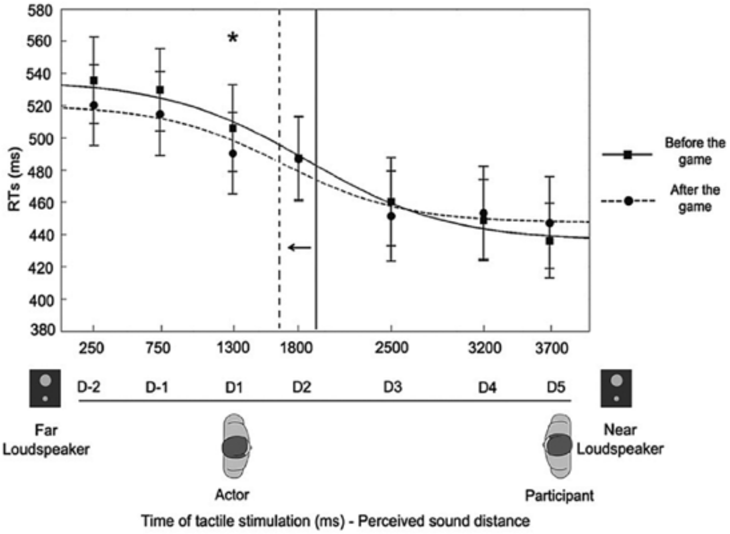


Figure 3. Vocal RTs as a function of the perceived distance of an approaching sound before (filled squares, solid line) and after (filled circles, dashed line) playing an economic game with a cooperative actor. The time until tactile stimulation is shown on the x-axis. Low values correspond with a sound that was far from the body whereas high values correspond with a sound that was close to the body. The boundary of peripersonal space (solid vertical line) shifted towards the actor after the game (dashed vertical line; taken from Teneggi et al., 2013, with permission).

In most of the studies in which the boundary of peripersonal space has been investigated, the participants remained static and the stimuli were either static or dynamic (e.g., perceived to approach or recede from the participant). There are, however, reasons to believe that the speed at which humans move through the environment may also modulate the effective boundary of peripersonal space. In a neurophysiological study by Fogassi et al. (1996), the extent of the RF of a certain type of neuron (i.e., “somatocentered” neurons) in terms of their distance from the observer (i.e., a monkey) was observed to depend on the velocity of an approaching visual stimulus (that is, increasing the velocity of the visual stimulus increased the extent of the RF in depth for certain neurons). It would be interesting here to investigate whether an extension of peripersonal space could be observed with moving observers who approach static stimuli at different speeds.

Another way in which the space that is currently reachable can be manipulated is by placing a (transparent) barrier (e.g., a window) in front of an individual, which effectively reduces the reachable space, but not the visible space. One might expect that knowing that a body part cannot be reached by a visual stimulus ought to decrease the extent of the peripersonal space border and the distance at which audiotactile and visuotactile interactions are especially efficient in evoking a rapid response. Knowing that an approaching visual stimulus will collide with a barrier in front of you instead of with your body may serve to reduce anxiety for the physical consequences of approaching visual stimuli, and therefore reduce attention to threat-related stimuli (see Eysenck, Derakshan, Santos, & Calvo, 2007, for the influence of anxiety on attention). Interestingly, many animals maintain a certain distance between themselves and other animals (or objects; i.e., “flight distance”, see Hediger, 1955 and Sommer, 1959; “personal space”, e.g., Felipe & Sommer, 1966). The distance that an animal maintains between itself and other animals (or objects) depends on the perceived threat of the other animals (i.e., preserving a smaller distance from other individuals from the same species, while maintaining a larger distance from predators). The distance of the boundary of peripersonal space from an observer varies between individuals. In one recent study, the size of the hand-blink reflex (HBR) was investigated with electrical stimulation of the wrist while participants were instructed to hold their hand at different distances from their face. An increase in the hand-blink reflex was observed as the hand got closer to the participant’s face (Sambo & Iannetti, 2013). An increase in the HBR with decreasing distance was taken as a measure of peripersonal space. Interestingly, the differences in the size of the peripersonal space between participants in this case were related to variability in trait anxiety (Eysenck, Derakshan, Santos, & Calvo, 2007). Although Sambo and Iannetti (2013) observed a relation between the size of defensive peripersonal space and trait anxiety, they did not find a relation between claustrophobic fear and defensive peripersonal space. Interestingly, a relation between peripersonal space and claustrophobic fear was observed when the magnitude of pseudo-neglect was used as a measure of peripersonal space (Lourenco, Longo, & Pathman, 2011). When indicating the center of a horizontal line, neurologically healthy individuals normally show a slight leftward bias called pseudo-neglect. This bias changes into a rightward bias as the distance between the line and the observer increases (Longo & Lourenco, 2006, 2007). The slope of the change from a leftward into a rightward bias as the distance between the line and the observer increases can be taken to represent the extent of peripersonal space. Shallower slopes indicate a larger peripersonal space and steep slopes represent a smaller peripersonal space. Individual differences in peripersonal space, as indicated by the slope of the bisection bias, were related to claustrophobic fear (Lourenco, Longo, & Pathman, 2011).

Some evidence for the idea that peripersonal space may contract by modulating the reachable space comes from a neurophysiological study by Caggiano et al. (2009). These researchers found that some mirror neurons (that is, neurons that are responsive to both the performance and viewing of goal-directed actions) responded when the monkey viewed goal-directed actions that were performed outside reachable space, whereas others responded selectively to goal-directed actions performed within reach (see also section 3.3). Crucially, when a transparent barrier was placed in front of the monkey, those neurons that had previously responded to goal-directed actions that were performed within the animal's peripersonal space did not respond anymore to the viewing of goal-directed actions that were performed at the same distance from the monkey but behind a transparent barrier. Thus, when the border changed the previously reachable space of the monkey to non-reachable space, the "periperisonal" mirror neurons did not respond anymore (see also Costantini, Ambrosini, Tieri, Sinigaglia, & Committeri, 2010, for similar behavioral effects in humans).

Furthermore, those neurons that previously responded to the viewing of actions in extrapersonal space now responded to those actions that were performed in non-reachable peripersonal space (behind the transparent barrier). These results suggest that it was the currently reachable space that was important in evoking a response in these "extrapersonal" and "peripersonal" mirror neurons.

In humans, however, studies of the influence of a barrier between stimuli from different modalities on multisensory interactions have shown different results. For example, Farnè et al. (2003) reported that the strength of visuotactile extinction in brain-damaged patients was not modulated by the presence of a transparent barrier between the visual and tactile stimuli. Similarly, the size of the visuotactile crossmodal congruency effect was not found to be influenced by the presence of a transparent barrier between the visual and tactile stimuli in Kitagawa and Spence's (2005) study using the crossmodal congruency task. It remains to be seen whether different experimental paradigms may be more sensitive in terms of capturing the influence of a transparent barrier on multisensory interactions. One thing to note about these studies is that the transparent barrier did not separate or create different regions of space in terms of depth. It would therefore be interesting to see if, and how, visuotactile interactions are modulated by the presence vs. absence of a transparent border (e.g., a window or windscreen) between visual and tactile stimuli placed at slightly different distances from the observer (much like when we look out a window and see, for example, a clueless pigeon flying towards us).

This approach has been used in a study of affordances in (virtual) peripersonal and extrapersonal space (Costantini, Ambrosini, Tieri, Sinigaglia, & Committeri, 2010). In this study participants had to perform reaching movements to a virtual mug that afforded either a left or a right precision grip. Before the presentation of the mug, a hand was shown indicating a left or a right hand precision grip (the instruction stimulus). Participants' movement onset times after the presentation of the mug were shorter when the instruction stimulus was congruent with the orientation of the mug (e.g., a left precision grip and a mug with the handle on the left), compared to when it was incongruent. This congruency effect was dependent on the perceived distance of the mug. The congruency effect was present when the mug was presented in peripersonal space, but not when it was presented in extrapersonal space. Placing a transparent border in front of the mug in peripersonal space was shown to remove this congruency effect. In a later study, the congruency effect was also shown to be present in extrapersonal (or non-reachable) space when the mug was within the reachable space of an avatar (Costantini, Ambrosini, Tieri, Sinigaglia, & Committeri, 2010; see also Cardellicchio, Sinigaglia, & Costatini, 2011). Although these studies of affordances do not directly relate to multisensory interactions in peripersonal and extrapersonal space, they do provide information on the interaction between the body and objects in the environment, the role of the distance between objects and the observer, and predictions about possible interactions with the environment.

3.1. Spatio-temporal properties of visuotactile and audiotactile interactions in peripersonal space

The specific spatio-temporal properties of dynamic stimuli can also influence the strength of the multisensory interactions that are observed. For example, when using dynamic (i.e., moving) visual stimuli that are seen to approach or recede from the participant's own body, information concerning time-to-contact can be used to predict when the detection of sensory information will be fastest (e.g., Canzoneri et al., 2012; Gray & Tan, 2002; Teneggi, Canzonero, Di Pellegrino, & Serino, 2013). When visual stimuli approach the body at a certain speed, predictions concerning the location of a tactile stimulus in 3-D space can be made (e.g., when and where contact with the body is expected), thus enhancing the speed with which stimuli at the expected time and location are processed relative to other times and locations. Similar results have now also been observed in a study of visuotactile interactions (Kandula, Hofman, & Dijkerman, 2015). In the latter experiment, movies of an arm reaching out to the left or right cheek of participants were presented prior to the onset of a tactile target that was presented to their left or the right cheek. Tactile targets were presented at different interstimulus intervals, either earlier, at the same time as, or later than the moment at which contact with the face was expected based on the speed of the reaching movement that was presented in the movie. The participants had to detect tactile targets that were presented to either the left or right cheek. As expected, responses to the tactile targets were fastest in the on-time condition.

In a follow-up experiment, this effect was found to depend on the time-to-contact information that could be derived from the movie of the moving arm (i.e., playing the reaching movie in reverse did not elicit an effect of SOA or spatial congruency). Such findings underline the importance of considering the relation between the temporal and spatial properties of multisensory stimulus presentation when trying to understand the dynamic multisensory interactions that, by now, have been documented in 3-D space.

3.2. Multisensory interactions involving mirror-reflected stimuli in peri- and extrapersonal space

The use of mirror-reflected stimuli in the investigation of multisensory interactions in peripersonal and extrapersonal space has also provided some interesting insights in terms of multisensory perception in the depth space. For example, Maravita et al. (2002) studied visuotactile congruency effects with stimuli that appeared to be presented in extrapersonal space. In one condition the stimuli were presented in peripersonal space but mirror-reflected to make them visually appear in extrapersonal space. In another condition the stimuli were presented physically in extrapersonal space at the same location that could be seen in the mirror reflection. A schematic side view of their setup is shown in Figure 4. Visuotactile congruency effects were shown to be significantly larger when the visuotactile stimuli were observed through a mirror and were known to be coming from peripersonal space as compared to when they were presented in extrapersonal space without the aid of a mirror. The results of this study have now been replicated in a more recent EEG study (Sambo & Forster, 2011). In the latter study, Sambo and Forster investigated visuotactile interactions in peripersonal space, with stimuli that were viewed through a mirror and thus appeared to be presented in extrapersonal space. This condition was then compared with the results from one of their earlier studies (Sambo & Forster, 2009) in which the stimuli were presented in extrapersonal space without the aid of a mirror. The researchers observed that the lateral spatial alignment between visual and tactile stimuli modulated the ERPs recorded over the somatosensory cortex in the mirror condition, but not in the extrapersonal condition.

Such results would therefore appear to suggest that the known position of mirror-reflected stimuli is more important than their visually perceived position in depth in terms of modulating multisensory spatial interactions. Once again, this result can be interpreted in terms of the spatial alignment in depth of visuotactile stimuli, but here in terms of the known rather than the perceived spatial locations in the depth space.

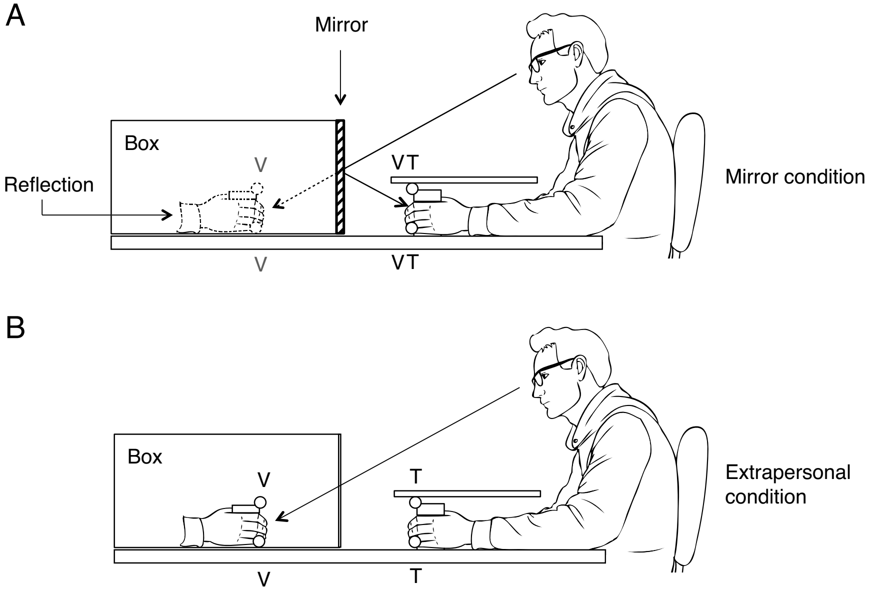


Figure 4. Schematic side view of the setup used in the study by Maravita, Spence, Sergent, & Driver (2002). Panel A shows the mirror condition, and panel B show the real extrapersonal condition. LEDs are indicated with a "V", and vibrotactile stimulators with a "T".

3.3. On the nature of the boundary of peripersonal space

Although the previously-mentioned observations, of stronger visuotactile and audiotactile interactions in peripersonal space as compared to when the component stimuli are presented from different locations in space, are usually taken to suggest that peripersonal space has a boundary defined by such multisensory interactions, it is important to note that these findings might simply reflect the importance of the particular spatial (mis)alignment of stimuli in depth in modulating these interactions (i.e., perhaps limited by the 3-D spatial RF size of neurons). Given that tactile perception is inherently bound to the body, the stronger multisensory interactions involving touch in peripersonal space require stimuli from different modalities to not only be aligned at any particular depth (presenting visuotactile stimuli in extrapersonal space is not meaningful), but specifically to be aligned with the body (or extensions of the body in the case of tool-use). It may be interesting to investigate the relation between the distance separating the unisensory component stimuli in depth and the strength of multisensory interactions involving touch. One can depict this relation as a steep curve with a sudden change in the strength of the interaction when a certain distance has been passed, or as a more gradual one (and if so, over what distance does this change occur). Alternatively, different spatial representation mechanisms may redundantly code stimuli presented within a narrow overlapping field at a border between peripersonal and extrapersonal space (similar to the suggested bilateral representation of a small part of the central visual field, e.g., Bunt, Minckler, & Johanson, 1977; Luo, Shan, Zhy, Weng, & He, 2011; see Figure 5 for an overview of the proposed shapes of the peripersonal border).

Canzoneri et al.'s (2012) results may be taken to suggest that the presentation of sensory information from slightly beyond the border of peripersonal space does not result in a sudden absence of all audiotactile interactions. Rather, there is a more gradual decline in the strength of any interactions that are observed, as demonstrated by a gradual increase in RTs. This effect would appear to be most pronounced for those sounds that are perceived to be approaching the participant, and less for those sounds that appear to recede from the participant. The observation of a gradual decline in RT with increasing distance between the unimodal components is in line with the idea that the border of peripersonal space is not a sharp one, but perhaps a (somewhat steep) gradient, suggesting that different neural mechanisms that code peripersonal and extrapersonal space may partly code the same region of space. Similar results have also been observed using the rubber-hand illusion (Lloyd, 2007). The strength of the illusion decreases as the distance between the rubber-hand and the participants' own hand increased with a significant decrease in the strength of the illusion for distances between the rubber hand and the participants' hand larger than ~30 cm (see also Sambo & Iannetti, 2013).

Neurophysiological evidence in support of the idea of there being an overlap in the coding of peri- and extrapersonal space comes from a study by Caggiano et al. (2009). These researchers observed groups of mirror neurons (that is, neurons that are responsive to both the performance and viewing of goal-directed actions) that were responsive to viewing actions performed at different distances from the animal. They reported that while certain neurons that they recorded from were responsive to those actions that were seen in peripersonal space, others were responsive to those actions seen in extrapersonal space. Some of the neurons were responsive to the viewing of actions in both regions of space.

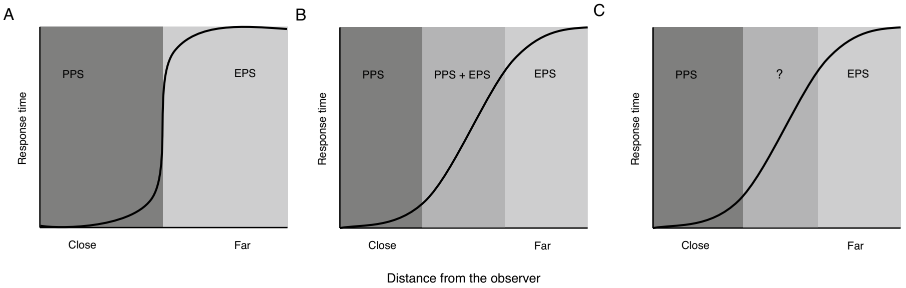


Figure 5. The putative existence of a border between peripersonal space (PPS) and extrapersonal space (EPS) might be taken to predict that visuotactile or audiotactile interactions would be stronger in peripersonal space. It can be imagined that this border might be relatively sharp (A) or else more gradual (B). Thinking more hypothetically, one might even consider whether there is a region of space that is included in the representation of both PPS and EPS (PPS + EPS in panel B), or a region that is not specifically included in either a PPS or EPP representation of space ('?', panel C).

Some hints concerning the nature of the border between peripersonal and extrapersonal space may also be found in a study by Gabbard et al. (2007) in which the participants had to estimate the border of the space that was reachable to them. Visual targets were presented at several distances in both peripersonal and extrapersonal space while the participants responded as to whether they thought that they would have been able to reach the visual stimulus or not. The results revealed that estimation errors peaked at the border of reachable space and declined at larger distances. Speculatively, these results might be seen as fitting in with the idea of overlapping spatial representations of peripersonal (reachable) and extrapersonal (non-reachable) space (see Figure 5B). In this situation, objects that are presented within the peripersonal space only region (dark gray region Figure 5B) are clearly reachable and objects that are presented only in extrapersonal space are clearly not reachable (light gray region Figure 5B). A small region of space surrounding the end of peripersonal space and the beginning of extrapersonal space (medium gray region Figure 5B) may, however, lead to a larger variation in reach estimates as these objects are coded both as within reach (within peripersonal space) and out of reach (extrapersonal space) and therefore to more errors, see also Bourgeois et al. (2014).

When thinking about the neuronal mechanisms that may give rise to the border in peripersonal space, studies such as Rizzolatti et al. (1981) immediately spring to mind. In this now-classic study, the researchers demonstrated that some bimodal neurons (pericutaneous and distant peripersonal neurons) were only responsive to those visual stimuli that were presented from within the animal's peripersonal space. Another type of neuron was found to be sensitive to those visual stimuli that were presented both far from and near to the animal (i.e., the distance at which stimuli were presented did not modulate neuronal responses in a systematic manner). Bearing such observations in mind, the visuotactile interactions that are documented in peripersonal space may also depend largely on overlapping RFs, in terms of their sensitivity to the depth from which stimuli are presented. Perhaps visuotactile integration depends on the coding of peripersonal space by bimodal neurons, whereas visuotactile interactions involving visual stimuli that are presented from further away from the observer would undoubtedly still occur, but would simply fail to meet the criterion for integration (that criterion being that multisensory responses should be different from the best of the unimodal responses; e.g., see Stein, Stanford, Wallace, Vaughan, & Jiang, 2004, p. 32).

As mentioned before, several studies have shown that distinct brain regions are involved in the processing of sensory information from different regions of space, but we usually do not perceive depth as a dichotomous entity. Drawing a distinction between the space that is reachable and that which is not would appear to be useful in terms of the possible (motor) interactions with the environment and the perception of stimuli that are somehow relevant in terms of their proximity to the body of the observer (in either a positive or negative way). In terms of perceiving and localizing stimuli, and performing fine-grained motor actions, however, a continuous representation of space would seem to make more sense. Interestingly, the existence of these two different spatial representations (dichotomous vs. continuous) has also been supported by the results of several behavioral studies. For example, the results of the sound localization in depth experiment in the study by Canzoneri et al. (2012, Figure 1B; Teneggi, Canzoneri, Di Pellegrino, & Serino, 2013, Figure S1) indicate that participants perceive the location of simulated approaching or receding unimodal auditory stimuli in terms of a spatial continuum in depth. During audiotactile stimulation, however, RTs showed a somewhat steeper decrease at a certain distance from the observer, indicative of there being some kind of boundary in peripersonal space. Such a border was not observed when audiovisual stimuli were used (see Teneggi, Canzoneri, Di Pellegrino, & Serino, 2013, Supplementary information), underlining the idea that those multisensory interactions that involve tactile stimulation may display a (dichotomous) border for peripersonal space (again, perhaps, due to the asymmetric nature of tactile perception).

As should by now be evident, studying multisensory interactions in peripersonal space in isolation is probably not all that informative when it comes to thinking about multisensory interactions in the depth space more generally. For that, one needs to compare the studies of multisensory interactions that have been documented in response to stimuli presented in other modalities (audiovisual), in other regions of space (e.g., extrapersonal space), or at least presented from varying depths.

4. Interim summary

Summarizing what we have seen thus far, a growing number of studies have provided evidence demonstrating the abundant multisensory interactions that take place between those auditory, visual, and tactile stimuli that happen to be presented within reachable space. The results of those studies in which the spatial alignment of the stimuli has been varied in the depth plane would appear to suggest that visuotactile and audiotactile multisensory interactions become more pronounced (as indicated by faster RTs) when the stimuli are aligned in terms of their distance from the (extended) body. In most of those studies in which visuotactile and audiotactile interactions were more pronounced in peripersonal as compared to extrapersonal space, the spatial alignment of the stimuli in the depth plane has also been varied. Such findings may therefore also be interpreted in terms of a reduction in multisensory integration as the distance between the component stimuli is increased. Something like this was also hinted at by Sambo and Forster (2009, p. 1556) when they stated that: "…these results show that the spatial relationship between visual and tactile stimuli modulate early ERPs, with enhanced amplitudes for tactile stimuli coupled with visual stimuli delivered near the site of tactile stimulation (i.e., perihand space) compared to ERPs obtained when visual stimuli are presented at a different location in peripersonal or extrapersonal space, as one would predict according to the spatial rule of multisensory integration". To further explore the role of spatial alignment in multisensory interactions in the depth space, we will now discuss audiovisual interactions in frontal periand extrapersonal space.

5. Studies that have focused on audiovisual interactions in frontal peri- and extrapersonal space

As shown in the previous sections of this review, the results of neurophysiological, neuropsychological, and psychophysical studies converge on the suggestion that the space directly surrounding the body (i.e., peripersonal space) and the space that lies just beyond reach is represented differently by the brain when looking at visuotactile and audiotactile interactions (e.g., Caggiano, Fogassi, Rizzolatti, Thier, & Casile, 2009; see Makin, Holmes, & Ehrsson, 2008, for a review). Given such a distinction, it is remarkable that so little attention has been paid to those multisensory interactions that also take place in extrapersonal space. Indeed, the focus on multisensory interactions in depth thus far has been almost exclusively on uncovering the nature of visuotactile and audiotactile interactions in peripersonal space. If the spatial rule were to apply in the depth plane, as it has repeatedly been shown to apply in the horizontal and vertical planes (at least when the participant's task is in some sense spatial; see Spence, 2013), it can be used to explain the asymmetry of multisensory interactions in depth that involved tactile stimulation. Given that tactile perception is inherently about those stimuli/events that happen to be situated on the observer's skin, spatial alignment in depth (more specifically, spatial alignment with the body) is required for certain multisensory interactions to occur. For purely audiovisual interactions no such an asymmetry would be expected, as the perception of auditory and visual stimuli is not constrained by the distance between the stimuli and the body.

5.1. Spatial properties of audiovisual interactions in peripersonal and extrapersonal space

Given that visual and auditory stimuli can be perceived from both close to and far from the observer, we believe that it will be interesting in future research to investigate whether audiovisual interactions are more pronounced in a certain region of space or more or less equal for different regions of space. For example, whereas auditory and visual information may be dominant in extrapersonal space, visual, somatosensory, and vestibular information may be dominant in peripersonal space (Previc, 1998). In their recent study of crossmodal exogenous attention in 3-D space, Van der Stoep et al. (2014) demonstrated that the presentation of spatially non-predictive auditory cues gave rise to an enhancement in the detection of visual targets at cued as compared to uncued distances (relative to the observer; see Ngo & Spence, 2010, for similar findings during visual search).

These results can be interpreted in terms of a shift of the participant's exogenous crossmodal spatial attention in the depth plane. The participants in this particular study had to make up/down judgments concerning the elevation of visual targets that were presented on the left or the right of a fixation cross in either peripersonal or extrapersonal space (see Figure 6 for a bird's-eye view of the experimental setup). An exogenous auditory cue was presented from one of four locations prior to the onset of a visual target presented near left, near right, far left, or far right. The cues could therefore either be valid or invalid in terms of their laterality, and in terms of their depth.

Van der Stoep et al.'s (2014) results indicated that lateralized spatial cuing effects were only documented when the cue and target were presented from the same depth (see Figure 7 for a summary of the results). Interestingly, the magnitude of this effect did not appear to differ between those targets that were presented in peripersonal as compared to extrapersonal space. Figure 7 shows the mean RTs in each cuing condition for those targets presented in peripersonal as compared to extrapersonal space.

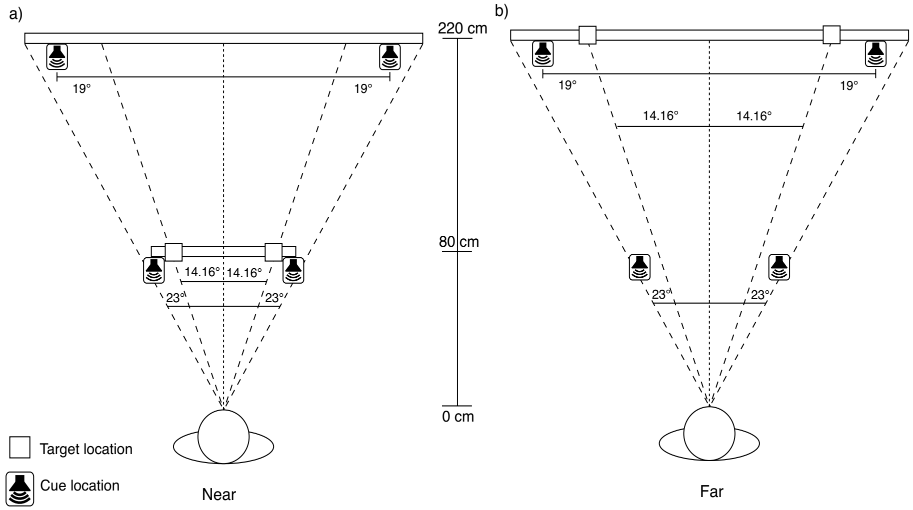


Figure 6. Panel A shows a bird's-eye view of the experimental setup in the condition in which visual targets were presented in peripersonal space. Panel B shows the setup for targets presented in extrapersonal space [taken from Van der Stoep, Nijboer, & Van der Stigchel, 2014, with permission].

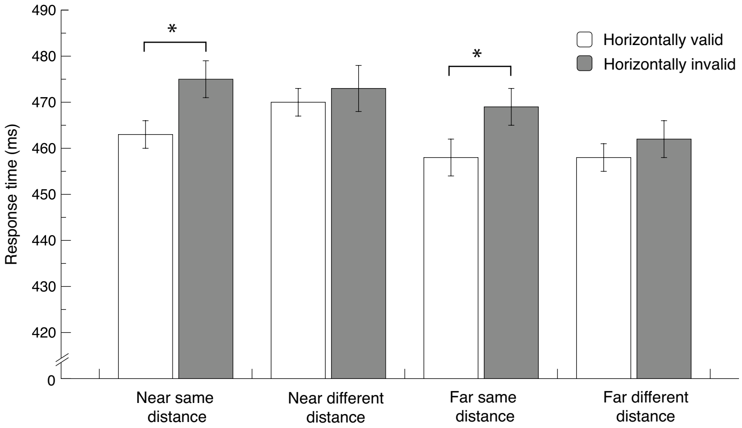


Figure 7. RTs to visual targets following the presentation of either a valid or invalid lateral cue at the same or a different distance in near and far space. Significant lateral cuing effects were only ever observed when the cue and the target were presented at the same distance (the asterisk indicates a significant difference, $p < .005$). Error bars represent standard errors of the mean with between-subject variance removed [data taken from Van der Stoep et al., 2014, with permission].

In addition to showing that crossmodal exogenous attentional cuing can operate n the depth plane, these results also show that, at least for crossmodal audiovisual exogenous spatial attention, there are no differences in terms of whether both of the stimuli are presented in peripersonal as compared to extrapersonal space. These results therefore do not support a distinction between peripersonal and extrapersonal space as defined by the strength of audiovisual interactions. Instead, they can be taken to suggest that auditory and visual RFs are defined in terms of their responsiveness to stimuli in 3-D space (azimuth, elevation, and depth).

Another recent study provided support for the idea that there might be more pronounced audiovisual interactions in extrapersonal as compared to peripersonal space (Van der Stoep, Van der Stigchel, Nijboer, & Van der Smagt, 2015; also see Previc, 1998). The participants in this particular study had to detect auditory, visual, and audiovisual stimuli that were presented in either peripersonal or extrapersonal space. There were four conditions in the experiment: One condition in which the stimuli were presented in peripersonal space (with a certain intensity), a condition in which the same stimuli were presented from a greater distance and were therefore of a lower intensity (stimuli were not corrected for visual angle and intensity), a condition in which the same stimuli were presented from a greater distance but were corrected for visual angle and intensity, and lastly a condition in which the stimuli were presented in peripersonal space but with the same decrease in intensity as the stimuli presented at a larger distance without correcting for intensity and visual angle (see Figure 8A for an overview of the conditions).

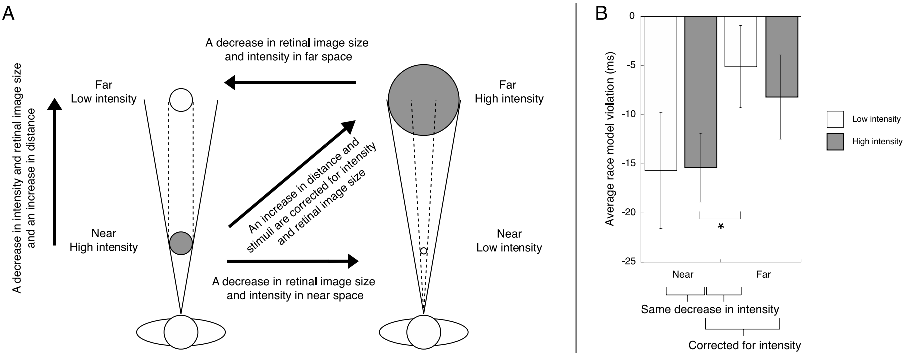


Figure 8. Left panel: A schematic overview of the conditions that were used in Van der Stoep et al.'s (2015) study. Right panel: The average amount of race model violation in each condition. The critical difference between the Near High and the Far Low condition was significant (*p* < .05). Error bars represent standard errors of the mean and contain between-subject variance [data taken from Van der Stoep, Van der Stigchel, Nijboer, & Van der Smagt (2015), with permission].

The results revealed that a decrease in intensity that was related to an increase in distance from the observer (presentation in extrapersonal space) enhanced audiovisual integration, whereas the same decrease in intensity for stimuli that were presented in peripersonal space did not result in an increase in multisensory integration (see Figure 8B). These results can therefore be taken to suggest that audiovisual interactions may be stronger in extrapersonal as compared to peripersonal space. The amount of audiovisual integration was, however, similar in both regions of space in a condition in which the stimuli were corrected for visual angle and sound pressure level. One could argue that these results reflect a preference of the brain to integrate weak auditory and visual stimuli that are presented at large distances relative to weaker stimuli presented from closer to the observer. Speculatively, this preference might be the result of the multisensory experience that stimulus intensity usually decreases as a function of increasing distance (i.e., increasing distance is lawfully related to decreasing retinal image size and perceived auditory and visual intensity).

Although the previously-mentioned studies on audiovisual interactions in peripersonal and extrapersonal space indicated that audiovisual interactions between those stimuli presented at different distances in the space in front of the observer do not necessarily demonstrate any asymmetric effects in depth, one might have expected that differences in multisensory interactions would occur as the distance at which stimuli are presented increased, given that differences in the arrival time of visual and auditory stimuli will also increase (see Spence & Squire, 2003).

5.2. Temporal properties of audiovisual interactions in peripersonal and extrapersonal space

Given that the temporal relation between auditory and visual information depends on the distance from which the stimuli are presented, any resulting differences in arrival times may influence the strength of audiovisual interactions between stimuli presented from different distances. Sugita and Suzuki (2003) investigated whether the point of subjective equality for visual and auditory stimuli (that were presented by means of headphones) depended on the perceived depth of the visual stimulus (see also Engel & Dougherty, 1971). In this case, the participants were explicitly instructed to imagine that the visual stimulus was the source of both the light and the sound. Increasing the distance of the visual stimulus increased the auditory delay that was necessary for subjective simultaneity for stimuli that were presented from distances up to ~20 m.

In a later study of the relation between subjective temporal alignment and the depth at which audiovisual stimuli were presented, Alais and Carlile (2005) investigated the influence of the direct-to-reverberant energy ratio on audiovisual distance perception. This is the ratio between the energy of a sound source that is perceived directly, and the energy that is perceived by means of the reflections in enclosed reverberant environments (Bronkhorst & Houtgast, 1999). The results showed that this ratio was crucial in determining the perceived distance of a sound source and the time delay that was required for participants to judge whether a visual stimulus was aligned temporally (i.e., synchronous) with auditory stimuli that were presented from different simulated distances (also see Gardner, 1968). As in the study by Sugita and Suzuki, the delay that was needed for temporal alignment between the auditory and visual stimuli scaled with approximately the speed of sound (3.43 ms/m) for distances up to 20 m. These results suggests that human observers are able to judge whether the auditory and visual components of an audiovisual stimulus are co-located in terms of their distance from the observer by using multisensory experience-based estimations of differences in arrival times (see also Silva et al., 2013). However, this effect appears to depend primarily upon strategies in which differences in arrival time are explicitly taken into account (e.g., Alais & Carlile, 2005; Arnold, Johnston, & Nishida, 2005; Kopinska & Harris, 2004; Harris, Harrar, Jaekl, & Kopinska, 2010; Lewald & Guski, 2004; Stone et al., 2001). No correction for differences in arrival times was observed when participants judged perceptual simultaneity without taking the distance of the stimuli into account (see Ichikawa, 2009, for unisensory visual temporal order judgments with stimuli spatially misaligned in depth).

Based on these results, it could be argued that humans may especially benefit from audiovisual integration when stimuli are presented at distances beyond 20 m as the correct determination of auditory distance becomes increasingly difficult to ascertain. On the other hand, one could argue that visual perception will be dominant at certain distances (Agganis, Muday, & Schirillo, 2010; Kitagawa & Ichihara, 2002), thus reducing any audiovisual integration that is observed merely to the level of statistical facilitation (Miller, 1982, 1986; Raab, 1962).

Taken together, the results of those studies of audiovisual interactions in peripersonal and extrapersonal space suggest that these interactions are not stronger at a certain distance from the observer. This notion is in line with the idea that asymmetric effects of multisensory interactions in depth depend both on spatial alignment of the unimodal component stimuli in depth and the nature of the sensory modalities that are involved. Auditory and visual perception is not constrained by the distance at which stimuli are presented, as is tactile perception. Consequently, audiovisual interactions do not show the same asymmetric effects as audiotactile and visuotactile interactions.

In that sense, multisensory interactions involving touch are distinct from purely audiovisual interactions, as the tactile component requires contact with the body. Thus, multisensory interactions involving touch may be especially pronounced close to the body, given that spatial alignment in depth will always require that all unimodal component stimuli be presented in peripersonal space in alignment with the body.

Although these studies do provide information concerning the multisensory interactions taking place in frontal space, it is important to remember that multisensory interactions can also occur between stimuli that are presented in the space behind the observer. Therefore, in order to gain a fuller understanding of multisensory interactions in depth, those multisensory interactions that take place in rear space also need to be addressed.

6. Rear peripersonal and extrapersonal space

The area of space behind the observer, rear space, cannot be seen directly, and therefore provides an interesting opportunity to investigate those interactions taking place between auditory and somatosensory information. Auditory perception covers the full 360° of space around an observer, whereas somatosensory perception covers the entire body surface (and thus covering 360° of (near) space).

6.1. Audiotactile interactions in rear space

The distinction between peripersonal and extrapersonal space can also be made when it comes to the space behind an observer. Indeed, the results of neurophysiological (Graziano, Reiss, & Gross, 1999) and neuropsychological studies (Farnè & Làdavas, 2002) have clearly shown that audiotactile spatial interactions are more pronounced in near rear space. Graziano et al. (1999) observed that multisensory neurons in the macaque were especially responsive to auditory stimuli that were presented close to the head. In humans, differences between rear peripersonal and rear extrapersonal space have been observed in those patients exhibiting audiotactile extinction (Farnè & Làdavas, 2002). Audiotactile extinction was stronger for auditory stimuli that were presented in rear peripersonal (or near) space compared to rear extrapersonal (or far) space. Similar results have since been obtained in healthy participants. For example, Kitagawa et al. (2005) reported that auditory distractors (white noise bursts) that were presented in rear peripersonal space (~20 cm) caused a greater distractor effect for tactile targets compared to auditory distractors presented in rear extrapersonal space

(~70 cm). Importantly, these differences could not be explained in terms of differences in sound pressure level or localizability.

Whereas several studies have shown the importance of spatial alignment for the occurrence of multisensory integration, the spatial alignment of audiotactile stimuli in depth in peripersonal space (i.e., front vs. rear peripersonal space) has been shown not to influence the amount of audiotactile integration (Zampini, Torresan, Spence, & Murray, 2007). Zampini et al. suggested that this is perhaps the result of the large auditory RFs of auditory-somatosensory neurons, making the spatial alignment between auditory and tactile information less relevant for integration to occur (but see Tajadura-Jiménez et al., 2009). However, influences of spatial alignment on multisensory integration have mainly been observed in studies where the spatial locations of stimuli were somehow relevant to the participant's task (Spence, 2013). Given that the task was to detect stimuli in either the auditory or somatosensory modality, it cannot be ruled out that spatial modulation of audio-somatosensory integration is possible.

A study of the Colavita[10] effect with audiotactile stimuli that were presented in front and rear space revealed an effect of lateral spatial alignment in rear space (i.e., stimuli presented from the same vs. opposite side; Occelli et al., 2010). Furthermore, this effect was only present when auditory stimuli consisted of white noise bursts as compared to pure tones. Compared to pure tones, complex sounds involve more frequencies and therefore they generally allow for better localization (e.g., Frens, Van Opstal, & Van der Willigen, 1995) and perhaps as a result also better multisensory binding based on the spatial location of stimuli (e.g., Occelli, Hartcher-O'Brien, Spence, & Zampini, 2010). Furthermore, given that complex sounds are so much more like ecologically valid sounds, humans may have much more multisensory experience with this type of sounds shaping the way in which the brain responds to multisensory stimulation. Interestingly, the difference between the strength of multisensory interactions with complex sounds and pure tones was especially pronounced in rear space (e.g., Occelli, Hartcher-O'Brien, Spence, & Zampini, 2010; also see Farnè & Làdavas, 2002). Auditory dominance in rear space could cause audiotactile interactions to be more sensitive to differences in the localizability of sounds.

6.2. Audiovisual and visuotactile interactions in front and rear space

When thinking of multisensory interactions in rear space, perhaps audiovisual interactions do not readily come to mind, given that we can often not see the space

---

[10] In a typical study of the Colavita effect, participants have to report the modality of unimodal (e.g., visual and auditory) and bimodal (e.g., audiovisual) stimuli and often report a certain modality more than the others (e.g., more reports of visual than of auditory or audiovisual stimuli; see Spence, Parise, & Chen, 2011, for a review).

behind us. Nevertheless, some researchers have investigated audiovisual interactions in both front and rear space. For instance, using a simulated driving task, Ho and Spence (2005) demonstrated that auditory exogenous and auditory endogenous spatial cues presented in front of, or behind, participants were able to decrease detection times of visual events that occurred at the cued location. The visual target behind the participants was visible through a mirror, providing ways to see visual targets that were presented in rear space. In a later study, using a similar setup, auditory spatial cues were shown to facilitate visual target discrimination at the same location (valid vs. invalid in front/rear space), whereas tactile spatial cues did not evoke any such effect (Ho, Tan, & Spence, 2006; but see Ho, Tan, & Spence, 2005). Among others, the lack of an effect of vibrotactile cues was suggested to be the result of less spatial alignment between cue and target. Vibrotactile cues were presented at the front or the back of the torso of participants, but visual targets were presented in the space on the rear left, 70 cm behind participants. In contrast, auditory cues were presented from the same location as the visual event in rear space (in left rear space 70 cm from the body of the participant).

Mirror reflected visual stimuli presented in rear space have also been used in some studies of visuospatial neglect (Ramachandran et al., 1999; Viaud-Delmon, Brugger, & Landis, 2007). Interestingly, it has been observed that some patients with left visuospatial neglect in frontal space do not neglect visual stimuli presented in rear space but that are seen through a mirror in frontal space. Although these findings do not provide information on multisensory interactions in rear space, they do provide evidence that attentional orienting can be specifically impaired in frontal space with intact attentional orienting in rear space. This further underlines the idea that rear space is a spatial region that is distinct from other regions of space.

Taken together, then, these results indicate that audiotactile, visuotactile, and audiovisual interactions can also occur in rear space, and again indicate that a distinction between (rear) peripersonal and extrapersonal space can be observed, possibly mediated by the spatial alignment of stimuli in terms of their distance from the observer in rear space.

7. Differences between multisensory interactions involving touch and audiovisual interactions

So far, we have demonstrated that depth related asymmetries in multisensory interactions that occur between different modalities can be explained by the spatial alignment of the unisensory component stimuli in depth and spatial alignment with the body. This may not be so surprising when thinking of the observation of, for example, bimodal, visuotactile neurons with visual RFs that are sensitive to visual stimulation within a certain distance (i.e., depth) from the tactile RF (e.g., Fogassi et al., 1996; Graziano, Reiss, & Gross, 1999). Although audiovisual interactions and multisensory interactions involving the tactile modality may be explained by differences in the particular spatial alignment of unimodal stimuli in depth, these multisensory interactions also differ in several ways that are determined by the particularities of each sensory modality.

Auditory, visual, and tactile stimuli are initially all encoded in different references frames: auditory stimuli are encoded in a head-centered reference frame, visual stimuli in an eye-centered or retinotopic reference frame, and tactile stimuli in a limb/bodypart centered reference frame (see Cohen & Andersen, 2002, and Spence & Driver (2004), for reviews). The spatial alignment of stimuli from different sensory modalities can only be determined if the spatial location of these stimuli can be compared in a common reference frame. Interestingly, a target location that is initially encoded in one reference frame can be transformed into a different reference frame (e.g., the location of an auditory stimulus location that is encoded in a head-centered reference frame may be transformed into an eye-centered reference frame by means of intermediate-reference-frames; Avillac et al., 2005). With multisensory interactions between auditory and visual stimuli, such reference frame transformations need to take both head and eye-movements into account. When one sensory modality is dominant in a certain brain region (e.g., touch in the ventral intraparietal area, VIP) the reference frame of the other modality is often transformed to that of the dominant modality (Avillac et al., 2005). In line with this idea is the observation of visual RFs of bimodal visuotactile neurons that shift with the position of the hand (e.g., Fogassi et al., 1996; Graziano et al., 1997), anchoring vision to the body (see Avillac, Deneve, Olivier, Pouget, & Duhamel, 2005, p. 947, for examples with different modalities).

Multisensory interactions involving touch are thus different from audiovisual interactions because they always involve stimulation or prediction of stimulation of the body. The importance of knowing what and when something touches the body seems especially relevant for interaction with the environment and staying out of harm's way (see Section 3.1). However, a clear distinction between the importance of audiovisual interactions on the one hand and audiotactile or visuotactile interactions on the other in spatial processing seems a bit artificial, given that auditory, visual, and tactile information are generally continuously presented to us in our daily lives. In the future, it could be interesting to examine the influence of the spatial alignment of different unimodal component stimuli in depth on the interaction between audition, vision, and touch to investigate the degree to which the spatial alignment of each of these unimodal components contribute to enhanced multisensory (trimodal) interactions in peripersonal space.

8. Conclusions

The importance of studying multisensory interactions in 3-D space is finally starting to be recognized by more researchers. Still, it is evident that research on multisensory interactions that takes the distance at which stimuli are presented into account is still lacking. The majority of studies of multisensory interactions in the depth plane have focused on their distinct nature in peripersonal space. Multisensory interactions between pairs of stimuli that include the tactile modality are often found to be strongest when the stimuli are all presented from within peripersonal space as compared to when the component stimuli are presented from different regions of space. In sum, these asymmetrical spatial effects can be explained by the particular spatial alignment of the unimodal component stimuli, and the fact that tactile perception is limited to body. This means that for multisensory interactions involving touch spatial alignment in depth inevitably requires spatial alignment with the body.

We are generally only able to perceive tactile stimulation when the source of stimulation is presented on, or from, the body (see Spence, 2011, for a review; using a tool could be considered to lead to tactile perception in extrapersonal space, but this will always be mediated by, for example, mechanoreceptors in the skin). The nature of certain multisensory interactions in peripersonal space may therefore be attributable to the asymmetry in being able to perceive information presented from different distances. Indeed, studies on active tool use that increase reachable space have revealed similar visuotactile interactions for stimuli that are presented at distances normally considered as extrapersonal space.

Second, a closer look at the multisensory interactions that have been documented in studies concerning multisensory processing in frontal peripersonal, frontal extrapersonal, rear peripersonal, and rear extrapersonal space makes it evident that the majority of the asymmetries in multisensory interactions in the depth plane can be explained in terms of spatial alignment/misalignment in depth/lateral space. Furthermore, multisensory interactions between those sensory modalities that do not require stimulation in close proximity of the body should not be asymmetric in terms of the distance from which these stimuli are presented. This was indeed the finding of a recent study by Van der Stoep et al. (2014). Although the spatial alignment in depth did affect the presence of a crossmodal cuing effect, this effect did not differ for visual targets that were presented in near as compared to far space. Table 1 provides an overview of the multisensory interactions that are observed in the regions of space and the modulatory factors that were discussed in this review and shows clearly how audiotactile and visuotactile interactions are constrained by the fact that tactile perception is inherently close to the body. This seems to provide an explanation for differences between peripersonal and extrapersonal space both in front of and behind the observer.

It seems as though the importance of visuotactile and audiotactile interactions in peripersonal space has so far mainly been explained in rather non-specific terms concerning their relevance in dealing with interactions with the environment close to the observer. Although these explanations clearly make sense, they do not provide a clear mechanism that underlies the special role of such interactions in peripersonal space.

Given the important role of multisensory experience with the environment in the development of multisensory spatial interactions (Wallace & Stein, 1997, 2001, 2007), it may be relevant to consider the role of depth in the development of multisensory spatial interactions. Sources of tactile stimulation are normally located on the body, which provides an opportunity for the body to build a coherent multisensory spatial representation, perhaps calibrating the sensory systems of audition, touch, proprioception, and vision, through their interactions in depth space (see Graziano, Hu, & Gross, 1997, p. 2289). This may result in multisensory interactions that are sensitive not only to the horizontal and lateral spatial alignment of sensory information but also to their spatial alignment in depth. In contrast with this idea is the seemingly asymmetrical effect of spatial alignment on audiotactile interactions that has recently been documented when stimuli are presented in frontal as compared to rear peripersonal space (see Occelli, Hartcher-O'Brien, Spence, & Zampini, 2011, for a review). Whereas the relative spatial alignment in the horizontal (or azimuthal) plane of auditory and tactile stimuli does not seem to be very important when it comes to studying audiotactile interactions in frontal peripersonal space (but see Farnè & Làdavas, 2002, and Spence, Nicholls, Gillespie, & Driver, 1998), it seems crucial for such interactions when the very same stimuli are presented in rear peripersonal space. This may, in part, be the result of the dominance of vision in frontal space, making spatial alignment of audiotactile stimuli in frontal space less important for multisensory interactions as compared to spatial alignment of audiotactile stimuli in rear space.

The distance from the body at which multisensory interactions involving the tactile modality are enhanced may depend on the body part that is stimulated (something that has also been suggested by Tajadura-Jiménez et al., 2009). Although audiotactile interactions seem to be more pronounced with the tactile stimulation of the head compared to the hands (Tajadura-Jiménez et al., 2009), it is still unclear how the spatial alignment between audiotactile and visuotactile stimuli influences multisensory interactions with tactile stimulation of different body parts (perhaps RF sizes around certain body parts allow for a larger distance between the unimodal component stimuli while still resulting in multisensory interactions/integration, see, for example, Teramoto, Nozoe, & Sekiyama, 2013).

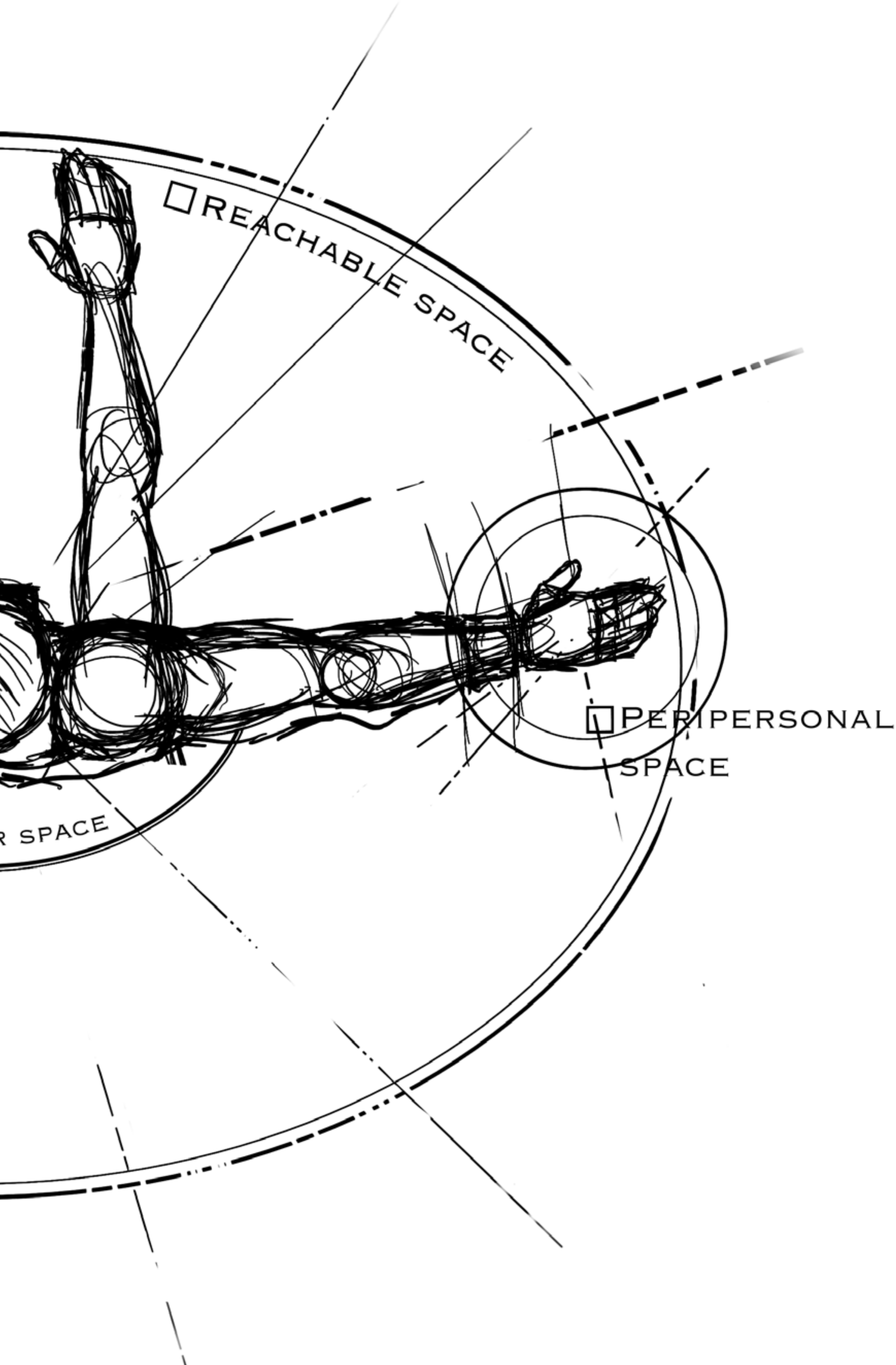
Besides the role of the spatial alignment of stimuli presented in depth in terms of modulating multisensory interactions, there may also be a role for spatial attention. For instance, one could argue that spatial attention is more focused in conditions of spatial alignment or better captured as compared to when stimuli are presented from different positions (see Spence, 2010). It remains to be seen what the relative contribution of attention is to the processing of multisensory information in such a complex spatial environment (e.g., Alsius, Navarra, & Soto-Faraco, 2007; Fairhall & Macaluso, 2009; Talsma, Doty, & Woldorff, 2007; Talsma and Woldorff, 2005; Van der Stoep, Van der Stigchel, & Nijboer, 2015; see Koelewijn, Bronkhorst, & Theeuwes, 2010, and Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010, for reviews).

To conclude, it would seem important to take several factors into account when investigating multisensory interactions with a focus on either peripersonal or extrapersonal space. The spatial alignment in depth of stimuli presented in different modalities (more specifically spatial alignment with the body in the case of multisensory interactions involving touch), the distribution of attention in depth, and the asymmetric nature of multisensory interactions involving the tactile modality may all contribute to the behavioral effects that arise from such multisensory interactions.

Table 1 Observed multisensory interactions within each region of space and modulatory factors.

	Frontal space		Rear space	
	Peripersonal space	Extrapersonal space	Near space	Far space
Observed				
Multisensory interactions	VT, AT, AV	AV, VT	AT	AV
Modulatory factors	Lateral space, depth, social context, anxiety	Lateral space, depth	Lateral space, depth	Lateral space, depth, stimulus intensity

V = Visual, T = Tactile, A = Auditory


☐ Reachable space
☐ Peripersonal space
R SPACE

Chapter 7

Exogenous orienting of crossmodal attention in 3-D space: Support for a depth-aware crossmodal attentional system

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 Review, 21, 708-714

Abstract

The aim of the present study was to investigate exogenous crossmodal orienting of attention in three- dimensional (3-D) space. Most studies in which the orienting of attention has been examined in 3-D space concerned either exogenous intramodal or endogenous crossmodal attention. Evidence for exogenous crossmodal orienting of attention in depth is lacking. Endogenous and exogenous attention are behaviorally different, suggesting that they are two different mechanisms. We used the orthogonal spatial-cueing paradigm and presented auditory exogenous cues at one of four possible locations in near or far space before the onset of a visual target. Cues could be presented at the same (valid) or at a different (invalid) depth from the target (radial validity), and on the same (valid) or on a different (invalid) side (horizontal validity), whereas we blocked the depth at which visual targets were presented. Next to an overall validity effect (valid RTs < invalid RTs) in horizontal space, we observed an interaction between the horizontal and radial validity of the cue: The horizontal validity effect was present only when the cue and the target were presented at the same depth. No horizontal validity effect was observed when the cue and the target were presented at different depths. These results suggest that exogenous crossmodal attention is "depth-aware," and they are discussed in the context of the supramodal hypothesis of attention.

Introduction

In a typical exogenous crossmodal spatial-cueing study, a spatial cue is presented on the same side as the target or on the opposite side. Exogenous spatial cues presented in one modality are able to speed up responses to target stimuli presented in a different modality when they are presented at approximately the same location as the target, as compared to when they are presented at opposite locations (i.e., a validity effect). These crossmodal exogenous cueing effects have now been reported for all combinations of visual, auditory, and tactile stimuli (Spence & MacDonald, 2004), suggesting that exogenous spatial attention operates in a supramodal fashion (e.g., Eimer & Van Velzen, 2002; Farah, Wong, Monheit, & Morrow, 1989).

In most studies in which exogenous crossmodal orienting of attention has been examined, cueing effects were assessed in a single plane of depth. Yet, in real-life environments, visual and auditory sources can appear at various distances from the observer. Studies with healthy individuals in which exogenous intramodal attention was examined at different depths have indicated that exogenous visual cues are able to attract visual attention to a specific three-dimensional (3-D) location (Atchley, Kramer, Andersen, & Theeuwes, 1997; Bauer, Plinge, Ehrenstein, Rinkenauer, & Grosjean, 2011; Theeuwes & Pratt, 2003). Although it is known that exogenous intramodal attention operates in depth, evidence is scarce for exogenous crossmodal orienting of attention in depth. In a study by Ho and Spence (2005), the facilitating effects of exogenous and endogenous crossmodal attention were investigated in a simulated driving setup with cues and targets presented in front of and behind the participant. Responses were faster when the cue and the target originated from the same location in space. Although this indicates that attention can be crossmodally attracted to a location in 3-D space either in front of or behind a driver, it does not show whether exogenous crossmodal attention can be shifted in different planes of depth in front of the participant. Studies that did look into attentional shifts to different planes in frontal depth either concerned exogenous or endogenous intramodal orienting of attention (see for example Downing & Pinker, 1985, and Couyoumdjian, Di Nocera, & Ferlazzo, 2003 for intramodal endogenous attention in frontal depth).

Although they recruit overlapping brain networks (Kim et al., 1999), exogenous and endogenous orienting of attention appear to have different properties (e.g., they have different time courses and are differently affected by cognitive load; see Berger, Henik, & Rafal, 2005). These differences indicate that exogenous and endogenous crossmodal forms of attention do not necessarily work in the same way, and may interact under, for example, high task demands.

So far, previous studies have thus provided support for the idea that intramodal exogenous and intra- and crossmodal endogenous attention are able to shift in depth. However, it is currently unclear whether and how crossmodal exogenous attention can be shifted in frontal space. In order to investigate this, we used the orthogonal spatial-cueing paradigm (Spence & Driver, 1997) and presented auditory cues and visual targets in either a near or a far depth plane. We blocked the plane of depth at which targets were presented, because otherwise near-space targets would occlude far-space targets, and we randomized the depth at which cues were presented. If exogenous crossmodal attention is "depth-aware," we expected to find a validity effect in the horizontal dimension (i.e., a classic validity effect) when cues were presented at the same depth as the target, but not when the cue and target were presented at different depths. In contrast, if exogenous crossmodal attention is not "depth-aware," the validity effects should not differ for cues presented at the same depth as or at different depths from the target.

Materials and method

Participants

On the basis of previous studies on attention in 3-D space (Atchley, Kramer, Andersen, & Theeuwes, 1997; Couyoumdjian, Di Nocera, & Ferlazzo, 2003; Theeuwes & Pratt, 2003), in which the samples varied between 10 and 24 participants, we included 16 healthy participants (13 female, three male; mean age = 22.44 years, *SD* = 1.90) who received course credits for their participation. All participants reported normal or corrected-to-normal visual acuity and no hearing problems, and showed normal performance on a short left– right sound localization task (see below). The experiment was performed in accordance with the Declaration of Helsinki, and participants signed informed consent before the start of the experiment.

Apparatus

To project the visual stimuli on a black canvas (near, 75 × 60 cm; far, 170 × 170 cm), we used a Toshiba TLP-T621 LCD projector (60 Hz). Four speakers (Harman/Kardon HK206, frequency response: 90–20000 Hz) were used to present the auditory cues. A chinrest was used to stabilize the participant's head and to keep the distance between the participant and the projection largely stable across participants.

Stimuli, task, and procedure

The loudness, and even more so the direct-to-reverberant energy ratio of a sound, provides information about the distance of a sound source and enables us to estimate its approximate distance in closed environments (Bronkhorst & Houtgast, 1999). In order to investigate the influence of exogenous auditory cues from different locations in 3-D space on visual information processing, the cues in our experiment should vary on both properties, depending on the distance between the auditory source and the observer. This would ensure that the brain received enough information to estimate the approximate location of the sound source, and possibly to attract attention to that location. Auditory cues consisted of a 75-ms, 2000-Hz tone (10-ms rise and fall of the signal) of 100 dB(A) SPL, as measured with an audiometer directly in front of the speakers. We used a sine wave as the auditory cue, to ensure that left and right could be distinguished, but not elevation (up and down; Frens & Van Opstal, 1995). This was important, because participants had to indicate whether a visual target was presented above or below the vertical center of the screen in the main experiment. The sine waveform ensured that the auditory cue could not be used as a landmark for visual target localization. Auditory cues that were presented in far space had a lower SPL than did auditory cues that were presented in near space, as measured with an audiometer from the distance at which the ears of the participant were located during the experiment [near space ~90 dB(A) SPL, far space ~80 dB(A) SPL]. In addition to objective measurements of SPL and inspection of the direct-to-reverberant profile of the auditory cues, we also behaviorally confirmed these properties in a pilot study in which we examined whether the 3-D localization performance of the auditory cues was above chance[11].

Before the start of the main experiment, each participant performed a short sound localization task to verify that he or she was able to hear whether auditory cues were presented on the left or the right side of the room. The sound localization task consisted of 20 presentations of the auditory cue from a random speaker (five presentations from each of the four locations). Participants had to maintain fixation on a fixation cross (0.5° × 0.5°, 0.20 cd/m^2 as measured with a PhotoResearch SpectraScan PR 650 spectrometer) presented on a light gray background (4.79 cd/m^2) in the center of a screen in near space (at 80 cm distance). The participants were instructed to indicate

---

[11] Ten participants who were not part of the main experiment sample (five female, five male; mean age = 25.70 years, *SD* = 3.13) were tested with a four-choice localization task in the same setup as in the main experiment. They were instructed to maintain fixation on a black fixation cross (0.5° × 0.5°, 0.20 cd/m^2) presented on a light gray background (4.79 cd/m^2) in near space (80 cm). The average accuracy of 3-D localization was significantly above chance ($M = .55$, $SD = .16$) [$t(9) = -6.00$, $p < .001$].

whether the sound was coming from the left or the right side of the room by using two buttons. All participants performed the hearing task with above-chance accuracy.

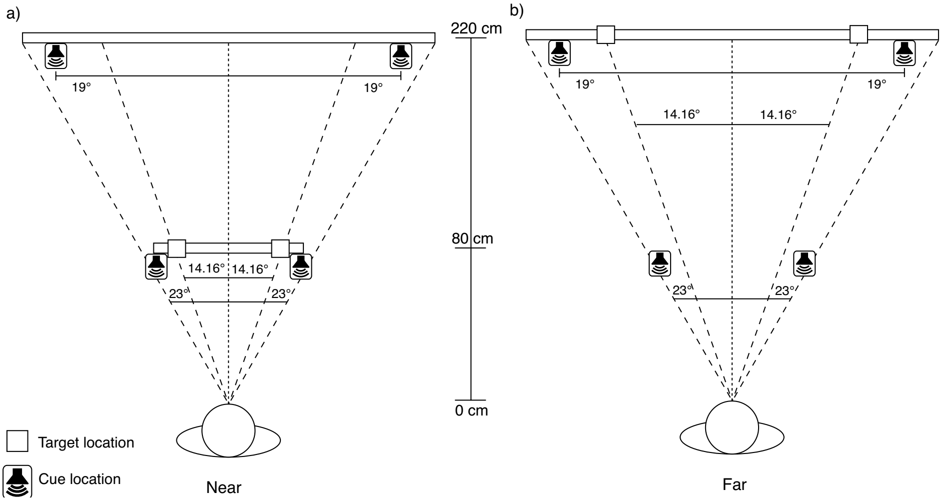


Figure 1. Schematic top views of the experimental setup in the near-space (left panel) and far-space (right panel) conditions.

As in the hearing task, participants in the main experiment were tested in a darkened room, with only the light of the projector illuminating the room. Visual targets were projected in either near (80 cm) or far (220 cm) space and were corrected for visual angle. Presenting stimuli at the same visual angle in different planes of depth has also been done in other studies in which attention was investigated in depth with rather large distances between depth planes (e.g., Couyoumdjian et al., 2003; Downing & Pinker, 1985). In the near-space condition, both screens were present, but targets were only projected on the near-space screen. In the far-space condition, the near-space screen was removed from the setup (target sizes were corrected for visual angle digitally). Speakers were positioned at four locations: near-left, near-right, far-left, and far-right space. Schematic top views of the experimental setup are shown in Figure 1. Each block, visual targets were presented in one space only (near or far space), whereas auditory cues were randomly presented from one of the four speakers located in near or far space. The order of the regions of space in which visual targets were presented first was blocked and counterbalanced across participants. Each speaker was placed outside the light of the projection of the beamer on the left and right sides of the screen. As a result, the speakers in near space were located 23° from the fixation cross and 19° in far space. The experiment started with 20 practice trials for the participants to get used to the task.

Participants were instructed to gaze at a gray fixation cross (size: 0.91° × 0.91° in near and in far space, 0.38 cd/m^2 in near space) presented on a black background (<0.15 cd/m^2 in near and far space) in the center of the screen at a height of 34 cm above the table[12]. After 1,000 ms, an auditory cue was presented at one of the four locations while the fixation cross remained on the screen. Auditory cues could be presented on the same side as (valid) or on the opposite side from (invalid) the visual target, and at the same (valid) or at a different (invalid) depth. The procedure of the experiment is shown in Figure 2.

Also, in a no-cue condition, no cue was presented before the target appeared. Targets were presented in two regions of space, resulting in ten conditions. Each condition contained 80 trials, adding up to a total of 800 trials. A break was provided after 200 trials during a block, and participants could press the space bar to continue. After 400 trials, the visual targets were presented in the other region of space (first far and then near, or vice versa), and another 400 trials were presented. The stimulus onset asynchrony varied between 90 and 250 ms and was always followed by the presentation of a target. The targets were filled gray circles with a diameter of 2.600 in either near space (0.38 cd/m^2) or far space. The target location was randomized and could be presented either above or below the vertical center of the screen, to the left or the right of the fixation cross. The horizontal distance from the fixation cross to the target was 14.160, and the vertical distance from the fixation cross to the middle of the target (either above or below) was 3.9°. The target disappeared upon response. Participants were instructed to press the number-pad "5" key for an upper target, and the number-pad "2" key for a lower target. The maximum response duration was set to 2,000 ms, after which the target disappeared automatically. The intertrial interval consisted of the presentation of the background alone, with a duration of 1,200 ms.

---

[12] Note that we did not track eye movements, and therefore that we cannot draw any conclusions about whether any exogenous cueing effects were due to covert or overt exogenous orienting of attention. The results will still be able to indicate whether attention can be exogenously shifted in 3-D space in a crossmodal way.

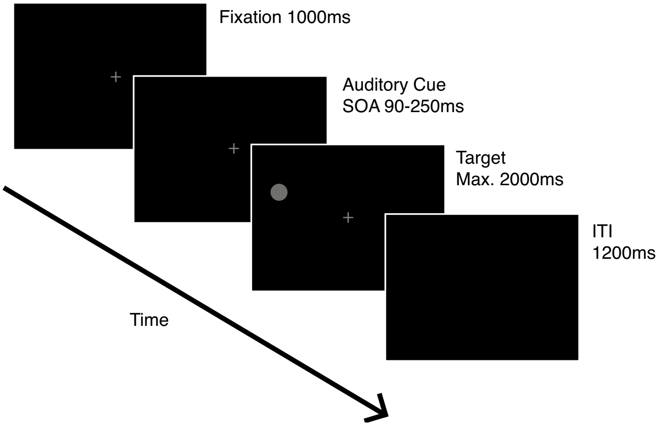


Figure 2. A schematic representation of the procedure of the experiment.

Data analysis

Preprocessing

Practice trials were excluded from both the accuracy and the response time (RT) analyses. We only analyzed the RTs of correct trials. In addition, trials on which the RT was below 100 ms or above 1,000 ms were removed from further analysis, since they were considered to be the results of anticipation or of not attending to the experiment, respectively. RTs were regarded as outliers when they exceeded two- and-a-half standard deviations above or below the group mean of a condition. On average, 6% of the trials were removed from further analysis when targets were presented in near space, and 5% of the trials were removed when targets were presented in far space.

Statistical analysis

First, to investigate whether the overall effect of horizontal cue type (i.e., a "2-D" cueing effect), we performed a repeated measures analysis of variance (ANOVA) with the within-subjects factor Horizontal Cue Type (no cue, valid cue, invalid cue), with RTs being averaged over radial cue validities and distances of the target.

Second, to more closely investigate how the distance between the cue and the target in the radial plane influenced RTs, we performed a 2 × 2 × 2 repeated measures ANOVA on RTs and accuracy, with the factors Target Space (near, far), Horizontal Cue Validity (valid, invalid), and Radial Cue Validity (valid, invalid). Note that we did not include the no-cue condition in the design, because the no-cue condition could not be valid or invalid in the horizontal and radial dimensions. Paired-samples t-tests were done to compare differences between each of the conditions, and the resulting p values were Bonferroni corrected where applicable.

Results

Left–right sound localization task
None of the participants were excluded from the analysis on the basis of the short left–right hearing task, and the group as a whole performed well above chance (mean accuracy = .97, *SE* = .01), $t(15) = 42.37$, $p < .001$. All participants performed the task with an accuracy of at least .85.

Accuracy
The repeated measures ANOVA with the factors Target Space (near, far), Horizontal Cue Validity, and Radial Cue Validity revealed no significant main effects or interactions (all *p*s > .05). The average accuracy was .976 (*SE* = .005). All participants performed with an accuracy of at least .91.

Response times
The results of the repeated measures ANOVA indicated a significant effect of cue type [valid, invalid, or no cue: $F(1.120, 16.799) = 109.946$, $p < .001$, Greenhouse–Geisser epsilon = .560, partial $\eta^2 = .880$]. Pairwise comparisons indicated that RTs were significantly slower in the no-cue condition (*M* = 521 ms, *SE* = 15) than in the invalid condition (470 ms, *SE* = 13, p < .001) and the valid condition (462 ms, *SE* = 13, p < .001). In addition, RTs were significantly faster on valid than on invalid trials (*p* = .001). This indicated that both types of auditory cues facilitated RTs to the target, relative to when no auditory cue was presented (i.e., an alerting effect), with valid cues resulting in the fastest responses.

The analysis with the factors Target Space, Horizontal Cue Validity, and Radial Cue Validity revealed a significant main effect of horizontal cue validity [$F(1, 15) = 24.374$, $p < .001$, partial $\eta^2 = .619$]. RTs on horizontally validly cued trials were significantly faster than those on horizontally invalidly cued trials (462 ms, *SE* = 13, vs. 470 ms, *SE* = 13). We did not find a main effect of target space [$F(1, 15) = 1.587$, $p = .227$, partial $\eta^2 = .096$] and of radial cue validity [$F(1, 15) = 0.055$, $p = .818$, partial $\eta^2 = .004$]. The interaction between horizontal cue validity and radial cue validity was significant [$F(1, 15) = 6.390$, $p = .023$, partial $\eta^2 = .299$]. The magnitude of the horizontal validity effect depended

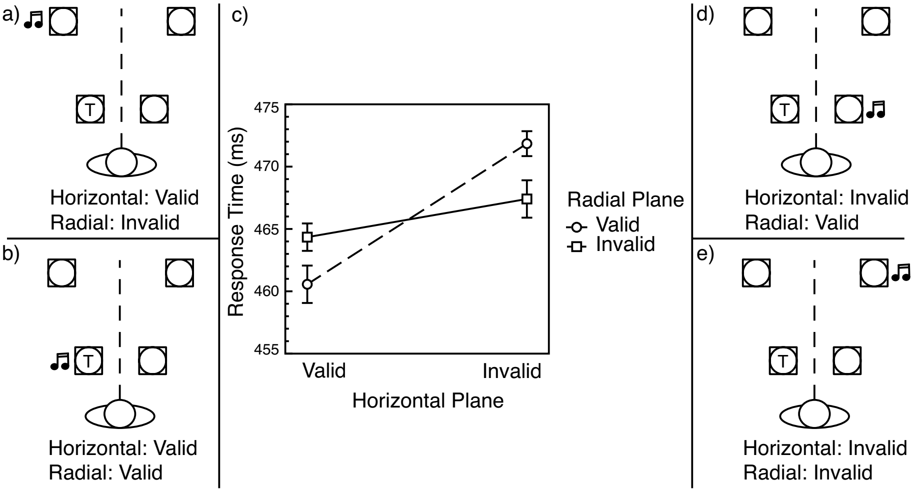


Figure 3. Schematic representations of each cue condition and their corresponding data-point depicted in panel c are shown in panels a, b, d, and e. A note represents an auditory cue location and "T" indicates a target location (target distance was blocked). Panel c depicts the mean response times for valid and invalid trials in the horizontal and radial plane. Error bars represent standard errors of the mean without between-subject variability for graphical purposes (Cousineau, 2005).

on whether the radial distances between the cue and the target were the same. We collapsed the near- and far-space conditions in subsequent analyses, because of the lack of a main effect of target space. Figure 3c shows the mean RT in each condition. The difference between the horizontal valid (461 ms, *SE* = 13) and horizontal invalid (472 ms, *SE* = 14) conditions was 11 ms when the cue was presented at the same depth as the target (radial valid). When the cue and the target were presented at different depths (radial invalid), the difference between the horizontal valid (464 ms, *SE* = 14) and horizontal invalid (467 ms, *SE* = 13) conditions was 3 ms. None of the other interactions were significant (all *p*'s > .1).

Planned comparisons revealed a validity effect (horizontal invalid– horizontal valid) when the cue and target were presented at the same distance [mean horizontal validity effect = 11 ms, *SE* of the difference = 2 ms; $t(15) = -5.360$, $p < .001$], but not when they were presented at different distances [mean horizontal validity effect = 3 ms, *SE* of the difference = 2 ms; $t(15) = -1.359$, $p = .350$]. The size of the validity effect was significantly larger when the cue and target were presented at the same distance (11 ms), as compared to when they were presented a different distances (3 ms) [$t(15) = 2.528$, $p = .023$]. The difference between radial valid and radial invalid cues was not significantly different when cues were horizontally valid (mean radial validity effect = 4 ms) [$t(15) = -1.744$, $p = .194$], nor when cues were horizontally invalid (mean radial validity effect = −4 ms) [$t(15) = 2.092$, $p = .105$].

Discussion

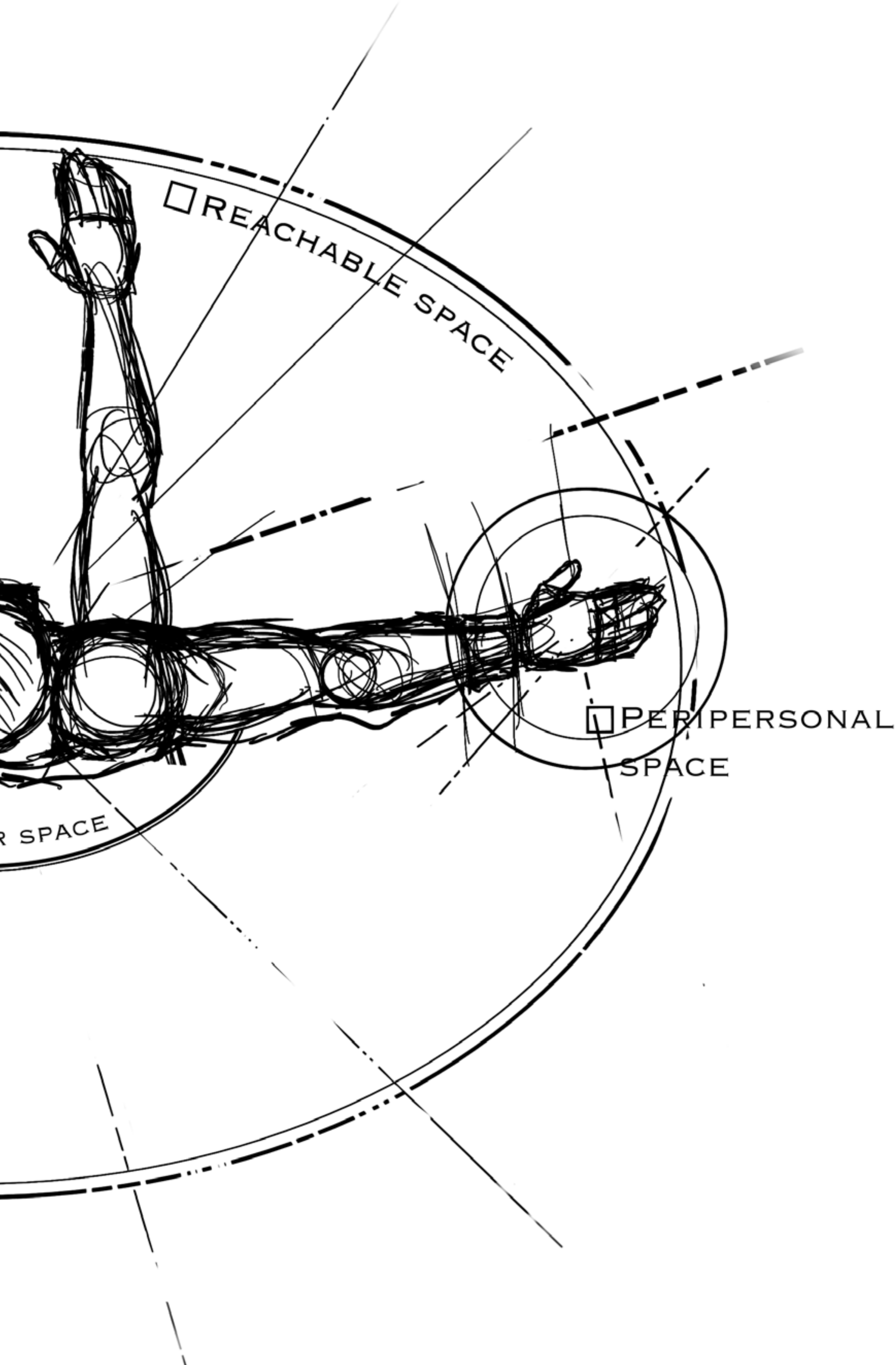
The aim of the present study was to investigate the nature of exogenous crossmodal attention in 3-D space. We presented visual targets in either near or far space, and exogenous auditory cues from one of four locations: near left, near right, far left, or far right. The results indicated an overall cue validity effect in the horizontal dimension. More interestingly, the presence of a horizontal validity effect was dependent on whether the cue and the target were presented at the same depth: A validity effect was only present when the cue and the target were presented at the same depth, and not when the cue and the target were presented at different depths. These findings suggest that exogenous crossmodal attention is "depth-aware." In contrast, if exogenous crossmodal attention were not "depth-aware," we should have observed a horizontal validity effect when the cue and the target were presented both at the same and at different depths from the target. This was not the case.

In our study, the horizontal validity effect for cues and targets presented at the same depth did not differ between visual targets presented in near and far space. This finding is in contrast with those from other studies on attentional orienting in depth, in which an asymmetry was observed between the cueing effects for targets in near and targets in far space (e.g., Chen, Weidner, Vossel, Weiss, & Fink, 2012; Downing & Pinker, 1985). In these studies, in which endogenous attention was manipulated, participants were faster to respond to targets that were presented between the participant and the focus of endogenous attention, as compared to targets that were presented beyond the focus of endogenous attention. The lack of an asymmetry in attentional reorienting in depth in the present study may be explained by the fact that we blocked target distance: In our study, participants had no need to attend to multiple planes of depth, which might have caused an endogenous focus on one depth plane (i.e., 100% endogenous validity), possibly overruling any asymmetry of attention in depth.

The conclusion that exogenous crossmodal attention is "depth-aware" seems to be in contrast with the results of the short four-choice localization task. Although participants could localize the cue significantly above chance, their accuracy was rather low (mean = 55%). Still, the depth of the cues in the main experiment influenced the presence of the horizontal validity effect. This suggests that despite participants' being poor at consciously locating this type of sound (a sine wave), the brain did process the depth information of auditory sources.

Our findings are in line with the results from studies in which exogenous intramodal orienting of attention was investigated in 3-D space (Atchley et al., 1997; Theeuwes & Pratt, 2003). For example, Theeuwes and Pratt also found that the validity effect was stronger when the cue and target were presented at the same distance from the observer. Here, however, we extended these findings by showing that the crossmodal exogenous cues were also able to automatically attract attention to different depth planes in "real" 3-D space. Altogether, the previous and present results therefore seem to fit with the theory of a supramodal attentional system (Eimer & Van Velzen, 2002; Farah et al., 1989; Macaluso, Frith, & Driver, 2002) that processes spatial information from the auditory and visual modalities, despite differences in spatial reference frame (retinotopic [Gardner, Merriam, Movshon, & Heeger, 2008] vs. head-centered [Andersen, 1997]). A candidate region for supporting such a supramodal attentional system is the posterior parietal cortex, and more specifically the multisensory lateral intraparietal area (area LIP; Andersen, 1997). Crossmodal interactions also seem to depend on feedforward and feedback connections between unimodal and multisensory areas (Macaluso & Driver, 2005).

Taken together, our results indicate that the exogenous orienting of crossmodal attention is "depth-aware," and they contribute to the further understanding of crossmodal interactions in 3-D space.


☐Reachable space
☐Peripersonal space
r space

Chapter 8

Exploring near and far regions of space: Distance-specific visuospatial neglect after stroke

Van der Stoep, N., Visser-Meily, J. M. A., Kappelle, L. J., De Kort, P. L. M., Huisman, K. D.,
Eijsackers, A. L. H., Kouwenhoven, M., Van der Stigchel, S., & Nijboer, T. C. W. (2013).
Exploring near and far regions of space:
Distance-specific visuospatial neglect after stroke.
Journal of Clinical and Experimental Neuropsychology, 35(8), 799-811.

Abstract

Visuospatial neglect has been observed in the horizontal (left/right) and vertical (up/down) dimension and depends on the distance at which a task is presented (near/far). Previously, studies have mainly focused on investigating the overall severity of neglect in near and far space in a group of neglect patients instead of examining subgroups of neglect patients with different types of distance-specific neglect. We investigated the spatial specificity (near vs. far space), frequency, and severity of neglect in the horizontal and vertical dimensions in a large group of stroke patients. We used three tasks to assess neglect in near (30 cm) and far (120 cm) space: a shape cancellation, letter cancellation, and a line bisection task. Patients were divided into four groups based on their performance: a group without neglect (N–F–), a near only neglect (N+F–), a far only neglect (N–F+), and a near and far neglect group (N+F+). About 40% of our sample showed neglect. Depending on the task, N+F– was observed in 8 to 22% of the sample, whereas N–F+ varied between 8% and 11%, and N+F+ varied between 11% to 14% of the sample. The current findings indicate that horizontal and vertical biases in performance can be confined to one region of space and are task dependent. We recommend testing for far space neglect during neuropsychological assessments in clinical practice, because this cannot be diagnosed using standard paper-and-pencil tasks.

Introduction

Hemispatial neglect, also known as neglect, is a disabling disorder that frequently occurs after right-hemisphere stroke (Bowen, McKenna, & Tallis, 1999; Ringman, Saver, Woolson, Clarke, & Adams, 2004), suggesting a special role for the right hemisphere in spatial attention (Halligan, Fink, Marshall, & Vallar, 2003; Heilman, Watson, & Valenstein, 2003; Shulman et al., 2010). It refers to the failure to report, respond, or orient to stimuli on the contralesional side of space or body that cannot be accounted for by primary sensory or motor deficits (Halligan & Marshall, 1991; Heilman et al., 2003; Robertson, & Halligan 1999). Neglect has been observed in several sensory modalities (visual, auditory, and tactile, e.g., Barbieri & de Renzi, 1989) and is associated with poor functional recovery (Cherney, Halper, Kwasnica, Harvey, & Zhang, 2001; Jehkonen, Laihosalo, & Kettunen, 2006). Spontaneous recovery of neglect appears to occur mainly during the first 12 to 14 weeks after stroke (Nijboer, Kollen, & Kwakkel, in press), although 30% to 40% of the neglect patients did not fully recover and still had neglect a year after stroke (Cassidy, Lewis, & Gray, 1998; Jehkonen, Laihosalo, Koivisto, Dastidar, & Ahonen, 2007; Nijboer, Van de Port, Schepers, Post, & Visser-Meily, 2013). Clinical manifestations of neglect vary widely, which is consistent with the idea of neglect being a multicomponent syndrome (Vuilleumier et al., 2008): Patients can experience impairments in visual, auditory, tactile, and/or motor abilities (Bisiach, Cornacchia, Sterzi, & Vallar, 1984; Laplane & Degos, 1983; Pierson-Savage, Bradshaw, Bradshaw, & Nettleton, 1988) in perceptual as well as in representational space (Berti & Frassinetti, 2000; Bisiach & Luzzatti, 1978; Halligan & Marshall, 1991; Meador, Loring, Bowers, & Heilman, 1987; Mennemeier, Wertman, & Heilman, 1992). Perceptual neglect has been linked to lesions in the inferior parietal lobe, whereas lesions in the dorsolateral prefrontal cortex have been associated with visuomotor neglect (Verdon, Schwartz, Lovblad, Hauert, & Vuilleumier, 2010).

Although neglect may be characterized as a disorder in which patients show a spatial attentional bias in the horizontal dimension, this has also been found in the vertical dimension (Pitzalis, Di Russo, Spinelli, & Zoccolotti, 2001; Pitzalis, Spinelli, & Zoccolotti, 1997; Rapcsak, Cimino, & Heilman, 1988). Additionally, the presence and/or severity of (horizontal and vertical) neglect can depend on the distance at which visual information is presented: within reaching distance (i.e., near space) versus beyond reaching distance (i.e., far space; Aimola, Schindler, Simone, & Venneri, 2012; Pitzalis et al., 2001). Several group studies have reported that neglect was more severe in far than in near space, as measured with line bisection paradigms (Cowey, Small, &

Ellis, 1994, 1999; Keller, Schindler, Kerkhoff, von Rosen, & Golz, 2005) and visual search paradigms (Butler, Eskes, & Vandorpe, 2004), but the opposite has been found as well, as measured with line bisection and target cancellation tasks (Aimola et al., 2012). The effect of distance on neglect is not always the same across different tasks in the same patient, as was illustrated by a study that showed an influence of distance on severity of neglect on a line bisection task, but not on a cancellation task (Keller et al., 2005). Previc (1998) suggested that different neuroanatomical structures are involved in processing sensory information from near compared to far space. More specifically, the inferior parietal cortex appears to be more involved in near space processing, whereas the superior and medial temporal cortex appear to be more involved in far space processing (for more candidate structures see Table 2 in Previc, 1998). Brain regions that are associated with near and far neglect by means of lesion analyses are in line with this idea of a dorsal (near)–ventral (far) stream distinction (Aimola et al., 2012; Butler et al., 2004).

All in all, these studies suggest that neglect can be present in horizontal (left/right) and vertical (up/down) space and that the presence and/or severity can be influenced by the distance from the observer at which a task is presented (near/far space). However, the influence of distance on the attentional biases in both of these dimensions has not yet been investigated simultaneously in stroke patients. Although previous studies have investigated neglect in near and far space (e.g., Aimola et al., 2012; Pitzalis et al., 2001), they did not measure biases in performance in horizontal and vertical space in the same task, they used relatively small sample sizes, and the samples in these studies consisted predominantly of patients with right-hemisphere lesions, who sometimes suffered from brain tumors. The sample included in the current study contains not only about the same number of patients with left- and right-hemisphere lesions, but also some patients with lesions in both hemispheres. This enables a broader view on the presence of distance-specific neglect in patients with left-sided, right-sided, or bilateral brain damage. Additionally, the presence of brain tumors was an exclusion criterion. The presence and severity of neglect appear to depend on several factors such as (a) the spatial location and orientation of the perceived stimuli, (b) the type of task used to assess neglect, (c) the type of visuospatial operation that is required to perform the test, and (d) the lesion location.

The aim of the current study was to investigate the frequency, spatial specificity (near vs. far space), and severity of neglect in two spatial dimensions (horizontal and vertical space in the cancellation tasks and horizontal space in the line bisection task) in a large sample of stroke patients. In addition, we investigated the relation between the various visuospatial operations in different regions of space and in what way neglect on a specific task was related to neglect on another task.

Methods

Participants

The criteria for admission in a rehabilitation center in the Netherlands are: (a) the patient cannot be discharged home, but is expected to return home in view of the prognosis and availability of the caregivers; (b) the patient is able to learn and is sufficiently motivated; (c) the patient has sufficient vitality; (d) the rehabilitation goals are complex and need a multidisciplinary approach; (e) return to work may be possible; and (f) a relatively high rate of rehabilitation is possible. A group of 109 participants (61 without neglect, 48 with neglect on at least one task) were selected from stroke patients consecutively admitted for inpatient rehabilitation to Rehabilitation Center De Hoogstraat located in Utrecht in The Netherlands, according to the following inclusion criteria: (a) age between 18 and 85 years; (b) no severe deficits in communication and/or understanding; (c) normal or corrected-to-normal visual acuity; and (d) the ability to perform at least two tasks from our near/far neglect screening at both distances. An overview of the characteristics of the neglect and no neglect group is shown in Table 1. The groups did not differ in terms of time post stroke in weeks, $t(61.00) = -1.55$, $p = .13$, age, $t(107) = 0.35$, $p = .73$, sex, $\chi^2(1) = 0.60$, $p = .44$, the number of left- and right-hemisphere lesions, $\chi^2(1) = 2.66$, $p = .103$, Barthel Index score, $t(82) = 0.31$, $p = .76$, and Mini-Mental State Examination (MMSE) score, $t(61) = -0.29$, $p = .77$.

There are currently no normative data for our screening tasks. To be able to indicate whether the performance of patients deviated from healthy individuals, we needed to compare their performance with a sample of healthy control subjects. Therefore, we recruited 28 healthy individuals (16 male, mean age = 42.32 years, *SD* = 20.31) as a control group for performance on the shape cancellation, letter cancellation, and line bisection tasks in both near and far space.

Table 1 Characteristics per group (neglect vs. no neglect; L = Left hemisphere, R = Right hemisphere). Patients in the no neglect group did not show neglect on any of the tasks.

Clinical variables	Results neglect (*SE*)	Results no neglect (*SE*)
Group size	61	48
Time post-stroke (weeks)	13.14 (5.49)	4.58 (.50)
Age in years	58.25 (1.59)	59.06 (1.74)
Gender (% Male)	66%	58%
Hemisphere of stroke		
Left (*N*)	22	24
Right (*N*)	37	21
Both (*N*)	2	2
Unknown (*N*)	0	1
Barthel Index	14.69 (2.61)[a]	15.86 (2.53)[b]
MMSE	26.49 (0.60)[c]	26.19 (0.86)[d]

Note. Neglect versus no neglect. Patients in the no neglect group did not show neglect on any of the tasks. MMSE = Mini-Mental State Examination. a Information on the Barthel index was available from 80% of the neglect group. b Information on the Barthel index was available from 73% of the no neglect group. c Information on the MMSE was available from 61% of the neglect group. d Information on the MMSE was available from 54% of the no neglect group.

Stimuli, task, and procedure

To test for the spatial specificity, frequency, and severity of neglect, a neglect screening was administered to all patients. Spatial specificity was defined in terms of whether a patient showed neglect in a single region of space (i.e., near or far only) or in both regions of space (i.e., near and far). The procedure that we used to indicate whether a patient showed neglect in a region of space is described in the *Data*

Preprocessing

The neglect screening included three tasks that are often used to test for the presence of neglect in the clinical practice: two target cancellation tasks (shape cancellation, letter cancellation) and a line bisection task. Each task was performed in two conditions. In the near space condition, stimuli were presented on a monitor at a distance of approximately 30 cm, whereas in the far space condition the stimuli were presented on a monitor at a distance of approximately 120 cm. Stimuli were enlarged in the far space condition to control for visual angle. The order of the tasks and the distance at which a task was first presented was randomized across patients. All subjects (i.e., patients and healthy control subjects) were tested individually in a quiet room, were seated in front of a monitor, and received specific instructions per task.

The shape cancellation task consisted of a field of 54 targets shapes (0.6° × 0.6°) among 75 distractor shapes of various sizes (with widths ranging from 0.95° to 2.1° and heights ranging from 0.45° to 0.95°). The stimulus presentation was corrected for visual angle and was approximately 18.5° wide and 11° high at both distances. Subjects were instructed to find all the target shapes presented on the screen and to click on them. A circle appeared on the screen around the location of each mouse click and remained on screen during the test.

The letter cancellation task consisted of five strings of 34 letters (0.6° × 0.6°). The stimulus presentation was controlled for visual angle and was approximately 19° wide and 5.7° high at both distances. Subjects were asked to cross out 40 target letters among the distractor letters, by moving the cursor with a mouse and clicking on the target letters. A circle appeared on the screen around the location of each mouse click and remained on screen during the task.

The line bisection task consisted of three horizontally oriented lines that were evenly distributed across the screen in vertical space. The middle line was presented in the horizontal and vertical center of the screen, the top line was presented above the vertical center of the screen and shifted to the right, whereas the bottom line was presented below the vertical center of the screen and shifted to the left. The amount of vertical shift was always 28% of the line length, and the amount of horizontal shift was always 15% of the line length in both near and far space. Lines were controlled for visual angle and were approximately 22° long and 0.2° thick. Subjects were asked to indicate the center of each line by moving the cursor with the mouse and clicking on the subjective midpoint of each line, starting at the topmost line and working their way down. This task was performed four times in a row, resulting in a total of 12 lines presented in each region of space.

When a patient could not use the dominant hand (e.g., due to hemiplegia, hemiparesis, etc.), the non-dominant hand was used. This was feasible as they were accustomed to work with their nondominant hand during other rehabilitation programs (e.g., physical therapy, occupational therapy, etc.).

Data preprocessing

For the cancellation tasks we used a difference score of the amount of omissions between the left and the right part of the stimulus field of at least two as a rough indication of neglect. We based this difference score on the average difference score of the healthy control subjects plus three standard deviations. On the shape cancellation task the average difference score of the control group was 0.107 (*SD* = 0.32) in near space and 0.04 (*SD* = 0.19) in far space. The three-standard-deviation cut-off was 1.05 in near and 0.60 in far space on the shape cancellation task. On the letter cancellation task, the average difference score of the control group was 0.25 (*SD* = 0.52) in near space and 0.21 (*SD* = 0.50) in far space. This resulted in a three-standarddeviation cut-off of 1.81 in near space and 1.71 in far space on the letter cancellation task. A difference score of at least two omissions falls outside of the normal range of our healthy control group on both the shape and the letter cancellation, and we therefore used this as an indication of neglect. This method of establishing a normal range for each task has also been reported in other studies (e.g., Stone et al., 1991). Our specific cut-off is also in line with other studies in which a difference score of at least two was used to provide an indication of neglect (e.g., Nijboer, Kollen, et al., 2013; Nijboer, Van de Port, Schepers, Post, & Visser-Meiley, 2013). We divided patients into four groups based on whether their difference score indicated neglect or not: no neglect (N–F–), neglect in near space only (N+F–), neglect in far space only (N–F+), or neglect at both distances (N+F+).

The horizontal and vertical normalized center of cancellation (respectively, CoC-x and CoC-y) was calculated for both cancellation tasks based on all the targets in a task (Rorden & Karnath, 2010). The CoC takes both the *amount* and the *location* of cancelled targets into account. It is therefore more indicative than the number of omissions in each half of the stimulus field. CoC-x and CoC-y scores could range from –1 to 1. When targets are missed in the upper left corner of the stimulus field, the CoC-x shifts to the right (closer to 1), and the CoC-y shifts down (closer to –1). In contrast, when targets are missed in the lower right part of the stimulus field, the CoC-x shifts to the left (closer to –1), and the CoC-y shifts up (closer to 1). A CoC-x or y score of zero indicated that there was no spatial bias in the number of missed targets. To analyze the CoC-x and y, we used the absolute value of the normalized CoC, because our sample contained both patients with left-sided neglect and those with right-sided neglect. Left-sided neglect would result in positive CoC-x values, and right-sided neglect would result in negative CoC-x values. Analyzing the CoC on a group level without using absolute values would distort the results, because positive values will cancel out negative values, and this would not show the overall effect of spatial bias in patients with neglect. Analyzing both the CoC-x and y allowed to us to investigate biases in both horizontal and vertical space. We presented the task in both near and far space, and this allowed us to analyze performance in three dimensions for each cancellation task.

To analyze performance on the line bisection, we calculated the deviation between the actual center and the subjective midpoint of each line in degrees of visual angle.

Next, we calculated the average deviation on all 12 lines in each region of space and compared this to the control group of 28 healthy control subjects. Negative deviation values indicated a shift of the subjective midpoint to the left of the actual center, whereas positive deviation values indicated a shift to the right of the actual center. As on the cancellation tasks, performance indicated neglect when the average score of a patient was outside the normal range of the control group (mean ± 3 *SD*s) in a region of space. In the control group the average deviation from the actual center was –0.13 degrees of visual angle (*SD* = 0.20°) in near space and –0.15 degrees of visual angle (*SD* = 0.24°) in far space. The normal range (mean ± 3 *SD*s) was between –0.74 and 0.48 degrees of visual angle in near space and between –0.86 and 0.56 degrees of visual angle in far space. Patients were divided into neglect groups based on whether their average deviation score fell utside the normal range in a region of space. This again resulted in four groups: patients without neglect (N–F–), with neglect in near space only (N+F–), with neglect in far space only (N–F+), or with neglect at both distances (N+F+). To further analyze the results of the line bisection task we then used the absolute values of the average deviation from the center, because of the same reasons that were mentioned before (i.e., the sample contained patients with left- and right-sided neglect). We included the line bisection task in the current study to investigate the influence of distance on neglect using a more perceptual task and to be able to compare performance of patients in near and far space between cancellation and bisection tasks.

Statistical analyses

For each task, a 2 × 4 mixed repeated measures analysis of variance (ANOVA) was done with space (near/far) as within-subjects factor and neglect type (N–F–/N+F–/N–F+/N+F+) as between-subject factor. In order to gain power for the analyses, we analyzed both patients with left-sided neglect and those with right-sided neglect together. Thus we used the *absolute* values of the CoC-x and CoC-y as dependent variables in the analyses of the cancellation tasks. Although the relation between neglect type and CoC-x may appear circular because the groups are based on a difference score of missed targets in the left and right part of the task, this is not the case. A difference score of, for example, two can lead to totally different CoC-x values, depending on the location of the missed targets. We also used the absolute value of the average deviation in degrees of visual angle from the center of the line in the analysis of the line bisection test for the same reasons as on the cancellation tasks. It is important to note that the absolute values now only reflect the severity of imbalance in performance on the cancellation tasks and only the amount of deviation from the center of the line, not its direction (to the left or the right). Whenever a main effect of space or the interaction between space and neglect type was significant, three independent t-tests were done for each region of space to compare the performance of each neglect group to the N–F– group in that region. For each of the tasks, we also compared the performance of the N–F– group (brain damage, but no neglect) with that of the healthy control group using independent-samples *t*-tests to investigate whether their performance on each of the tasks was within the normal range.

To investigate whether the allocation of patients to different neglect types differed across tasks, we performed a Pearson chi-squared test on the frequencies of neglect in each group on each of the three tasks post hoc. There are conflicting findings on the influence of lesion location (left vs. right hemisphere) on the severity of neglect for cancellation tasks with either verbal (i.e., letters) or nonverbal stimuli (i.e., shapes). Some researchers observed more severe neglect for nonverbal stimuli after right-sided lesions than for verbal stimuli (e.g., Leicester, Sidman, Stoddard, & Mohr, 1969; Weintraub & Mesulam, 1988), whereas others found no differences (e.g., Caplan, 1985). To investigate the relation between the hemisphere of stroke and the presence of neglect on the shape and the letter cancellation, we performed a post hoc Pearson chi-squared test on the neglect frequency after left- and right-hemisphere lesions.

All reported *p* values of the follow-up tests are two-tailed and corrected for the number of comparisons with the formula described in Motulsky (1995): $p = 1 - (1 - p)^n$.

Results and discussion

Frequency of neglect types

In total, 107 patients were tested with the shape cancellation test, whereas 81 patients were tested with the letter cancellation test, and 82 patients were tested with the line bisection test. Based on their difference score on the shape cancellation task, 30% of the patients were diagnosed with neglect, whereas this was 47% on the letter cancellation test, and 30% on the line bisection test. Note that the total number of patients that was assessed differs for each of the tasks, and that neglect type could differ across tasks (see the *Neglect Consistency* section). Overall, approximately 72% of the patients that showed neglect on the shape cancellation task had an impaired performance in far space (the percentage of patients with N–F+ and N+F+), whereas this was approximately 53% on the letter cancellation task, and 64% on the line bisection task. The percentage of patients that had neglect in near space (the percentage of patients with N+F– and N+F+) was 72% on the shape cancellation task, 76% on the letter cancellation task, and 72% on the line bisection task. The percentages of patients in each of the groups on each of the tasks are shown in Table 2. The results of the Pearson chi-squared test on the frequencies of distance-specific neglect diagnoses (N–F–, N+F–, N–F+, N+F+) on each of the tasks indicated that the distribution of frequencies across neglect types did not differ across tasks, $\chi^2(1) = 10.61$, $p = .10$.

Table 3 shows the percentage of patients with left- and right-sided neglect based on difference scores and the nonabsolute values on the three tasks. As can be seen from this table, predominantly left-sided neglect was observed in the N+F+ group, whereas no clear pattern was observed in the region-specific groups (N+F–, N–F+). To be able to analyze the performance of both left- and right-sided neglect patients, absolute performance measures were used.

Spatial specificity and severity of neglect

As mentioned above, spatial specificity is defined by the presence of neglect in a region. We used difference scores on the cancellation tasks to create distance-specific neglect groups and analyzed the horizontal and vertical CoC of these groups to investigate whether the performance between groups was significantly different on this more subtle measure. Severity of neglect is reflected by the CoC value, with numbers closer to 1 or –1 indicating that performance is more lateralized.

Table 2 Percentage of patients in each group on the three tasks.

	Shape Cancellation (N=107)	Letter Cancellation (N=81)	Line Bisection (N=82)
N-F-	70%	53%	70%
N+F-	8%	22%	11%
N-F+	8%	11%	9%
N+F+	14%	14%	11%

Note. N–F– = without neglect, N+F– = near only neglect, N–F+ = far only neglect, N+F+ = near and far neglect.

Table 3 Percentage of patients with left- and right-sided neglect in each neglect group on the three tasks.

Task	Side	N+F-	N-F+	N+F+
Shape Cancellation	Left	56%	38%	87%
	Right	44%	63%	13%
Letter Cancellation	Left	50%	44%	55%
	Right	50%	57%	37%
Line Bisection	Left	22%	86%	89%
	Right	78%	14%	11%

Note. N+F– = near only neglect, N+F+ = near and far neglect. One patient showed an inconsistent pattern on the letter cancellation task: right-sided neglect in near space, and left-sided neglect in far space.

Shape cancellation

The characteristics of each of the four groups based on the performance on the shape cancellation task are shown in Table 4. The groups did not differ in terms of time post stroke in weeks, *F*(3, 103) = 0.07, *p* = .97, age, *F*(3, 103) = 0.15, *p* = .93, sex, $\chi^2(1) = 0.76$, *df* = 3, *p* = .89, and the number of left- and right-hemisphere lesions, χ^2 (1) = 4.95, *p* = .184. A repeated measures ANOVA for the absolute CoC-x showed a significant main effect of space, *F*(1, 103) = 4.21, *p* = .04, indicating that patients deviated more from the horizontal center in near space (mean absolute CoC-x = .05, *SE* = .01) than in far space (mean absolute CoC-x = .04, *SE* = .01) by missing targets either on the left or on the right part of the stimulus field. The results also revealed a significant main effect of neglect type, *F*(3, 103) = 26.52, *p* < .001, showing that the neglect groups differed in their performance on the test. The N+F+ group showed the largest average deviation from the horizontal center (mean absolute CoC-x = .137, *SE* = .01) compared to the other three groups. The performance of the N–F– group did not differ from the healthy control group on the absolute CoC-x in near, *t*(101) = 0.89, *p* = .38, and in far space, *t*(74) = 0.71, *p* = .48. The average values for the absolute CoC-x and -y for each neglect group and each region of space are shown in Table 5.

More importantly, we found a significant interaction between space and neglect type, *F*(3, 103) = 4.74, *p* < .01. The mean absolute CoC-x for each group in each region of space on the shape cancellation task is shown in Figure 1 (left panel). We used planned comparisons to compare the performance of each of the neglect groups with the N–F– group in each region of space (because of the main effect of space and the interaction between space and neglect type). Surprisingly, the N+F– group did not significantly deviate from the N–F– group, *t*(8.03) = –2.28, *p* = .15, uncorrected *p* = .05. As expected, the performance of the N–F+ group did not significantly differ from that of the N–F–

group in near space, $t(8.28) = –0.27$, $p = .99$, and performance of the N+F+ group did significantly differ from that of the N–F– group, $t(14.01) = –3.66$, $p < .01$. The N+F– group did not differ from the N–F– group in far space, $t(8.18) = –1.22$, $p = .53$, whereas the N–F+ and the N+F+ group did: respectively, $t(7.08) = –5.00$, $p < .01$, and $t(14.00) = –2.92$, $p = .03$.

Table 4 Characteristics of each of the groups that are based on the performance on the shape cancellation task.

Clinical variables	N-F- (*SE*)	N+F- (*SE*)	N-F+ (*SE*)	N+F+ (*SE*)
Group size	75	9	8	15
Time post-stroke in weeks	10.41 (3.82)	7.29 (11.02)	6.09 (11.69)	7.87 (8.53)
Age in years	58.61 (1.43)	60.22 (4.12)	56.25 (4.37)	58.27 (3.19)
Gender (N Male)	48	5	5	8
Hemisphere of stroke				
Left hemisphere (*N*)	33	2	5	4
Right hemisphere (*N*)	37	7	3	11
Both (*N*)	4	-	-	-
Unknown	1	-	-	-

Note. N–F– = without neglect, N+F– = near only neglect, N–F+ = far only neglect, N+F+ = near and far neglect.

Table 5 Average values (standard error) for each condition and each group of the absolute CoC-x and y on the shape cancellation task in near and far space.

	Near abs. CoC-x	Far abs. CoC-x	Near abs. CoC-y	Far abs. CoC-y
N-F-	.004 (.007)	.001 (.004)	.006 (.011)	.003 (.005)
N+F-	.046 (.056)	.006 (.012)	.038 (.021)	.003 (.005)
N-F+	.004 (.007)	.030 (.016)	.005 (.005)	.017 (.012)
N+F+	.159 (.164)	.115 (.151)	.038 (.036)	.040 (.033)

Note. N–F– = without neglect, N+F– = near only neglect, N–F+ = far only neglect, N+F+ = near and far neglect. CoC-x and CoC-y = horizontal and vertical normalized center of cancellation, respectively. abs. = absolute. Standard errors in parentheses.

For the absolute CoC-y, we found a main effect of space, $F(1, 103) = 5.73$, $p = .02$, and neglect type, $F(3, 103) = 29.70$, $p < .01$, driven by an interaction between space and neglect type, $F(3, 103) = 12.11$, $p < .01$. Patients' absolute CoC-y deviated slightly more from the vertical center in near (mean absolute CoC-y = 0.02, $SE = 0.002$, than in far space (mean absolute CoC-y = 0.02, $SE = 0.002$). The N+F+ group showed the largest deviation from the vertical center (mean absolute CoC-y = 0.04, $SE = 0.003$) followed by the N+F– group (mean absolute CoC-y = 0.02, $SE = 0.004$), the N–F+ group (mean absolute CoC-y = 0.01, $SE = 0.005$), and the N–F– group (mean absolute CoC-y = 0.004, $SE = 0.002$). The performance of the N–F– group did not differ from that of the healthy control group on CoC-y in near, *t*(101) = –0.58, *p* = .56, and in far space, *t*(28.81) = 0.76, *p* = .45. The average absolute CoC-y for each group in each region of space on the shape cancellation task is shown in Figure 1 (right panel, note the difference in scale of the *y*-axis compared to the left panel).

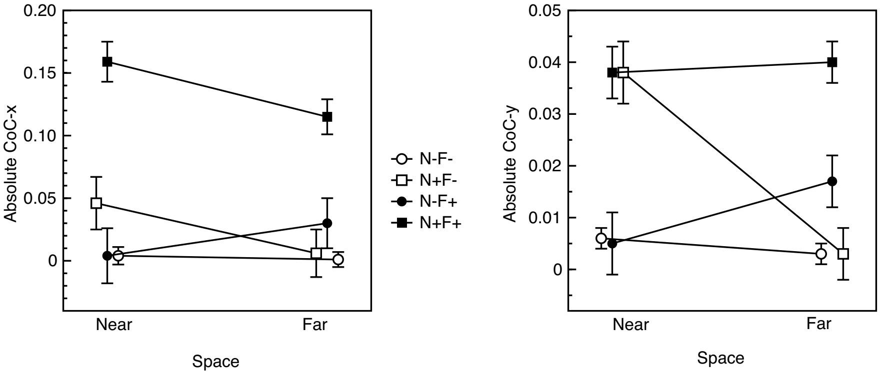


Figure 1. Average absolute CoC-x (left panel) and CoC-y (right panel) for each group at both distances for the shape cancellation task. CoC-x and CoC-y = horizontal and vertical normalized center of cancellation, respectively. N–F– = without neglect, N+F– = near only neglect, N–F+ = far only neglect, N+F+ = near and far neglect. Error bars depict standard error of the mean.

To further investigate the interaction between space and neglect type, we used planned comparisons to compare performance between each of the neglect groups and the N–F– group in each region of space. In near space, the N+F– group deviated from the N–F– group, *t*(8.54) = –4.33, *p* < .01, as did the N+F+ group, *t*(14.54) = –3.34, *p* = .02, but not the N–F+ group, *t*(81) = 0.35, *p* = .98. In far space, the N–F+ group and the N+F+ group deviated significantly from the N–F– group, *t*(7.49) = –3.24, *p* = .04, and *t*(14.25) = –4.34, *p* < .01, respectively, but the N+F– group did not, *t*(11.48) = –0.22, *p* = .86.

In sum, we observed that performance of patients with neglect could be impaired in the horizontal and in the vertical dimension, and that this impairment could be present in both regions of space (near and far), or in only one region of space (near or far). In addition, when performance was impaired in both regions of space, the horizontal bias was more severe than when performance was impaired in only one region of space.

Letter cancellation

A repeated measures ANOVA on the absolute CoC-x showed no significant effect of space, *F*(1, 77) = 0.02, *p* = .89, and no significant interaction between space and neglect type, *F*(3, 77) = 2.59, *p* = .06. A significant main effect of neglect type, *F*(3, 77) = 18.17, *p* < .01, was obtained. In line with the SC, the N+F+ group showed the largest deviation from the horizontal center (mean absolute CoC-x = .12, *SE* = .01) followed by the N–F+ group (mean absolute CoC-x = .03, *SE* = .02), the N+F– group (mean absolute CoC-x = .03, *SE* = .01), and the N–F– group (mean absolute CoC-x = .01, *SE* = .01). All three neglect groups performed significantly different from the N–F– group (all *p* < .05). The performance of the N–F– group did not differ from the healthy control group on CoC-x in near, *t*(69) = –0.92, *p* = .36, and in far space, *t*(69) = –0.19, *p* = .85. Although the interaction between space and neglect type was not significant, the means of each group showed the same interaction pattern as that on the shape cancellation task and are shown in Figure 2 (left panel; note that the groups do not consist of the same patients as those on the shape cancellation test, see section *Neglect Consistency*). The average values of the absolute CoC-x and y for near and far space and for each region of space are shown in Table 6.

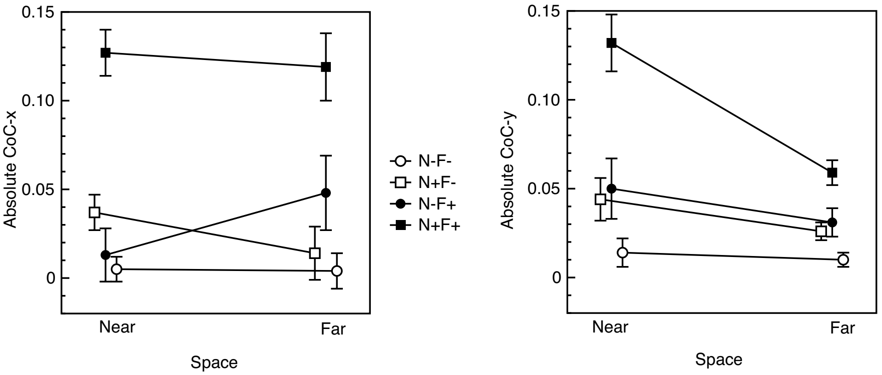


Figure 2. Average absolute CoC-x (left panel) and CoC-y (right panel) for each group at both distances for the letter cancellation task. CoC-x and CoC-y = horizontal and vertical normalized center of cancellation, respectively. N–F– = without neglect, N+F– = near only neglect, N–F+ = far only neglect, N+F+ = near and far neglect. Error bars depict standard error of the mean.

On the absolute CoC-y, a significant main effect of space, *F*(1, 77) = 15.49, *p* < .01, and neglect type, *F*(3, 77) = 23.17, *p* < .01, was obtained, as well as a significant interaction between space and neglect type, *F*(3, 77) = 4.55, *p* < .01. On average, patients deviated more from the vertical center in near space (mean absolute CoC-y = 0.06, *SE* = 0.007) than in far space (mean absolute CoC-y = 0.03, *SE* = 0.003). The largest deviation from the vertical center of the stimulus field was found in the N+F+ group (absolute CoC-y = 0.10, *SE* = 0.009), followed by the N–F+ group (absolute CoC-y = 0.04, *SE* = 0.010), the

N+F– group (absolute CoC-y = 0.04, *SE* = 0.007), and the N–F– group (absolute CoC-y = 0.01, *SE* = 0.005). The performance of the N–F– group did not differ from that of the healthy control group on CoC-y in near, *t*(67.93) = 0.35, *p* = .73, and in far space, *t*(69) = –1.00, *p* = .32.

The interaction between space and neglect type indicated that the amount of deviation from the vertical center depended both on the region of space in which the task was performed and on the group in which the patient was placed based on the difference score of at least two omissions between the left and the right part of the stimulus field. The mean absolute CoC-y for each group in near and far space is shown in Figure 2 (right panel).

Planned comparisons in each region of space were done to compare the absolute CoC-y between each of the neglect groups and the N–F– group. In near space the N+F– group deviated significantly from the N–F– group, *t*(21.85) = –3.17, *p* = 0.01, as did the N+F+ group, *t*(10.17) = –3.29, *p* = .02, but not the N–F+ group, *t*(8.45) = –1.80, *p* = .29[13]. In far space the N+F– group did not deviate from the N–F– group, *t*(21.81) = –2.16, *p* = .12 (see Footnote 1), whereas the N–F+ group, *t*(11.36) = –3.34, *p* = .02, and the N+F+ group, *t*(11.10) = –4.39, *p* < .01, did.

To summarize, the results of the letter cancellation indicate that the horizontal spatial bias in near and far space did not differ between groups. However, the pattern of horizontal bias in near and far space for each of the groups was similar to what we observed on the shape cancellation task. The performance of patients with neglect in both regions of space was more impaired than that of patients with neglect in one region of space. Despite the lack of differences in horizontal bias, we did find differences in vertical spatial bias in near and far space between groups. Interestingly, the group with neglect in both regions of space showed a more severe bias in the vertical dimension in near space than in far space, whereas the groups with neglect in one region only showed approximately the same amount of vertical bias in near and far space.

Table 6 Average values (standard error) for each condition and each group on the absolute CoC-x and y on the letter cancellation task in near and far space.

	Near abs. CoC-x	Far abs. CoC-x	Near abs. CoC-y	Far abs. CoC-y
N-F-	.005 (.007)	.004 (.010)	.014 (.008)	.010 (.004)
N+F-	.037 (.010)	.014 (.015)	.044 (.012)	.026 (.005)
N-F+	.013 (.015)	.048 (.021)	.050 (.017)	.031 (.008)
N+F+	.127 (.013)	.119 (.019)	.132 (.016)	.059 (.007)

Note. N+F– = near only neglect, N–F+ = far only neglect, N+F+ = near and far neglect. CoC-x and CoC-y = horizontal and vertical normalized center of cancellation, respectively. abs. = absolute. Standard errors in parentheses.

---

[13] Levene's Test for Equality of Variances indicated that the variance between the respective groups was not equal. We therefore used the adjusted *t*-value and degrees of freedom in this t-test, resulting in a less significant *p*-value, which was then also Bonferroni corrected.

Hemisphere of lesion and cancellation content

The frequency of neglect after left- and right-sided lesions was not related to the content of the cancellation task (verbal vs. nonverbal) that was used to measure neglect, $\chi^2(1) = 0.72$, $p = .40$. The percentages of neglect after left- and right-sided brain lesions were about the same across the two cancellation tasks: shape cancellation, left 34% versus right 66%; letter cancellation, left 25% versus right 75%. These differences in occurrence of neglect between left- and right-sided lesions are in line with the literature (see Bowen et al., 1999).

Line bisection

The repeated measures ANOVA revealed a significant main effect of space, $F(1, 78) = 39.23$, $p < .01$, and neglect type, $F(3, 78) = 21.30$, $p < .01$. On average, patients deviated more from the center of the lines in far space (mean absolute deviation = 1.19°, *SE* = 0.12°) than in near space (mean absolute deviation = 0.88°, *SE* = 0.12°). The N+F+ group showed the largest deviation (mean absolute deviation = 2.45°, *SE* = 0.26°), followed by the N– F+ group (mean absolute deviation = 0.74°, *SE* = 0.29°), the N+F– group (mean absolute deviation = 0.69°, *SE* = 0.26°), and the absolute deviation = 0.26°, *SE* = 0.10°). The absolute average deviation in degrees from the center of the line in each region of space and for each group is shown in Table 7. The average absolute deviation from the center in the healthy control group was 0.23° (*SE* = 0.03°) in near space and 0.21 (*SE* = 0.05°) in far space. The performance of the N–F– group did not differ from that of the healthy control group in near, $t(83) = -0.04$, $p = .97$, and in far space, $t(83) = 0.31$, $p = .76$.

Table 7 Average values (standard error) for each condition and each group for the absolute deviation from the center in degrees of visual angle on the line bisection task in near and far space.

	Near abs. deviation from center (degrees)	Far abs. deviation from center (degrees)
N-F-	.247 (.102)	.278 (.105)
N+F-	.926 (.258)	.451 (.264)
N-F+	.225 (.292)	1.250 (.300)
N+F+	2.120 (.258)	2.785 (.264)

Note. N–F– = without neglect, N+F– = near only neglect, N–F+ = far only neglect, N+F+ = near and far neglect. abs. = absolute. Standard errors in parentheses.

More importantly, we found a significant interaction between space and neglect type, $F(3, 78) = 38.79$, $p < .01$. To examine the interaction more closely, several planned independent t-tests for near space and far space were done. In near space, the N+F– group performed significantly worse than the N–F– group, $t(64) = -11.08$, $p < .01$, but not the N–F+ group, *t*(62) = 0.34, *p* = .98, and the N+F+ group, *t*(8.01) = –2.37, *p* = .13. In far space, the performance of the N+F– group did not differ from that of the N–F– group, *t*(64) = –2.35, *p* = .07, whereas the N–F+ and the N+F+ group did, *t*(6.46) = –6.80, *p* < .01, and *t*(8.02) = –3.15, *p* = .04, respectively. The mean line bisection performance of each group in each region of space is shown in Figure 3 and shows again the same interaction pattern as that seen on the shape and letter cancellation tasks (note that the groups do not consist of the same patients; see the next section).

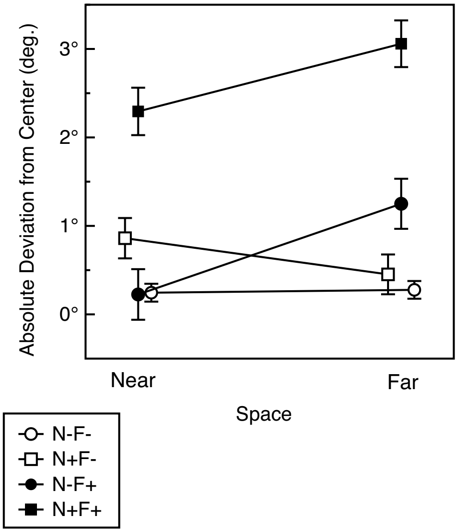


Figure 3. Average absolute deviation from the center of the line in degrees of visual angle for each group for both distances on the line bisection task. Error bars depict standard error of the mean.

Neglect consistency

To investigate whether neglect was consistent across tasks, we examined how many patients consistently showed neglect on each of the tasks. In order to do so, we selected those patients that performed all neglect tasks (shape cancellation, letter cancellation, line bisection) in each region of space (*N* = 65). Of these 65 patients that performed each task of the screening (shape cancellation, letter cancellation, line bisection), 37% did not show neglect on any of the tasks, 32% showed neglect on only one task, 15% of these patients showed neglect on two tasks, and another 15% of these patients showed neglect on three tasks. Of the patients that showed neglect on one task, 90% had neglect in one region of space (N+F– or N–F+). Of the patients that showed neglect on three tasks, 50% had neglect in both regions of space (N+F+), whereas the other 50% had a different type of neglect on at least one of the tasks (e.g., N+F+ on the shape cancellation and line bisection, yet N+F– on the letter cancellation). The results indicate that neglect for one region of space is often found on one specific task, whereas neglect for both regions of space is often found on several tasks. The latter might be due to more severe neglect in the N+F+ group (as shown by the results from each of the tasks), resulting in more consistent performance impairments.

General Discussion

The aim of the current study was to investigate the frequency, spatial specificity (near vs. far space), and severity of neglect in two spatial dimensions in a large sample of stroke patients. We tested patients in two regions of space (near: 30 cm vs. far space: 120 cm) with three frequently used tasks for measuring neglect (a shape cancellation, a letter cancellation, and a line bisection task). We analyzed biases in both the horizontal and the vertical dimension for the cancellation tasks.

With respect to the frequency of neglect, approximately 40% of the patients were diagnosed with neglect, depending on the task that was used to assess neglect. Although there is a large variation in the reported frequency of neglect, our findings are in line with the percentage of neglect frequency as reported by Bowen and colleagues (1999) in their systematic review. The current results also indicate that the presence and the severity of neglect can depend on the distance at which a task is presented as well as on the type of task that is used to test for neglect. Patients could have neglect in only one region of space, or in both regions of space. This has also been observed in previous (small group) studies in which the influence of distance on neglect was investigated (e.g., Aimola et al., 2012; Halligan & Marshall, 1991; Vuilleumier, Valenza, Mayer, Reverdin, & Landis, 1997). Of the patients with neglect, the percentage of patients with neglect at one distance (N+F– or N–F+) varied between 8% and 22%, and the percentage of patients with near and far neglect (N+F+) varied between 11% and 14%, depending on the task that was used to assess neglect. A rather large part of the patients showed far neglect (25% N–F+ and 47% N+F+ based on the shape cancellation task). Far space neglect is often not assessed in clinical practice, and our results indicate that a large proportion of patients with neglect may have an attentional impairment for stimuli presented in far space. Even more important, about one fourth of the patients with neglect were impaired in far space only (N–F+), which we would not be able to identify when these patients were tested in near space only with paper-and-pencil tasks.

Interestingly, the patients with neglect showed not only a horizontal spatial bias (an imbalance in the number of missed targets between the left and right part of the visual field in the cancellation tasks), but also a vertical spatial bias (an imbalance in the number of missed targets between the upper and lower part of the task in the cancellation tasks). Vertical neglect has been found before (Butter, Evans, Kirsch, & Kewman, 1989; Halligan & Marshall, 1989; Pitzalis, Spinello, & Zoccolotti, 1997), but our results indicate that horizontal and vertical attentional biases can co-occur.

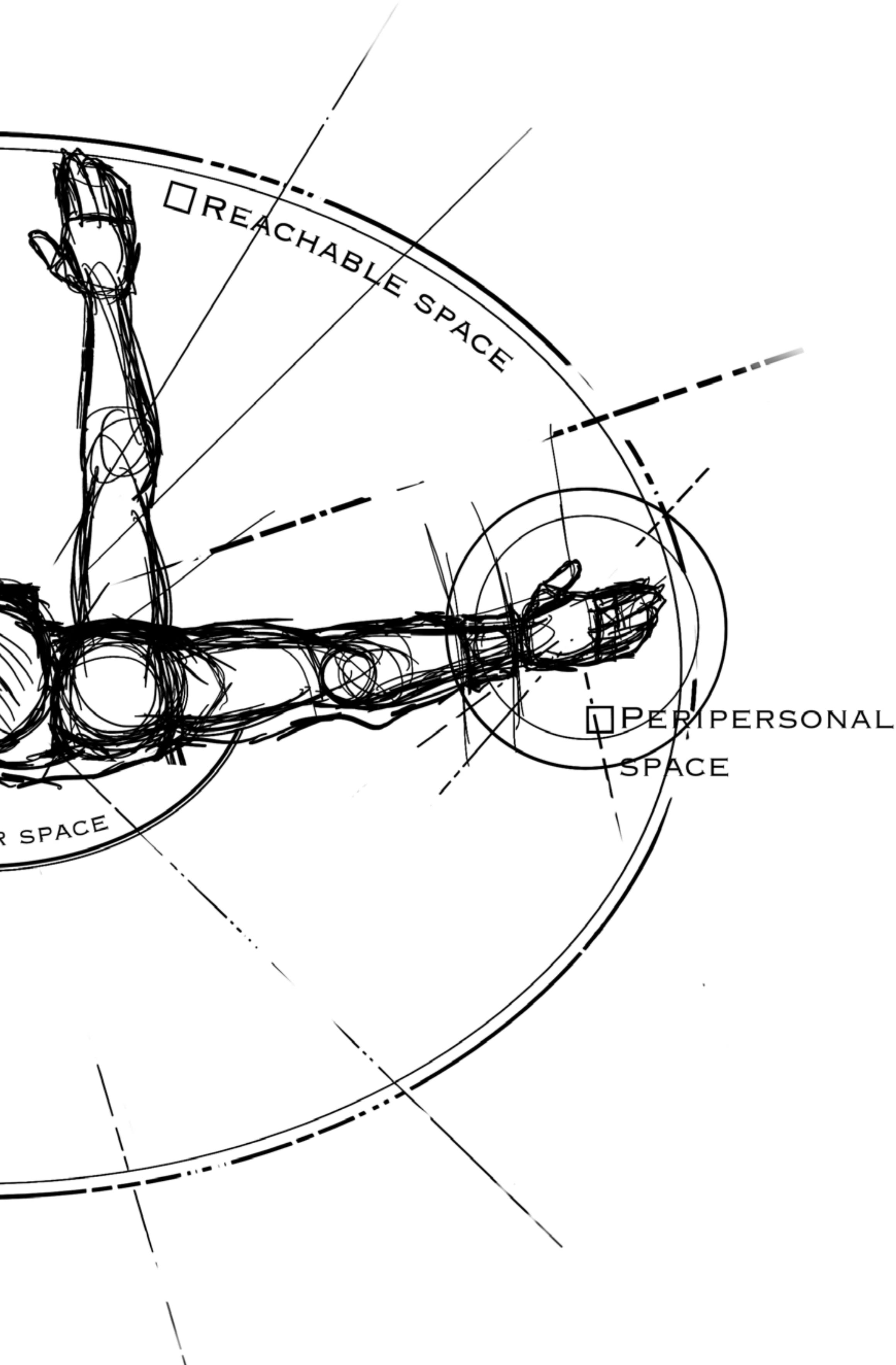
Concerning the severity of neglect, in general, patients with neglect at both distances (N+F+), showed more severe neglect than patients that showed neglect at one distance (N+F– or N–F+). The severity of neglect was approximately the same for the near and the far neglect groups. A correlation between the size of a brain lesion and the severity of neglect has been reported previously (Leibovitch et al., 1998). Leibovitch and colleagues suggested that this relation might be the result of damage to more areas that are involved in attentional processing. The dorsal and ventral streams have been associated with the processing of information coming from near and far space (Previc, 1998) and with action and perception (Milner & Goodale, 2008), respectively. On a speculative basis, we suggest that patients with neglect in both near and far space may have larger lesions that include both parietal (dorsal) and temporal (ventral) cortical regions that are involved in near and far space attentional processing (Previc, 1998). On the other hand, patients with neglect confined to one region of space may have smaller lesions that include only one of these areas (e.g., Aimola et al., 2012; Butler et al., 2004). Different neglect tasks may depend more heavily on either action or perception, and therefore more on one of the two streams. A brain lesion might affect either one or both of the streams, which may result in different performance impairments across tasks.

In line with this reasoning, patients in one of the neglect groups based on one task were often not found in the same neglect group on another task. This was especially true for the N+F– and the N–F+ group. There was a higher consistency of neglect type across tasks when patients had both near and far neglect, which might be because their performance was worse than that of the single space neglect groups. These findings are in agreement with a previous study in which it was shown that the correlation between performance on cancellation and line bisection tasks is generally low (Schubert & Spatt, 2001). This might be the result of different operations that are required to perform the task. Cancellation tasks might depend more on visual exploration and require a dominant motor response (hence the term cancellation), whereas line bisection tasks may depend more on a perceptual estimation with a less dominant motor component. This difference might explain why some patients showed neglect on one task, but not on another task, suggesting that their neglect is only present when certain operations are required. This was also reported by Keller and colleagues (2005), who found no influence of distance on severity of neglect on a cancellation task, whereas they did find an effect of distance on performance on a line bisection task in the same patients. We found that the severity of neglect depended on distance in both cancellation and line bisection tasks, but most often not within the same patient.

Although we did not test for the presence of hemianopia, we do not expect that patients with visual field deficits influenced the current results. We know that hemianopia (one of the common visual field defects, Suchoff et al., 2008) can cause a contralateral line bisection bias, and that performance on cancellation tasks is not lateralized in patients with hemianopia (Barton & Black, 1998; Doricchi, Onida, & Guariglia, 2002). Still, we would not expect a difference in the bisection bias between near and far space on a line bisection task, because we corrected the lines that needed to be bisected for visual angle. More importantly, the amount of deviation from the center in patients with hemianopia may be larger than that in healthy subjects but is smaller than that in patients with neglect (Barton & Black, 1998). In this study, we used the mean bisection bias ± 3 standard deviations of a healthy control group as a cut-off for neglect. In other studies, the performance of hemianopia patients fell within a 3-standard-deviation range from the mean of their control group (Barton & Black, 1998; Doricchi et al., 2002). We therefore assume that the performance on the line bisection of any patient with hemianopia in our sample was below the cut-off and was therefore not included in any of the neglect groups.

On the line bisection task, we observed that there were more patients with signs of right-sided than left-sided neglect in the near neglect group. This was mainly due to three right-hemisphere patients with ipsilesional deviations on this task (all other patients showed contralesional deviations). One possibility is that their performance reflects ipsilesional neglect, which is more commonly found on line bisection tasks than on cancellation tasks after frontal subcortical lesions (e.g., Kim et al., 1999). Another possible explanation for their pattern of performance may be the use of compensatory scan strategies. During their stay in the rehabilitation center, patients with signs of neglect are made more aware of their attentional bias and are trained to start scanning on their neglected side of space in order to learn to compensate for their neglect in everyday life. In several cases we observed that patients actually started to miss information that was presented in their non-neglected side of space after several training sessions. Although we have not systematically investigated the influence of this training on search behavior, the current results may in part be influenced by the effects of overcompensation due to this training in the rehabilitation center. Patients are trained in near space, and it is currently unknown whether the effects of the scan training in near space are transferred to far space. Strikingly, one of the patients in our sample showed an inconsistent neglect pattern on the letter cancellation task: right-sided neglect in near space, and left-sided neglect in far space. Although highly speculative, the deviating pattern of neglect in this patient would be in line with the fact that search strategies are not transferred from near to far space, as this would explain the presence of right-sided neglect in near space due to overcompensation and the presence of left-sided neglect in far space.

In sum, the current study adds to the growing body of research on distance-specific neglect by showing that both horizontal and vertical neglect can be distance specific in a large group of stroke patients and that the type of neglect depends on the type of task that is used to assess this disorder. Far neglect was present in a relatively large part of the patients with neglect, and this cannot be observed with the standard paper-and-pencil neglect tests that are widely used in clinical practice. Overall, our results suggest that it is important to take the three dimensions of space into account when testing for the presence of neglect.


☐ Reachable space
☐ Peripersonal space
r space

Chapter 9

Audiovisual integration in near and far space: Effects of changes in distance and stimulus effectiveness

Van der Stoep, N., Van der Stigchel, S., Nijboer, T. C. W., & Van der Smagt, M. J. (in press).
Audiovisual integration in near and far space:
Effects of changes in distance and stimulus effectiveness.
Experimental Brain Research.

Abstract

A factor that is often not considered in multisensory research is the distance from which information is presented. Interestingly, various studies have shown that the distance at which information is presented can modulate the strength of multisensory interactions. In addition, our everyday multisensory experience in near and far space is rather asymmetrical in terms of retinal image size and stimulus intensity. This asymmetry is the result of the relation between the stimulus-observer distance and its retinal image size and intensity: an object that is further away is generally smaller on the retina as compared to the same object when it is presented nearer. Similarly, auditory intensity decreases as the distance from the observer increases. We investigated how each of these factors alone, and their combination, affected audiovisual integration. Unimodal and bimodal stimuli were presented in near and far space, with and without controlling for distance-dependent changes in retinal image size and intensity. Audiovisual integration was enhanced for stimuli that were presented in far space as compared to near space, but only when the stimuli were not corrected for visual angle and intensity. The same decrease in intensity and retinal size in near space did not enhance audiovisual integration, indicating that these results cannot be explained by changes in stimulus efficacy or an increase in distance alone, but rather by an interaction between these factors. The results are discussed in the context of multisensory experience and spatial uncertainty, and underline the importance of studying multisensory integration in the depth space.

Introduction

Integration of information from different senses has been extensively studied in the last few decades (see Stein et al. 2010, Figure 6; Murray, Spence, & Harris, 2013; Van der Stoep, Nijboer, Van der Stigchel, & Spence, 2015, Figure 1). Several rules, or principles, have emerged from neurophysiological studies that are known to be important for multisensory integration (MSI) to occur (or to enhance it, e.g., King & Palmer, 1985; Stein & Meredith, 1993). Typically, the effects of multisensory stimulation are more pronounced (e.g., shorter response times, RT) when stimuli from different senses are *spatially* and *temporally aligned* (Holmes & Spence, 2004; though see Spence, 2013; see Stein & Stanford, 2008, for a review). In addition, the principle of *inverse effectiveness* states that the benefit of multisensory interactions is largest when responses to unisensory stimuli are weak (e.g., Meredith & Stein, 1983; Holmes, 2007). In neurophysiological studies, relative increases in multisensory responses with decreasing stimulus intensity have been observed in multisensory neurons (in terms of relative increase in spike rates in single- or multi-unit recordings; e.g., Alvarado, Vaughan, Stanford, & Stein, 2007).

In humans, however, results from behavioral studies have shown conflicting findings, mainly with respect to the spatial rule (Spence, 2013) and the principle of inverse effectiveness (Holmes, 2007, 2009). Concerning the principle of inverse effectiveness, several behavioral studies have reported inconsistent results regarding the relation between stimulus intensity (or signal-to-noise ratio) and either the amount of multisensory response enhancement (i.e., faster or more accurate responses to multisensory stimuli in comparison with responses to unimodal stimuli) or MSI (i.e., enhancement beyond what would be expected based on an independent channel model, Miller, 1982, 1986; e.g., Leone & McCourt 2013; Stevenson, Fister, Barnett, Nidiffer, & Wallace, 2012; Ross, Saint-Amour, Leavitt, Javitt, & Foxe, 2007). For example, Leone and McCourt (2013) observed larger benefits of MSI (i.e., shorter RTs) when multisensory stimuli were composed of unimodal stimuli of *higher* intensity as compared to when stimuli were of lower intensity. These studies demonstrate that it is difficult to consistently replicate some of the neurophysiological observations regarding the principle of inverse effectiveness in behavioral studies in humans.

Although the majority of behavioral studies in humans have looked into the principles governing MSI at a fixed distance from the observer, there are several reasons to believe that the integration of information from certain sensory modalities is more or less effective depending on the region of space from which multisensory stimuli are presented (for review, see Van der Stoep, Nijboer, Van der Stigchel, & Spence, 2015). For example, multisensory interactions involving touch are more pronounced in near (or peripersonal space, the space directly surrounding different body parts; e.g., Rizzolatti, Scandolara, Matelli, & Gentilucci, 1981; Occelli, Spence, & Zampini, 2011) as compared to far space (the space further away from the body that is out of reach). In contrast, audition and vision seem to be the dominant senses in far space (or action-extrapersonal space; Previc, 1998). This difference in distance-based sensory modality dominance may not be surprising when thinking of the differing behavioral functions that are related to near and far space. Grasping and manipulating objects require direct contact between the body and the environment and are typical behavioral functions bound to near space. One of the dominant functions in far space is spatial orienting, a much more audiovisual-based function which does not necessitate contact between the environment and the body in the same way as, for example, grasping. Furthermore, different neuroanatomical systems seem to be involved in the processing of near (e.g., Graziano & Gross, 1994; Graziano, Reiss, & Gross, 1999) and far space (e.g., Aimola et al., 2013; Committeri et al., 2007; for review see Previc, 1998). Brain regions coding near space seem to be more related to audiotactile/visuotactile and motor processing (e.g., Serino, Canzoneri, & Avenanti, 2011), whereas brain regions coding far space seem to be more related to spatial orienting and audiovisual integration (Previc, 1998).

In addition to possible distance-based sensory modality dominance, a change in the distance from which stimuli are presented also changes the arrival time of auditory and visual stimuli. For instance, increasing the distance between audiovisual stimuli and the observer increases the difference in arrival times of auditory and visual stimuli that are caused by the difference in the speed of light and sound. It has been shown that the difference in the speed of light and sound is taken into account when judging the simultaneity of audiovisual sources in depth (e.g., Alais & Carlile 2005; Arnold, Johnston, & Nishida, 2005; Sugita & Suzuki, 2003). This effect, however, seems to depend on whether estimating external temporal alignment is task relevant (see also Heron, Whitaker, McGraw, & Horoshenkov, 2007). These studies indicate that fairly accurate estimates about the distance of multisensory stimuli can be made, possibly as a result of prior multisensory experience with the environment.

Finally, increasing the distance between a multisensory stimulus and an observer also decreases the retinal image size and intensity of visual stimuli (e.g., through the inverse-square rule, Warren, 1963) and the intensity and direct-to-reverberant ratio of auditory stimuli (Alais & Carlile, 2005; Bronkhorst & Houtgast, 1999). Because of this relation between distance and stimulus properties and audiovisual dominance in far space, audiovisual integration may be specifically enhanced when an increase in distance and a decrease in retinal image size and stimulus intensity go hand in hand.

In order to investigate the possible interplay between distance (region of space), retinal image size, and stimulus intensity, audiovisual stimuli of two different stimulus intensities and sizes were presented both in near and far space. A near space condition was used with an audiovisual stimulus of a certain size and intensity (*Near High*) and a far space condition in which the same stimulus was presented at a larger distance from the observer (*Far Low*). To be able to disentangle the influence of distance, retinal image size and intensity, and their interaction on MSI, a condition was added in which the audiovisual stimulus in near space had the same decrease in retinal image size and intensity (*Near Low*) as the stimulus in the *Far Low* condition. To balance the design, a final condition was constructed (*Far High*) which consisted of the same near space stimulus (*Near High*) at a larger distance, but corrected for retinal image size and intensity (*Far High*).

Thus, by comparing these four conditions, the effects of changes in distance, in stimulus efficacy, and distance-related changes in stimulus efficacy on audiovisual integration could be investigated. We hypothesized that audiovisual integration may be more pronounced in far space, given the potential dominance of audiovisual information in far space, and the lawful relation between changes in distance and changes in stimulus efficacy.

Methods

Participants

Twenty participants were tested in the experiment (11 male, mean age = 24.58 years, *SD* = 3.55). All participants reported to have normal or corrected-to-normal visual acuity and normal hearing. The experiment was conducted in accordance with the Declaration of Helsinki. All participants signed an informed consent form before taking part in the study and received either monetary compensation or course credits for their participation.

Apparatus and stimuli

To project visual stimuli in near and far space, we used an Acer X1261P projector (60 Hz) that was placed behind the observer projecting down in a small angle. The projection screen in near space was located at a distance of ~80 cm from the observer and at ~208 cm in far space. These distances were chosen as they were considered to be within the reachable (near/peripersonal space) and unreachable space (far/extrapersonal space) that were discussed in the "Introduction" section. In previous studies, it has been shown that perceptual processing can differ depending on the region of space (peripersonal vs. extrapersonal) in which the stimuli are presented (e.g., Aimola, Schindler, Simone, & Venneri, 2012; Halligan & Marshall, 1991; Occelli, Spence, & Zampini, 2011; Previc, 1998; Van der Stoep et al., 2013; Van der Stoep, Nijboer, Van der Stigchel, & Spence, 2015). When stimuli were presented in far space, the projection screen in near space was placed out of sight. Six speakers (Harman/Kardon HK206, Frequency response: 90–20,000 Hz) were used to present auditory stimuli at three locations (left, center, right) in both near and far space (resulting in six locations). An overview of all conditions is shown in Figure 1.

The auditory stimuli consisted of 100-ms white noise bursts (with a 15-ms rise and fall of the signal). In the *Near High* and the *Far High* condition, auditory stimuli were presented at ~70 dB(A) SPL as measured from the location of the participant. In the *Near Low* and *Far Low* condition, auditory stimuli were presented at ~60 dB(A) SPL as measured from the location of the participant. This difference in intensity was chosen as the increase in distance between the speaker and the observer resulted in a decrease of approximately 10 dB(A) SPL. The reverberant room was 2.60 m wide, 5.10 m long, and 2.80 m high. The difference in the direct sound and the sound reflections between near and far space provided a cue for the distance of the sound. When the auditory stimulus was corrected for intensity, the differing reflections between near and far space provided information about the distance of the stimulus [see supplementary Figure 1 for the impulse response of sound presented at the central location (see Figure 2) in near (80 cm; S1 top) and far space (208 cm; S1 bottom)]. The visual stimuli consisted of a gray-filled circle (*Near High* and *Far High*: 3° in diameter, ~6.6 cd/m^2, *Near Low* and *Far Low*: 1.16° in diameter, ~1.98 cd/m^2; intensities were measured from the location of the participant using a PhotoResearch SpectraScan PR 650 spectrometer) that was presented on a dark gray background (near and far space: ~1.3 cd/m^2). The fixation cross was a gray plus sign (near and far space: 0.7° × 0.7°, ~6.6 cd/m^2). Visual target locations were 13.8° to the left and right of fixation, and in the center of the screen at the location of the fixation cross. The speakers were placed directly behind the locations of the projected visual stimuli in both near and far space to ensure perfect spatial alignment. Audiovisual stimuli were always presented spatially and temporally aligned. A schematic bird's-eye view of the experimental setup is shown in Figure 2.

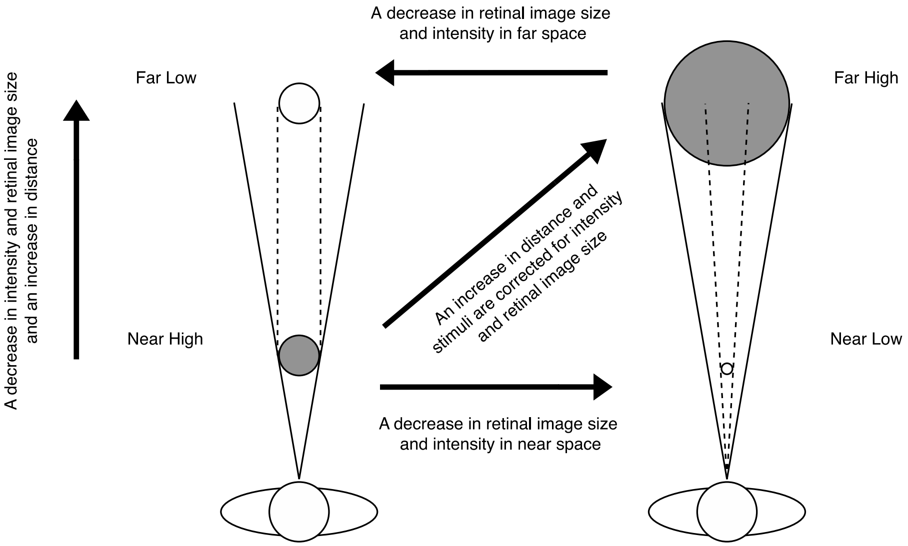


Figure 1. A schematic overview of the four conditions that were used in the experiment. Note that the retinal image size and the brightness of the visual stimulus, as well as the sound pressure level of the auditory stimulus at the observer's location, were the same in the *Near High* intensity and *Far High* intensity conditions and in the *Near Low* intensity and the *Far Low* intensity conditions

The experiment consisted of two blocks: 360 trials in near space and 360 trials in far space. The distance from which stimuli were presented was blocked and counterbalanced across participants. Stimuli originating from the same distance were all presented in a single block. There were two intensity conditions for each distance resulting in a *Near High*, a *Near Low*, a *Far High*, and a *Far Low* intensity condition. Within each distance block, two breaks were introduced: one after 120 trials and another one after 240 trials. Participants were able to continue the experiment by pressing the space bar. Both near and far blocks contained 180 high-intensity and 180 low-intensity trials. Each of those 180 trials consisted of 60 auditory, 60 visual, and 60 audiovisual targets, and each of those 60 trials contained 20 trials in which the target was presented to the left of the fixation cross, 20 trials to the right of the fixation cross, and 20 trials presented at the location of the fixation cross.

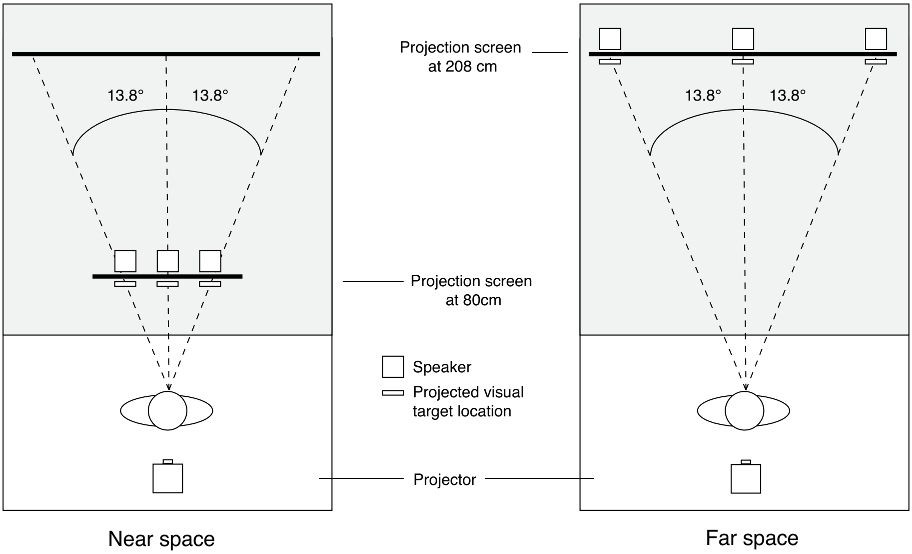


Figure 2. A schematic bird's-eye view of speaker and projected target locations in near space (left panel) and far space (right panel) setup.

Procedure

The experiment was conducted in a room where the light from the projector was the only source of illumination. Participants were asked to place their chin in a chinrest. Before the start of the experiment, participants had to report the location of three auditory stimuli (white noise) that were presented at three different locations in near space (left, center, and right) as a quick test to confirm their ability to localize these sounds. All participants were able to correctly report the locations of the auditory stimuli.

Each trial started with the presentation of the fixation cross. After a random duration between 750 and 1250 ms, the fixation-cross disappeared, and a blank screen was presented. After another random duration between 200 and 250 ms, either an auditory, visual, or audiovisual target was presented for 100 ms after which the target disappeared. The target modality, location, and intensity were randomized across trials. Participants were instructed to press a button on a response box as soon as they detected either a visual, auditory, or audiovisual stimulus to the left or the right side of the fixation cross (Go trials), but to withhold their response when a stimulus was presented in the center (No-go trials). Given that participants only had to respond to lateral targets, but not to central targets in the experiment, their ability to correctly localize the auditory stimuli was reflected by the amount of errors made in responding to auditory targets (i.e., whether they were able to differentiate between auditory stimuli presented to the lateral locations and the central location). The response window was 2,000 ms from target onset.

Data analysis

Go trials with response times between 100 and 1000 ms and No-go trials without a response were considered correct. Only response times on Go trials between 100 and 1000 ms were analyzed, because responses faster than 100 ms were considered to be the result of anticipation and responses slower than 1000 ms the result of not paying attention to the task. This led to the removal of 2.60% of the data in the near space conditions and 2.53% of the data in far space. In the near space condition, 0.77% of the Go trials (High = 0.37%, Low = 1.17%) and 6.25% of the No-go trials (High = 7.25%, Low = 5.25%) were removed, and in far space, 0.65% of the Go trials (High = 0.50%, Low = 0.79%) and 6.29% of the No-go trials (High = 6.92%, Low = 5.67%) were removed. The median response times of each participant in each condition were used in the RT analysis as RT distributions are generally skewed and the median is less affected by the presence of outliers.

A 2 × 3 × 2 repeated measures ANOVA with the factors Target Space (Near, Far), Target Modality (A, V, AV), and Intensity (High, Low) was used to analyze RTs. The Greenhouse–Geisser correction was used whenever the assumption of sphericity was violated. To test for differences in RT between conditions, we used two-tailed paired samples *t-tests* (*p* values were corrected using the Bonferroni method where indicated: corrected *p* value = *p* × number of tests).

To investigate the increase in the speed of detection in the audiovisual condition compared to the fastest unimodal condition, we calculated the amount of absolute multisensory response enhancement (aMRE) and the amount of relative multisensory response enhancement (rMRE) for each condition (*Near High*, *Near Low*, *Far High*, and *Far Low*):

$$\text{absolute MRE} = \min(\text{median}(RT_A), \text{median}(RT_V)) - \text{median}(RT_{AV})$$

$$\text{relative MRE} = [\min(\text{median}(RT_A), \text{median}(RT_V)) - \text{median}(RT_{AV}))]/[\min(\text{median}(RTA), \text{median}(RT_V))] \times 100\%$$

Note that the median RT of each condition was used to calculate the amount of aMRE and rMRE for each participant. The amount of rMRE was also calculated because of possible unimodal baseline differences in the different intensity conditions (low and high) and the different space conditions (near and far). Two-tailed planned pairwise comparisons were used to test for the differences between the *Near High*, *Far High*, *Near Low*, and *Far Low* conditions.

To test whether the observed response enhancement for audiovisual targets could be explained by statistical facilitation, the cumulative distributive function (CDF) of RTs to auditory, visual, and audiovisual targets was calculated for each condition (*Near High*, *Near Low*, *Far High*, and *Far Low*). Using these CDFs, the upper bound of statistical facilitation predicted by a race model was calculated using the race model inequality (Raab 1962; Miller 1982, 1986; Ulrich, Miller, & Schröter, 2007):

$$P(RT_{AV} < t) \leq P(RT_{A} < t) + P(RT_{V} < t)$$

The race model inequality represents the probability (*P*) of a given RT in the audiovisual condition that is less than or equal to a given time *t* in milliseconds based on the combined probabilities for a given RT in the unimodal conditions where *t* ranges from 100 to 1000 ms (assuming a maximum negative correlation of –1 between detection times of unimodal stimuli).

The performance in the audiovisual condition can be compared to the upper bound predicted by the race model inequality in two ways. First, RTs for a range of quantiles (i.e., 10, 20, up to 90%) of the CDF can be compared between the audiovisual CDF and the race model CDF. Second, probabilities between the audiovisual CDF and the race model CDF can be compared for a range of RTs. Shorter RTs or larger probabilities in the audiovisual condition compared to the race model indicate MSI. Comparing RTs at a range of quantiles reveals how much shorter RTs in the audiovisual condition are in an absolute sense (i.e., difference in ms) at comparable points of the audiovisual and race model CDF. In addition, the difference in probability indicates in which RT range differences between the AV CDF and the race model CDF occur. Given that both methods provide different information and have (in isolation) been used in previous studies (e.g., Girard, Pelland, Lepore, & Collignon, 2013; Stevenson, Krueger Fister, Barnett, Nidiffer, & Wallace, 2012; Van der Stoep, Van der Stigchel, & Nijboer, 2015), both methods were used to analyze the data in the current study.

Differences in RT between the audiovisual CDF and the race model CDF for each quantile (i.e., race model inequality violation) were analyzed using one-tailed pairwise comparisons for each quantile and each condition (*p* values were corrected for nine tests in each condition using the Bonferroni method). Differences in probability between the audiovisual CDF and the race model CDF for each RT point were analyzed using paired samples t-tests.

Four additional measures were extracted from the race model inequality violation curves in each condition of each participant: (1) the mean race model inequality violation value across nine quantiles, (2) an estimate of the area under the race model inequality violation curve that was based on differences in RT for each quantile, (3) the maximum race model inequality violation in terms of probability, and (4) the corresponding time point of this maximum race model inequality violation. These values were extracted for each condition and each participant and compared between conditions using planned pairwise comparisons.

Results

Accuracy

Overall accuracy was very high (overall $M = .965$, SE = .005, Go trials: $M = .993$, SE = .002, No-go trials: $M = .937$, SE = .010) indicating that participants were well able to localize and detect targets, and withhold their response when targets appeared at the central location.

Response times

A significant main effect of Target Modality [$F(2, 38) = 47.048$, $p < .001$, partial $\eta^2 = .712$] was observed. The RTs to audiovisual targets ($M = 337$ ms, $SE = 16$) were significantly shorter compared to visual ($M = 380$ ms, $SE = 16$, $t(19) = 8.655$, $p < .001$, $d = .6$) and auditory targets ($M = 395$ ms, $SE = 19$, $t(19) = 9.337$, $p < .001$, $d = .591$). The difference between RTs to visual and auditory targets was not significant [$t(19) = .136$, $p = .138$]. The average median RTs for each condition are shown in Figure 3.

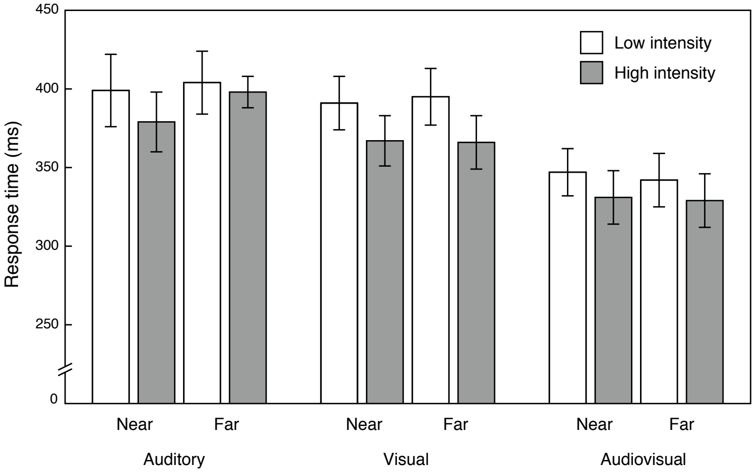


Figure 3. The average of median response times to auditory, visual, and audiovisual targets in near and far space for the low (white) and high (grey) intensity condition. Error bars represent *SE* of the mean.

There was also a main effect of Intensity [$F(1, 19) = 61.304, p < .001$, partial $\eta^2 = .763$]. RTs to high-intensity stimuli ($M = 362$ ms, $SE = 17$) were significantly shorter compared to RTs to low-intensity stimuli ($M = 380$ ms, $SE = 17$). No main effect of Target Space was apparent [$F(1, 19) = .098, p = .758$, partial $\eta^2 = .005$].

The interaction between Target Modality and Target Space was also significant [$F(2, 38) = 4.400, p = .019$, partial $\eta^2 = .188$]. This effect appeared to be driven by a larger difference in RTs between auditory and visual targets in far as compared to near space (mean difference near space = 10 ms, $SE = 8$, mean difference far space = 21 ms, $SE = 7$). The difference in RTs to auditory and visual targets was significant in far space [$t(19) = -2.792, p = .036, d = -.232$], but not in near space [$t(19) = 1.282, p = .645$]. One could argue that the difference between auditory and visual RTs is simply the result of stimulus characteristics. After all, at a distance of 208 cm, the auditory stimulus reaches the observer approximately 6.06 ms later than the visual stimulus. Indeed, if this delay in arrival time is subtracted from the RTs to auditory stimuli in far space, the RT difference between auditory and visual targets in far space ceases to be significant [$t(19) = -1.924, p = 0.192$].

The interaction between Target Modality and Intensity was also significant [$F(1.491, 28.323) = 5.280, p = .018, \varepsilon = .745$, partial $\eta^2 = .217$]. The difference in RTs between auditory and visual targets was significant for high-intensity stimuli (*M* Auditory High

= 388 ms, SE = 19 vs. *M* Visual High = 366 ms, *SE* = 16, $t(19) = 3.537$, $p = .004$, $d = .250$], but not for low-intensity stimuli (*M* Auditory Low = 402 ms, SE = 20 vs. *M* Visual Low = 393 ms, SE = 16, $t(19) = 0.972$, $p = .686$). The comparisons between RTs to auditory and audiovisual, and visual and audiovisual targets were significant for both the high and low intensity conditions (all *t*'s > 7, all *p*'s < .001, also see the "Multisensory response enhancement" section below). Furthermore, responses to high-intensity targets were faster as compared to low-intensity targets, as could be expected based on the intensity manipulation. The difference in RT between the high and low intensity conditions was significant for auditory (*M* difference = 13 ms, SE = 5), visual (*M* difference = 27 ms, SE = 3), and audiovisual targets (*M* difference = 14 ms, SE = 2); all *t*'s < −2.7, all *p*'s < .05). The difference in RT between the high and low intensity conditions was slightly larger for visual targets as compared to auditory and audiovisual targets.

The interaction between Target Space and Intensity [$F(1, 19) = .705$, $p = .412$, partial $\eta^2 = .036$], and the three-way interaction between Target Modality, Target Space, and Intensity were not significant [$F(2, 38) = 1.053$, $p = .359$, partial $\eta^2 = .052$].

In sum, we observed a general increase in the speed of responses in the audiovisual condition compared to the unimodal conditions, and high-intensity stimuli evoked faster responses compared to low-intensity stimuli. Overall, responses to auditory stimuli were slightly slower when presented in far space compared to near space, but this effect could be explained in terms of differences in arrival times. To investigate whether our correction for stimulus intensity and size across distance for both auditory and visual stimuli resulted in similar response times between the near and far condition, we compared RTs in the *Near High* and *Far High*, and the *Near Low* and *Far Low* conditions for auditory and visual targets. If the distance from which stimuli were presented had no particular influence on RTs per se, then no difference between these conditions should be observed. Indeed, we did not find any significant differences between the Near and Far conditions for neither low nor high stimulus intensities for auditory and visual targets (all *t*'s < 1.7, all *p*'s > .1).

Multisensory response enhancement

One-sample t-tests confirmed that there was a significant amount of absolute and relative MRE in each condition (all *t*'s > 4.9, *p*'s < .001, in the *Near Low*, *Near High*, *Far Low*, *Far High* Intensity condition), indicating that responses to multisensory targets were significantly faster when compared to the fastest response to either of the unimodal targets. The average aMRE and rMRE in each condition are shown in the left and right panel of Figure 4, respectively.

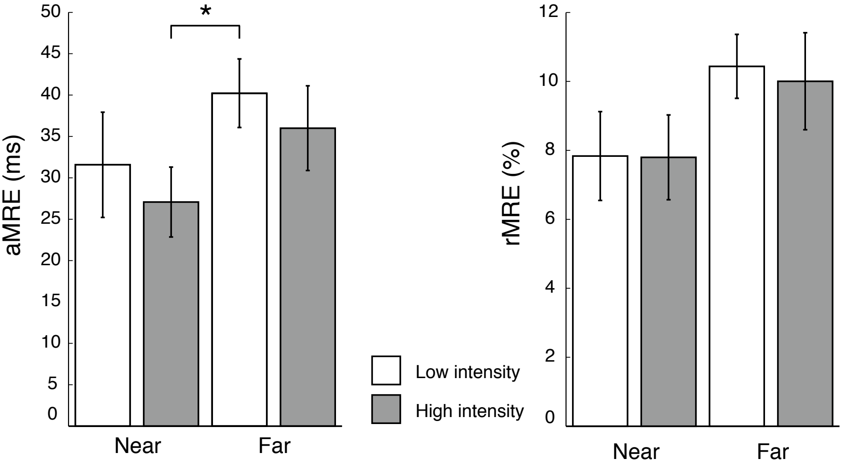


Figure 4. Left panel: The average amount of aMRE in each condition. Right panel: The average amount of rMRE in each condition. Error bars indicate standard error of the mean. Significant differences are indicated with an asterisk.

When the audiovisual stimulus was moved from near to far space, while keeping the stimulus properties the same (i.e., decreased intensity and retinal image size in the far space condition as measured from the location of the participant: *Near High* vs. *Far Low*), the average amount of aMRE was significantly larger in far space (*M Far Low* = 40 ms, SE = 4) as compared to near space (*M Near High* = 27 ms, SE = 4, $t(19) = -2.209$, $p = .040$, $d = -.702$). To investigate whether this increase in aMRE was the result of a decrease in retinal image size and stimulus intensity alone or whether the distance or the region of space from which the stimuli were presented also contributed to the effect, the amount of aMRE in the High intensity and the Low intensity conditions in near space was also compared. Interestingly, there was no difference in the amount of aMRE between the *Near High* and the *Near Low* intensity condition (*M Near High* = 27 ms, SE = 4 vs. *M Near Low* = 32 ms, SE = 6, $t(19) = .742$, $p = .467$). These results indicate that a decrease in stimulus intensity and retinal image size alone did not increase the amount of aMRE.

One could also argue that the difference in the distance from which the stimuli were presented is the only factor that contributes to the increase in aMRE in far space. Therefore, the amount of aMRE was also compared between the *Near High* and *Far High* condition in which the far space stimuli were corrected for intensity and retinal image size (*M Near High* = 27 ms, SE = 4 vs. *M Far High* = 36 ms, SE = 5). A significant difference between the Far Low and the Near Low condition would indicate that a change in distance alone could explain an increase in aMRE, but this was not the case [$t(19) = -1.518$, $p = .145$]. To check whether a decrease in intensity resulted in a generally larger amount of aMRE in far space, the *Far High* and *Far Low* conditions were compared, but the difference was also not significant [$t(19) = .757$, $p = .458$].

Thus, the decrease in stimulus intensity and retinal image size alone could not explain the observed pattern of aMRE, a result that does not seem to be in line with the principle of inverse effectiveness, but the use of only two intensities make it difficult to draw firm conclusions about this. Interestingly, the distance from which the stimuli were presented could also not explain these results, which suggests that aMRE is especially increased when an increase in the distance from which stimuli are presented and a decrease in retinal image size and stimulus intensity co-occured.

The rMRE was analyzed in the same way as the aMRE was. A similar, but less pronounced, effect as with the aMRE measure was observed for the rMRE measure (Figure 4 right panel). The difference between the *Near High* ($M = 7.8\%$, SE = 1.2) and *Far Low* condition ($M = 10.4\%$, SE = 0.9) was not significant [but in the right direction, $t(19) = -1.852$, $p = .080$, $d = -.402$]. An effect size of −.548 could be considered medium to large, which may indicate a lack of power. As in the aMRE measure, the other comparisons were not significant: *Near High* versus *Near Low* [$M = 7.8\%$, SE = 1.3, $t(19) = -.030$, $p = .976$], *Near High* versus *Far High* [$M = 10.0\%$, SE = 1.4, $t(19) = -1.514$, $p = .147$], and *Far High* versus *Far Low* [$t(19) = -.289$, $p = .776$]. The results of the rMRE analysis showed a similar pattern as the results of the aMRE analysis, but the critical difference, namely the difference between the *Near High* and the *Far Low* condition, failed to reach significance ($p = .080$). Both the absolute and relative MRE measures were reported here because it has been shown that a pattern of inverse effectiveness can depend on whether an absolute or a relative measure of multisensory response enhancement is analyzed (see for example Holmes 2007, for a discussion). Although it is unclear as to why the rMRE measure only showed a trend regarding the distance-based inverse effectiveness effect in the current study ($p = .080$), both the aMRE and the rMRE measure showed a similar pattern.

Race model inequality violation

To investigate whether multisensory response enhancement could be explained in terms of statistical facilitation alone (i.e., the race model) or better by MSI, the audiovisual CDF was compared to the race model CDF in each condition using paired samples *t-tests* at nine points of the CDF. We compared both differences in RT across a range of quantiles (see Figure 5) and differences in probability across a range of RTs based on the full CDF function (see Figure 7).

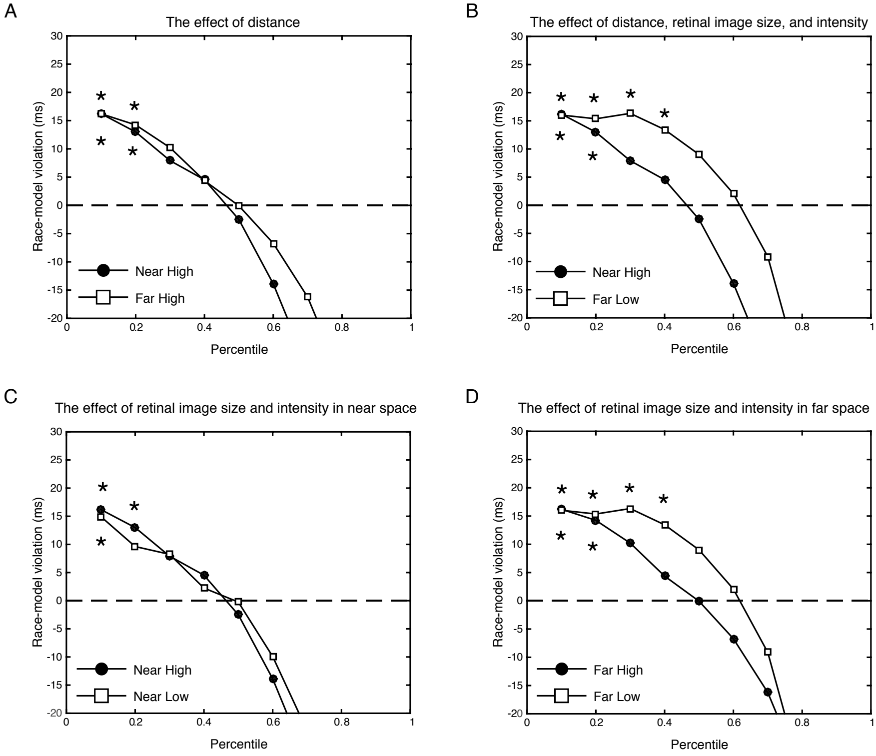


Figure 5. The difference curves show the amount of race race model inequality violation in each condition in terms of response time differences across the range of percentiles. A: The effect of distance on race model inequality violation, B: The effect of distance combined with a decrease in intensity and retinal image size, C: The effect of a decrease in intensity and retinal image size in near space, and D: The effect of a decrease in intensity and retinal image size in far space. Asterisks indicate *p*-values that were smaller than .05 after correction for multiple tests using the Bonferroni method.

The comparison between differences in RT at corresponding quantiles revealed significant violations at the 10th percentile in the *Near Low* condition. In the *Near High* condition, we observed significant violations at the 10th and the 20th percentile. Interestingly, in the Far Space Low Intensity condition (i.e., the same stimulus as in the *Near High* condition, but presented at a larger distance), we observed race model inequality violations over a broader range of percentiles, from the 10th to the 40th percentile. In the Far Space High Intensity condition, significant violations of the race model inequality were found at the 10th and the 20th percentile (for all four conditions: *t*'s > 2.9, *p*'s < .05, one-tailed, corrected using the Bonferroni method).

The average amount of race model inequality violation and the average net area under the curve for each condition are shown in the left and right panel of Figure 6, respectively. The race model CDF is very steep compared to a normal CDF as its probability sums up to 2 (and is cut of at 1). Therefore, race model inequality violations are almost always negative at the higher percentiles of the RT distribution, which leads the average amount of violation to be negative on average in all four conditions.

Nevertheless, differences between conditions still reflect differences in overall race model inequality violation. As with the amount of aMRE, the average amount of race model inequality violation was larger when the stimulus was moved further away from the participants while the stimulus properties remained constant [*M Near High* = –15.382, SE = 3.501 vs. *M Far Low* = –5.097, SE = 4.187, *t*(19) = –2.302, *p* = .033, *d* = –.594]. Interestingly, the same decrease in retinal image size and intensity did not result in an increase in the amount of race model inequality violation when stimuli were presented at the same distance [*M Near High* = –15.382, SE = 3.501 vs. *M Near Low* = –15.688, SE = 5.906, *t*(19) = –.069, *p* = .946, *d* = –.013]. Presenting the audiovisual stimulus at a larger distance, but correcting for retinal image size and intensity, did not result in a larger amount of race model inequality violation [*M Near High* = –15.382, SE = 3.501 vs. *M Far High* = –8.188, SE = 4.282, *t*(19) = –1.638, *p* = .118, *d* = –.409]. The difference between the *Far High* and the *Far Low* intensity condition was not significant [*M Far Low* = –5.097, SE = 4.187 vs. *M Far High* = –8.188, SE = 4.282, *t*(19) = .586, *p* = .565, *d* = .163], indicating that a decrease in intensity and retinal image size did not result in larger race model inequality violations in far space.

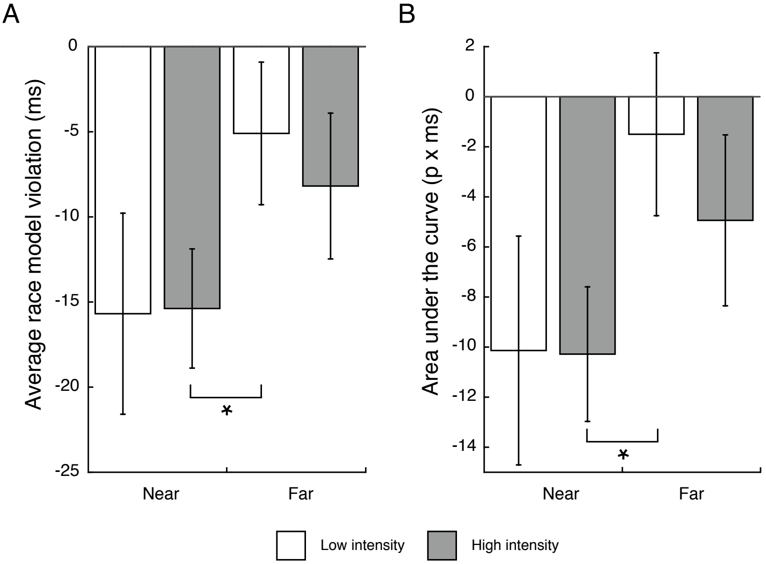


Figure 6. A: The average amount of race model inequality violation in each condition. B: The average area under the curve in each condition (right panel). More positive values reflect a larger amount of race model inequality violation. Significant differences are indicated with an asterisk (*p* < .05).

The results of the signed area under the curve measure were the same as the average amount of race model inequality violation. The average net area under the curve was more positive in the *Far Low* condition compared to the *Near High* condition (*M Near High* = –10.282, SE = 2.687 vs. *M Far Low* = –1.499, SE = 3.253, $t(19) = -2.487$, $p = .022$, $d = -.656$), indicating that audiovisual integration was enhanced when the audiovisual stimulus was presented in far space compared to a stimulus that was physically the same, but was presented in near space. None of the other comparisons were significant ($0 > t$'s > -1.8, all p's $> .1$).

The response time range of race model inequality violations

Comparing the probabilities between the audiovisual CDF and the race model at a range of RTs (100–1000 ms) revealed the RT range in which the race model was violated (see Figure 7). Only positive deviations from the race model were of interest (see the bar below each graph in Figure 7), as the race model inequality provides an upper bound of facilitation, not a lower bound. We observed significant violations in the *Near High* condition between 219 and 243 ms (range 25 ms, see panels a, b, or c of Figure 7). In the *Near Low* condition, there were no statistically significant race model inequality violations in terms of differences in probability (see panel c of Figure 7)[14]. The range of RTs in which violations occurred in the *Far Low* condition was from 216 to 282ms (range: 67ms, see panel a or d of Figure 7), and from 203 to 249 ms (range 47 ms) and 255–262 ms in the *Far High* condition (range 8 ms, see panel a or d of Figure7).

The RT range analyses indicate that an increase in distance mainly resulted in an increase in the width of the range in which race model inequality violations occurred, from a range of 25 ms in the *Near High* condition to an average range of 47 ms in the *Far High* condition (compare the gray and black bars in panel a of Figure 7). A decrease in intensity and retinal image size mainly resulted in a shift of the RT range of significant race model inequality violations to longer RTs (compare the gray and black curves in panel c and d of Figure 7). Race model inequality violations were significant starting with RTs larger than 219 ms in the *Near High* condition. Although no significant violations of the race model inequality in terms of probability differences were observed in the *Near Low* condition, a clear shift to later RTs can be seen in Figure 7C. In the *Far High* condition, race model inequality violations were significant starting from RTs larger than 203 and from 216 ms in the *Far Low* condition (a shift of 13 ms).

---

[14] When a different form of the race model inequality was used which does not assume a maximal negative correlation of –1 between signals [$P(RTAV \le P(RTA) + P(RTV) - P(RTA) \times P(RTV)$], violations were observed from 267 to 285 ms (range 18 ms).

The observed stronger MSI when the same stimulus was presented in far space (~200 cm) could thus be explained by the combined effects of an increase in distance on the one hand, and a decreased retinal image size and intensity on the other. This was apparent in larger race model inequality violations (see Figs. 6, 7), and also a larger range in which race model inequality violations occurred in terms of both quantiles (Figure 5) and a broader RT range (Figure 7).

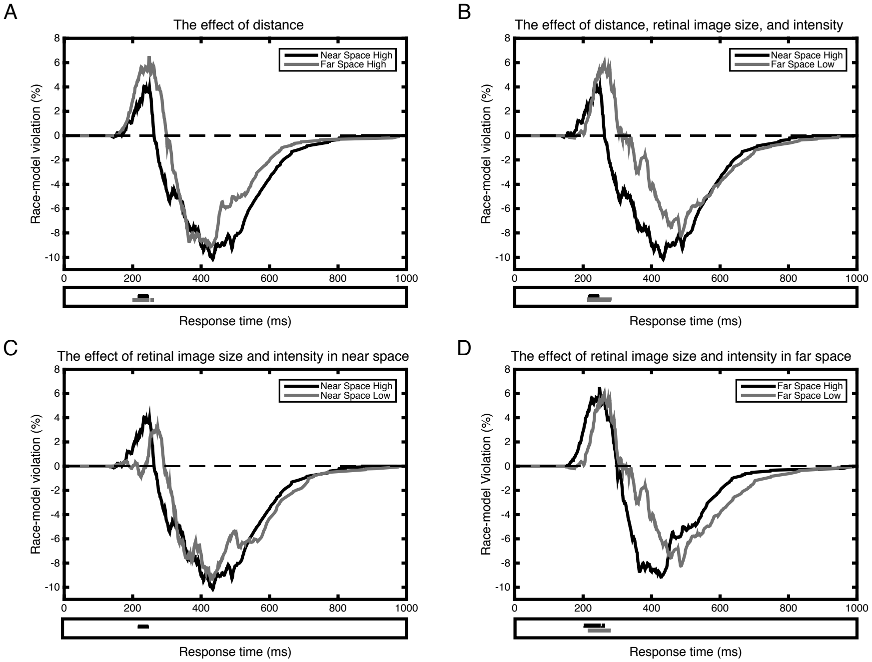


Figure 7. The difference curves that show the amount of race model inequality violation in each condition in terms of differences in probability across the full range of time points. A: The effect of an increase in distance on race model inequality violation, B: The effect of an increase in distance combined with a decrease in intensity and retinal image size, C: The effect of a decrease in intensity and retinal image size in near space, and D: The effect of a decrease in intensity and retinal image size in far space. The bars below each graph along the x-axis indicate significant positive violations of the race model inequality ($p < .05$, uncorrected).

Maximum race model inequality violation and shifts in the time point of the maximum violation

To further explore differences in race model inequality violation, the maximum probability difference and the corresponding time point at which these maximum race model inequality violations occurred were compared between the conditions (see

Table 1). Note that the average maximum probability difference reported here does not correspond with the observed maximum probability differences in Figure 7. The difference curves that are depicted in Figure 7 are averaged across subjects. As each subject showed race model inequality violations at slightly different time points, the average difference curves do not show the same maximum probability difference [that is why we also analyzed differences in RT for comparable points (quantiles) on the CDFs of each participant and each condition, see Figure 5].

Table 1 The average maximum probability difference between the AV CDF and the race model, and the average corresponding time point of the maximum race model inequality violation in each condition. The standard errors are presented between brackets.

Condition	Max. violation (%)	Time point of max. (ms)
Near Space High Intensity	14.03 (1.83)	285 (14)
Near Space Low Intensity	14.10 (2.07)	295 (22)
Far Space High Intensity	16.74 (2.56)	286 (13)
Far Space Low Intensity	18.45 (2.17)	319 (16)

There were no significant differences in the average maximum race model inequality violation between the conditions. There was, however, a significant difference in the time point of the maximum race model inequality violation between the *Near High* and the *Far Low* condition [$M\ T_{\text{Near High}} = 285$ ms, $SE = 14$, $T_{\text{Far Low}} = 319$ ms, $SE = 16$, $t(19) = -2.910$, $p = .009$, $d = -.506$]. The other differences in time points of the maximum race model inequality violation were not significant ($0 > t\text{'s} > -1$, all $p\text{'s} > .5$).

A decrease in retinal image size and intensity seemed to cause a shift of the time point of the maximum amount of race model inequality violation to longer RTs. The increase in the distance between stimuli and the observer seemed to slightly increase the average maximum amount of race model inequality violation, but these differences were not statistically significant. Only when stimuli were presented from far space and were of smaller retinal image size and intensity, did the shift in time point change significantly.

Discussion

Multisensory integration is generally assumed to be enhanced for weakly effective stimuli as compared to strongly effective stimuli (i.e., the principle of inverse effectiveness) (e.g., Holmes 2007, 2009; Leone & McCourt, 2013; Meredith & Stein, 1983). Yet, conflicting findings have been reported with respect to the principle of inverse effectiveness. What these studies have in common is that stimuli were always presented in a single depth plane. Given that increases in the distance between auditory and visual stimuli from the observer are generally related to decreases in retinal image size and intensity, audiovisual integration may be especially enhanced when decreases in retinal image size and intensity co-occur with increases in distance.

Our findings indicate that a decrease in retinal image size and stimulus intensity resulted in stronger audiovisual integration, but only when the stimulus was also presented at a larger distance from the observer (i.e., in far space). One could argue that moving a stimulus further away while keeping the stimulus physically the same causes an increase in audiovisual integration because of the principle of inverse effectiveness. This was, however, not the case, as the same decrease in intensity and retinal image size without an increase in distance did not result in enhanced audiovisual integration. Furthermore, the observed enhanced audiovisual integration in far space could not be explained solely based on the region of space in which the stimuli were presented, as an increase in the distance, while correcting the stimuli for intensity and retinal image size, did not significantly increase audiovisual integration.

A thorough analysis revealed how changes in retinal image size and intensity and changes in distance contributed to an increase in audiovisual integration in far space as compared to near space. A decrease in retinal image size and intensity without a change in the distance from which the stimuli were presented mainly resulted in a shift of the RT range to longer RTs in which race model inequality violations occurred. An increase in the distance while keeping the stimulus intensity and retinal image size constant caused an increase in the width of the RT range in which race model inequality violations occurred. Interestingly, however, the combination of these factors resulted in an effect that was different from the sum of these effects; an increase in the distance, and a decrease in stimulus intensity and retinal image size resulted in both an increase in the amount of race model inequality violation and an increase in the RT range in which violations occurred. It did not, however, result in a shift of the RT range in which violations occurred to longer RTs in far as compared to near space.

The current results thus indicate that stimulus efficacy and stimulus distance interact to increase audiovisual integration in far space. Changes in retinal image size and stimulus intensity alone did not increase audiovisual integration when the distance remained the same. However, given that only two retinal image sizes and stimulus intensities were used, which were also above threshold, it is difficult to take the current results as evidence against a general principle of inverse effectiveness. Although an increase in distance or a decrease in stimulus efficacy might independently affect audiovisual integration when large differences in distance or stimulus efficacy are used, the current findings indicate that their combined effects specifically enhance audiovisual integration at the distances used here (near: 80 cm, far: 208 cm). A comparison between supra-threshold stimuli and near-threshold stimuli in relation with distance-dependent changes in stimulus efficacy would require the presentation of stimuli at very large distance (this will also affect the differences in arrival times more strongly).

Multisensory interactions involving stimulation of the skin are often more pronounced when the source of multisensory stimulation is presented in peripersonal space as compared to extrapersonal space (see Van der Stoep, Nijboer, Van der Stigchel, & Spence, 2015 for a review). As auditory and visual perception does not depend on the distance between a stimulus and the body, no spatial asymmetry in depth is expected in terms of the strength of audiovisual interactions. Indeed, we did not observe any difference in audiovisual integration between information presented in near and far space when the audiovisual stimulus was corrected for intensity and retinal image size.

As for the cause of the current effects, there might be a role for multisensory experience. Several neurophysiological studies have shown the importance of prior experience with the environment in shaping the way the brain responds to multisensory stimuli (e.g., Wallace & Stein 1997, 2001, 2007). This is further underlined by results from behavioral studies in humans that indicate, for example, that the perceived temporal alignment of multisensory stimuli can be recalibrated based on prior multisensory experience (e.g., Vroomen, Keetels, de Gelder, & Bertelson, 2004; Machulla, Di Luca, Froehlich, & Ernst, 2012). A factor that changes the way we perceive multisensory information on a daily basis is the distance between audiovisual stimuli and the body. Typically, increasing the distance between stimuli and the observer results in a decrease in retinal image size and stimulus intensity when the stimulus properties of the source remain constant. Given that this relation is encountered in our everyday multisensory experiences, it might shape the way the brain responds to audiovisual stimuli. It has also been suggested that audition and vision are especially important for spatial orienting in far space (Previc, 1998). As a result of this distance-related multisensory experience, one might argue that relatively weak audiovisual stimuli in far space may be more helpful in spatial orienting as compared to the same stimuli when presented in near space. The observation of enhanced audiovisual integration in far space as compared to near space in a localization detection task is in line with this idea.

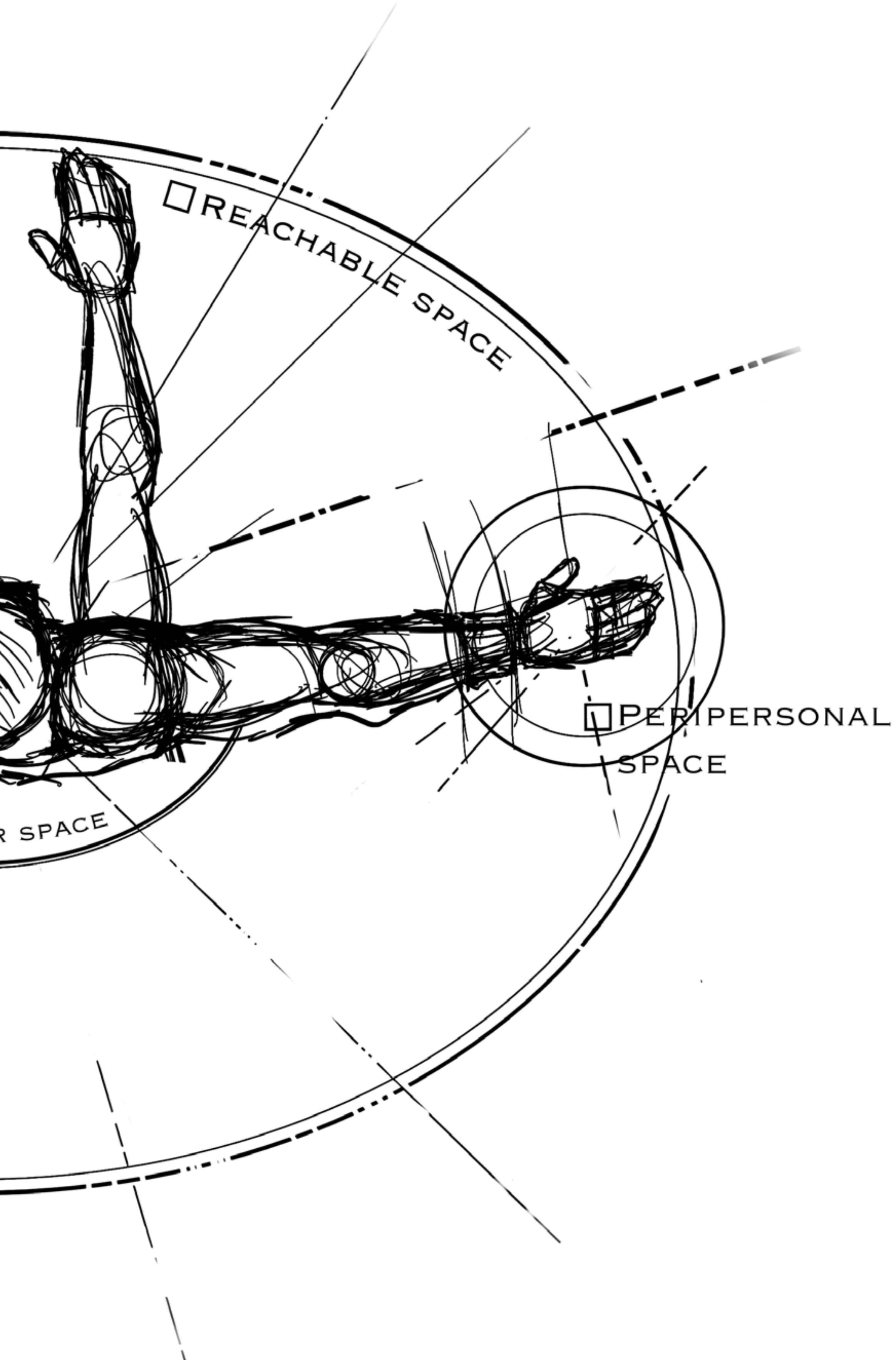
Another (underlying) factor that may contribute to the observed effects may be found in differences in spatial uncertainty of audiovisual sources between near and far space.

When comparing the retinal image size of, for example, a car in near space and far space, it is evident that the car covers a larger visual angle in near space as compared to far space. At a large distance, several cars could be observed in the same visual angle as that of the car in near space. Speculatively, audiovisual integration may be generally more helpful in spatial orienting in far space, when compared to near space, as the spatial uncertainty of the location of both visual and auditory information is higher (in this case the car).

The idea that spatial uncertainty enhances audiovisual integration is in line with the results from previous studies, in which it was shown that MSI is enhanced when the location of a multisensory target is unattended as compared to when it was attended (exogenous spatial attention: Van der Stoep, Van der Stigchel, & Nijboer, 2015; endogenous spatial attention: Zou, Müller, & Shi, 2012). Although seemingly unrelated to the explanation of our findings in terms of the uncertainty of the spatial location of stimuli, these studies demonstrated that MSI might not help as much during spatial localization when attention is already close to or at the target location as compared to when spatial attention is at a location that is far from the target location. In other words, MSI is stronger when spatial uncertainty is larger. In the current study, the broader window during which MSI occurred in far space as compared to near space may reflect a tendency for the brain to accumulate as much evidence as possible for the spatial location of the stimuli in far space because of a larger spatial uncertainty.

In the current study, the distance from which the target stimulus was presented was blocked. Therefore, participants could focus their attention endogenously at a certain depth, reducing spatial uncertainty in the depth space. It might be interesting in future research to investigate how the uncertainty of audiovisual information in depth affects audiovisual integration in near and far space. This may also depend on where attention is focused in terms of depth as some studies have observed faster responses to stimuli occurring between the body and the focus of attention, compared to stimuli presented beyond the focus of attention (e.g., de Gonzaga Gawryszewski, Riggio, Rizzolatti, & Umiltá, 1987). The current findings indicate that when endogenous attention is focused at the depth from which information is presented, audiovisual integration is enhanced in far space relative to near space when the stimulus is not corrected for retinal image size and intensity (i.e., a smaller retinal image size and a lower visual and auditory intensity in far space as compared to near space).

To conclude, we observed stronger audiovisual integration in far space when stimuli were not corrected for retinal image size and intensity as compared to the same audiovisual stimuli in near space. The increase in integration could not be explained in terms of inverse effectiveness alone, as the same decrease in retinal image size and intensity did not result in an increase in audiovisual integration in near space. The presentation of audiovisual stimuli in far space did seem to slightly increase the amount of audiovisual integration, but only when an increase in distance was combined with a decrease in retinal image size and intensity, was the enhanced integration in far space significantly different from that in near space. These results underline the importance of taking the distance from which information is presented into account when investigating multisensory interactions.


☐Reachable space
☐Peripersonal space
r space

Chapter 10

Summary and conclusions

The goal of this thesis was two-fold: 1) to investigate the relative contributions of crossmodal exogenous spatial attention and multisensory integration to multisensory response enhancement, and how attention and multisensory integration interact, and 2) to get a better understanding of how the distance from which information is presented affects (crossmodal exogenous) spatial attention and multisensory integration. This chapter provides an overview and discussion of the results from chapters 2 to 9 in light of the literature on crossmodal (exogenous) spatial attention and multisensory integration. The current findings also provide guidelines for the use of multisensory stimulation to enhance spatial orienting in clinical and applied settings. Therefore, the implications of each of our findings are also discussed in that context.

Into the depths of audiovisual interactions

The simultaneous presentation of information to different senses often results in behavioural benefits like faster detection times as compared to when only a single sense is stimulated. Given that both crossmodal exogenous spatial attention and multisensory integration are able to cause such behavioural benefits, it was investigated what the contribution of each of these two processes is to multisensory response enhancement (MRE, chapter 2). Multisensory response enhancement was caused by multisensory integration when sound and light were spatially and temporally aligned. This observation is in line with two principles of multisensory integration: the spatial and temporal rule (e.g., Stein & Stanford, 2008). When there was a small delay between the onset of the sound and the light (50 ms), both crossmodal exogenous spatial attention and multisensory integration contributed to MRE. As the delay between the onset of the sound and the light increased (100 and 200 ms onset difference), multisensory response enhancement was due to crossmodal exogenous spatial attention rather than multisensory integration. It should be noted that non-spatial alerting and response preparation can also contribute to the speeding-up of responses to multisensory stimuli (Los & Van der Burg, 2013). Overall then, the processes that contribute to MRE very much depend on the spatial and temporal structure of multisensory stimuli.

Besides the theoretical importance of these findings, they can also be of use in applied contexts. For example, in situations where responses should be as fast and as accurate as possible (e.g., in dangerous traffic situations while driving or while piloting an aircraft), one can imagine that in terms of absolute response times it can be very helpful to present a crossmodal exogenous auditory cue before the onset of a visual warning signal or an approaching (seen) threat. After all, a spatially aligned exogenous auditory cue decreased response times with ~80 ms when the spatial location of the visual target was task-relevant (see chapter 2, Figure 4A; for more on the use of sensory warning signals in the context of driving, Ho & Spence, 2006; Ho, Tan, & Spence, 2005; Meng & Spence, 2015; Spence & Ho, 2008). It has been shown, however, that the ability of exogenous cues to attract attention to their location is dependent on cognitive load (Spence & Santangelo, 2009). Whereas the presentation of unimodal (light or sound alone) and multimodal cues (sound and light) cause the same benefits in terms of response times when the cognitive load is low, only multimodal cues are able to still facilitate responses when an observer is involved in a secondary task (i.e., under high cognitive load). Auditory and visual stimuli that are integrated are thus more robust to high cognitive load in terms of their ability to attract exogenous spatial attention and facilitate behaviour than unimodal stimuli. Cognitive load is likely high in the case of driving, as the driver has to monitor multiple things at the same time. It therefore makes sense that the use of multisensory cues will result in the largest amount of facilitation whenever sensory warning signals are implemented in a car to reflexively attract a driver's spatial attention to a source of threat or important warning signals inside the car (Spence & Ho, 2008). Knowing at which stimulus intervals auditory and visual warning signals are still integrated while crossmodal exogenous spatial attention also starts to contribute to MRE can thus be very helpful in the design of optimal sensory warning signals.

Whereas in the previous section the attention attracting capacity of multisensory signals and their effects on the processing of later presented signals were discussed, here I will be talking about how attention can affect the (outcome of the) process of multisensory integration (e.g., can spatial attention enhance multisensory integration?). Previous studies of the effects of *endogenous* spatial attention on multisensory integration indicate that this is indeed the case whenever a multisensory stimulus is voluntarily attended (Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010). However, the opposite effects have also been observed (e.g., Santangelo, Ho, & Spence, 2008; Spence & Santangelo, 2009; Zou, Müller, & Shi, 2012). These studies thus indicate that endogenous spatial attention can modulate multisensory integration and, depending on the situation, can enhance or decrease multisensory integration. Whereas endogenous spatial attention is considered a top-down process, multisensory integration of simple lights and sounds is considered a bottom-up process (i.e., 'early' multisensory integration). Up until now it was unclear whether a bottom-up form of spatial attention (i.e., exogenous spatial attention) could also modulate early multisensory integration. The results of one study on this interaction did not point into the direction of any modulation (Vroomen, Bertelson, & de Gelder, 2001). However, as described in chapter 3, the circumstances of finding any effect of exogenous spatial attention on multisensory integration in that study were not optimal. By presenting a spatial cue that was separate from a multisensory target at a cue-target interval of about 200 ms, we were able to find that exogenous spatial attention and multisensory integration do interact. Interestingly, multisensory integration was reduced for exogenously attended spatial locations as compared with unattended locations. This effect was most pronounced when spatial orienting was task-relevant as compared to when it was task-irrelevant. Exogenous spatial attention is known to increase perceptual sensitivity, which is effectively similar to an increase in perceived intensity (Carrasco, 2011). According to the principle of inverse effectiveness, multisensory integration is enhanced when the unimodal stimuli are only weakly effective as compared to when they are strongly effective (Meredith & Stein, 1986). Exogenous spatial attention may thus modulate 'early' multisensory integration by increasing the perceived intensity of the audiovisual stimulus at attended as compared to unattended locations. In constrast with previous findings, our results indicate that interactions between attention and multisensory integration can already occur at an early level of multisensory processing (cf. Koelewijn, Bronkhorst, & Theeuwes, 2010). Despite multisensory integration being enhanced at exogenously unattended locations, responses to exogenous attended multisensory targets were faster than to exogenously unattended multisensory targets. Furthermore, it might be interesting to see whether increasing the cognitive load will affect the processing of exogenously unattended multisensory targets less than attended targets, as different amounts of integration were observed under these conditions.

At longer delays between the onset of an exogenous spatial cue and a multisensory target, responses to a target are often slower at cued as compared to uncued locations. This inhibitory after-effect is often labeled inhibition of return (IOR). This effect of IOR can be the result of inhibition at different levels of processing (e.g., at a perceptual or a motor level; Klein, 2000; Lupianez, Klein, & Bartolomeo, 2006) and seems to promote orienting to novel locations. In light of this idea, we hypothesized that multisensory integration might be enhanced at cued relative to uncued locations once IOR comes into play. Yet, the findings of our study of the effects of IOR on multisensory integration showed the opposite pattern. Although absolute response times to multisensory targets did not differ between cued and uncued locations, the amount of multisensory response enhancement and multisensory integration was *decreased* at *cued* locations (chapter 4). A further analysis of response times indicated that the difference in auditory and visual processing times was increased at cued as compared to uncued locations as a result of IOR for visual targets. This increase in the difference between unimodal response times caused the auditory stimulus to be more dominant in terms of processing times, which led to statistical facilitation rather than multisensory integration at the cued location. This observation is in line with previous studies of multisensory integration in which it was shown that integration is most pronounced when both sensory signals are about equally informative (e.g., Ernst & Banks, 2002; Otto, Dassy, & Mamassian, 2013).

In the studies in chapters 2 to 4 there was never more than one visual target onset at a given moment during a trial. The presence of multiple visual events at the same time can cause competition in, for example, the oculomotor system. It is thought that each visual element can evoke a vector in the oculomotor system (Tipper, Howard, & Jackson, 1997; Van der Stigchel, 2010). When two visual elements are presented close together is space, this competition can result in saccade averaging (Van der Stigchel & Nijboer, 2011). In that case saccades most often land at a location in between the two visual elements instead of on the center of one of the two elements. In chapter 5 we investigated how the onset of a non-lateralized sound affected this oculomotor competition. The participants were instructed to make an eye-movement as quickly as possible towards visual information that appeared on a monitor. On half of the trials a sound was presented through headphones simultaneously with the onset of visual information. We hypothesized that when the sound was automatically coupled to one of the two visual elements, the global effect should decrease compared to the no sound condition, as competition would favor one element more than the other. On the other hand, given that audiovisual coupling was totally ambiguous in this context because the sound was not lateralized, the sound could be associated with the onset of both visual elements. In that case both vectors that are evoked by the visual targets are boosted by the presence of the sound. This will not only cause the individual vectors to be enhanced, but also the averaged vector. Although this will not change the direction of the average saccade, it does increase the probability that saccade will land in between the two visual elements. This is what we observed. The global effect was stronger when a sound was simultaneously presented with the visual targets as compared to when no sound was presented. One could argue that there is no difference in the increase in activity of the average vector *relative* to the single vectors as a result of the sound. Yet, given that the enhancement of the single vectors because of the sound did *not* result in increased (or decreased) accuracy (as evident from same mean and *SD*, see chapter 5,

Experiment 1), and the fact that there was still room for the *average* vector to become more 'accurate' (i.e., closer to the exact average direction), a *relative* increase in saccades that are initiated towards the exact average is still expected.
Taken together, the studies described in chapters 2 to 5 have build upon and contributed to our understanding of the nature of audiovisual interactions. They have provided new insights into the circumstances that determine which multisensory processes contribute to multisensory response enhancement. Additionally, they revealed the importance of the specific spatial and temporal relation between auditory and visual stimuli in determining how crossmodal exogenous spatial attention and multisensory integration interact. We also showed that a spatially ambiguous sound could influence oculomotor competition evoked by multiple visual targets in a scene.

Multisensory interactions in the depth-plane
The second part of this thesis was concerned with the question whether and how the distance or region of space from which information was presented affects multisensory interactions. The current body of neurophysiological and neuropsychological evidence suggests that the space surrounding the body and the space that is further away from the body and out of reach are coded differently by the brain (Cowey, Small, & Ellis, 1994; Graziano, Hu, & Gross, 1997; Graziano, Reiss, & Gross, 1999; Gross & Graziano, 1995; Halligan & Marshall, 1991; Pegna et al., 2001; Previc, 1998). In his seminal paper on the neuropsychology of spatial perception, Previc (1998) suggested that the relevance of each of our senses differs depending on the region of space in which information is presented. For example, visual and somatosensory systems are especially relevant for the guiding of reaching movements and are therefore more dominant in peripersonal space as compared to extapersonal (action) space. In contrast, vision and audition are especially relevant for navigation, orientation, and scene memory in extrapersonal (action space). In our review of multisensory interactions in front and rear peripersonal and extrapersonal space (chapter 6) we highlighted that those multisensory interactions involving touch are especially enhanced when multisensory stimuli are presented close to the body. This asymmetry in depth can be explained by the fact that spatial alignment between tactile and visual (or auditory) information can only occur when all sources of stimulation are near the body. This idea fits well with the observation of multisensory neurons that integrate visual and tactile information that are presented at the same spatial location in 3-D space (e.g., Graziano & Gross, 1994). Because the distance from which this information is presented does not limit the perception of auditory and visual information, no asymmetry in depth is expected for

audiovisual interactions based on their spatial alignment in depth. This is exactly what we observed for crossmodal exogenous spatial attention (chapter 7) and audiovisual integration in near and far space (chapter 9).

In the study described in chapter 7 we presented exogenous auditory cues in near or far space on the left or right side of a central fixation cross prior to the onset of visual targets that were presented to the left or the right in near or far space. Overall, responses to targets that appeared at the same lateral location as the cue were faster compared to when cues and targets were presented at different lateral locations. This (classic) lateral validity effect is considered to be the result of a shift of crossmodal exogenous spatial attention to the cued location (Spence & Driver, 1997). More importantly, the lateral validity effect was only present when the auditory exogenous cue and the visual target were presented at the same distance from the observer (e.g., the presentation of a cue and a target in far space). This effect of spatial alignment in depth did not depend on the region of space in which the visual target was presented. That is, as long as the cue and the target were aligned in 3-D space target processing was facilitated. This is not surprising, given that visual and auditory perception is not bound to the body. These results indicate that crossmodal interactions between audition and vision in the context of crossmodal exogenous spatial attention depend on spatial alignment in depth, but not on the region of space in which the stimuli happen to be presented.

The results from our study of visuospatial neglect in peripersonal and extrapersonal space indicate that visuospatial attention can be selectively impaired (chapter 8). Using three differrent tasks to measure visuospatial neglect, four groups of patients could be identified: (1) patients with visuospatial neglect in peripersonal but not extrapersonal space, (2) patients with visuospatial neglect in extrapersonal but not peripersonal space, (3) patients with visuospatial neglect in both regions of space, and (4) stroke patients without any indication of visuospatial neglect. These results indicate that visual information that is presented in peripersonal and extrapersonal space is processed differently and can therefore be selectively impaired.

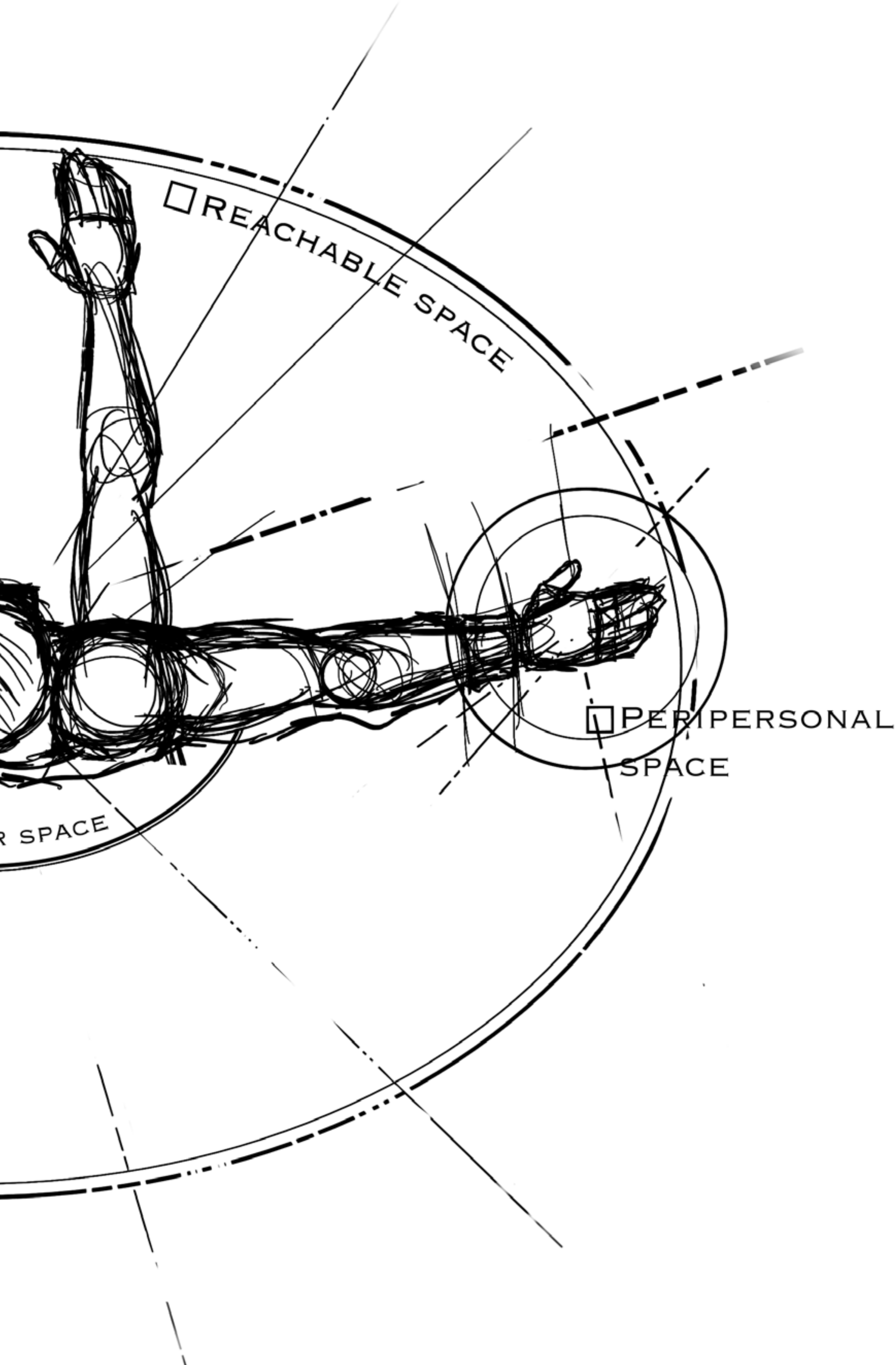
As was shown in chapter 7, the region of space in which information is presented does not necessarily affect shifts of crossmodal exogenous spatial attention as long as stimuli from different modalities are spatially aligned in 3-D space. This might be different for audiovisual integration though as the distance from which information is presented modulates several stimulus factors. For example, differences in auditory and visual arrival times will increase as a result of the differing speeds of sound and light. Several studies have shown that this difference can be taken into account when judging the simultaneity of auditory and visual onsets, but also that this difference is not automatically corrected for by the brain (Alais & Carlile, 2005; Arnold, Johnston, & Nishida, 2005; Heron, Whitaker, McGraw, & Horoshenkov, 2007; Sugita & Suzuki, 2003). As a result, auditory and visual onsets may fall outside the temporal binding window at large distances (Spence & Squire, 2003). Furthermore, the retinal image size of visual stimuli, the direct-to-reverberant ratio of sound (Bronkhorst & Houtgast, 1999), and the perceived intensity of both sound and light will change as the distance between the stimuli and the observer increase (Warren, 1963).

All these factors may have an impact on multisensory integration because they can change stimulus efficacy (principle of inverse effectiveness) and auditory and visual arrival times (temporal rule). As mentioned before, audiovisual information is thought to be more dominant in extrapersonal as compared to peripersonal space, while the opposite is true of visuotactile information (Previc, 1998). This may lead stronger audiovisual integration in far as compared to near space. The results of our study study described in chapter 9 were in line with this idea. We observed that audiovisual integration was enhanced for stimuli that were presented in far space, but only when the stimuli were not corrected for visual angle and intensity. Thus, when the exact same audiovisual stimulus was moved further away from the participant, audiovisual integration increased. At first sight this increase in integration seems to be easily explained in terms of the principle of inverse effectiveness: stimuli that are only weakly effective when presented in isolation will evoke a larger benefit of multisensory integration as compared to stimuli that are strongly effective (Meredith & Stein, 1983). The weaker stimuli in far space may therefore be more effectively integrated than the same stimuli in near space because of inverse effectiveness. However, the increase in audiovisual integration in our study could not be explained by a decrease in stimulus efficacy alone, as the exact same decrease in stimulus efficacy while presenting the stimuli in near space did not increase audiovisual integration. One could also argue that audiovisual integration was enhanced because auditory and visual information is dominant in far space. Yet, an increase in distance while the stimuli were corrected for retinal image size and intensity did not increase integration in far space. It was thus the combination of a decrease in stimulus efficacy and an increase in the distance from which information was presented that enhanced audiovisual integration.

Apart from the theoretical advances in our understanding of crossmodal exogenous spatial attention and multisensory integration in the depth-plane gained by these studies, they also provide opportunities for the use of multisensory stimulation in applied settings. For example, in a situation in which it is important to attract a driver's attention to frontal extrapersonal space (the location where threats are generally perceived during driving, such as when a car suddenly breaks in front of the driver), it seems best to present warning signals that come or appear to come from extrapersonal space, attracting attention to their location. Additionally, given that audiovisual integration is enhanced in far space, there may be a larger benefit of spatially and temporally aligned audiovisual warning signals in far as compared to near space.

Multisensory stimulation can lead to enhanced localization performance in both healthy subjects (Frassinetti, Bolognini, & Làdavas, 2002) and in subjects with brain damage (Frassinetti, Bolognini, Bottari, Bonora, & Làdavas, 2005; Frassinetti, Pavani, & Làdavas, 2002). Multisensory stimulation may thus alleviate some impairments that occur after brain damage. Although multisensory stimulation has been applied for some time now, it is starting to receive more attention now that the principles underlying multisensory interactions are beginning to become more clearly understood (Borghese, Bottini, & Sedda, 2013; Johansson, 2012). Given that distance-specific impairments in, for example, spatial attention have been observed after stroke (e.g., Aimola et al., 2012; chapter 8), it will be interesting to see whether such impairments are perhaps less pronounced during multisensory stimulation at specific distances. In one of the studies that we are currently conducting, we are investigating how audiovisual integration is affected by temporal-parietal brain lesions to gain a better understanding of how robust multisensory interactions are to such lesions (Van der Stoep, Van Engelen, Van der Stigchel, & Nijboer, in prep.). Our preliminary analyses indicate that although in some patients integration was reduced to the level of statistical facilitation in the contralesional visual hemifield, they still benefitted from multisensory stimulation as shown by faster responses to multisensory as compared to unimodal targets. These results could be taken to suggest that patients may still benefit from multisensory stimulation, even if multisensory integration has been affected.

Whereas past research on crossmodal spatial attention and multisensory integration has provided a foundation for the understanding of audiovisual interactions (e.g., Posner & Cohen, 1984; Stein & Meredith, 1993; Stein & Stanford, 2008; Spence & Driver, 2004; Spence & McDonald, 2004, to name a few), the studies in this thesis have revealed more details of the complex nature of the interactions between audition and vision. The previous chapters do not only indicate that interactions between exogenous spatial attention and multisensory integration can occur at very early stages of processing, they also underline the importance of investigating multisensory interactions in 3-D space.


☐ Reachable space
☐ Peripersonal space
r space

References

Agganis, B. T., Muday, J. A., Schirillo, J. A. (2010). Visual biasing of auditory localization in azimuth and depth. *Perceptual and Motor Skills*, *111*(3), 872–892.

Aimola, L., Schindler, I., Simone, A. M., & Venneri, A. (2012). Near and far space neglect: Task sensitivity and anatomical substrates. *Neuropsychologia*, *50*(6), 1115–1123.

Alais, D., & Carlile, S. (2005). Synchronizing to real events: subjective audiovisual alignment scales with perceived auditory depth and speed of sound. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(6), 2244–2247.

Alsius, A., Navarra, J., & Soto-Faraco, S. (2007). Attention to touch weakens audiovisual speech integration. *Experimental Brain Research*, *183*(3), 399–404.

Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Current Biology*, *15*(9), 839–843.

Alvarado J. C., Vaughan J. W., Stanford T. R., & Stein, B. E. (2007). Multisensory versus unisensory integration: Contrasting modes in the superior colliculus. *Journal of Neurophysiology*, *97*(5), 3193–3205.

Andersen, R. A. (1997). Multimodal integration for the representation of space in the posterior parietal cortex. *Philosophical Transactions of the Royal Society B*, *352*(1360), 1421–1428.

Arnold, D. H., Johnston, A., & Nishida, S. (2005). Timing sight and sound. *Vision Research*, *45*(10), 1275–1284.

Atchley, P., Kramer, A. F., Andersen, G. J., & Theeuwes, J. (1997). Spatial cuing in a stereoscopic display: Evidence for a "depth- aware" attentional focus. *Psychonomic Bulletin & Review*, *4*(4), 524– 529.

Avillac, M., Deneve, S., Olivier, E., Pouget, A., & Duhamel, J. R. (2005). Reference frames for representing visual and tactile locations in parietal cortex. *Nature Neuroscience*, *8*(7), 941–949.

Barbieri, C., & de Renzi, E. (1989). Patterns of neglect dissociation. *Behavioural Neurology*, *2*(1), 13–24.

Barton, J. J. S., & Black, S. E. (1998). Line bisection in hemianopia. *Journal of Neurology, Neurosurgery & Psychiatry*, *64*, 660–662.

Bassolino, M., Serino, A., Ubaldi, S., Làdavas, E., 2010. Everyday use of the computer mouse extends peripersonal space representation. *Neuropsychologia*, *48*(3), 803–811.

Bauer, D., Plinge, A., Ehrenstein, W. H., Rinkenauer, G., & Grosjean, M. (2011). Spatial orienting of attention in stereo depth. *Psychological Research*, *76*(6), 730–735.

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.

Bertelson, P., Pavani, F., Ladavas, E., Vroomen, J., & de Gelder, B. (2000). Ventriloquism in patients with unilateral visual neglect. *Neuropsychologia*, *38*(12), 1634–1642.

Bertelson, P., Vroomen, J., de Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics*, *62*(2), 321–332.

Berti, A., & Frassinetti, F. (2000). When far becomes near: Remapping of space by tool use. *Journal of Cognitive Neuroscience*, *12*(3), 415–420.

Bisiach, E., & Luzzatti, C. (1978). Unilateral neglect of representational space. *Cortex*, *14*(1), 129–133.

Bisiach, E., Cornacchia, L., Sterzi, R., & Vallar, G. (1984). Disorders of perceived auditory lateralization after lesions of the right hemisphere. *Brain*, *107*(1), 37–52.

Borghese, N. A., Bottini, G., & Sedda, A. (2013). Videogame based neglect rehabilitation: a role for spatial remapping and multisensory integration? *Frontiers in Human Neuroscience*, *7*(16), 1-3.

Bourgeois, J., Farnè, A., Coello, Y., 2014. Costs and benefits of tool-use on the perception of reachable space. *Acta Psychologica*, *148*(5), 91–95.

Bowen, A., McKenna, K., & Tallis, R. C. (1999). Reasons for variability in the reported rate of occurrence of unilateral spatial neglect after stroke. *Stroke*, *30*(6), 1196–1202.

Bronkhorst, A. W., & Houtgast, T. (1999). Auditory distance perception in rooms. *Nature*, *397*(6719), 517–520.

Bruce, C., Desimone, R., Gross, C. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*, *46*(2), 369–384.

Bunt, A. H., Minckler, D. S., Johanson, G. W. (1977). Demonstration of bilateral projection of the central retina of the monkey with horseradish peroxidase neuronography. *Journal of Comparative Neurology*, *171*, 619–630.

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 the United States of America*, *102*(51), 18751–18756.

Butler, B. C., Eskes, G. A., & Vandorpe, R. A. (2004). Gradients of detection in neglect: Comparison of peripersonal and extrapersonal space. *Neuropsychologia*, *42*(3), 346–358.

Butter, C. M., Evans, J., Kirsch, N., & Kewman, D. (1989). Altitudinal neglect following traumatic brain injury: A case report. *Cortex*, *25*(1), 135–146.

Caggiano, V., Fogassi, L., Rizzolatti, G., Thier, P., & Casile, A. (2009). Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. *Science*, *324*(5925), 403–406.

Calabro, F.J., Soto-Faraco, S., & Vaina, L.M. (2011). Acoustic facilitation of object movement detection during self-motion. *Proceedings of the Royal Society B*, *278*, 2840–2847.

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

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

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

Canzoneri, E., Ubaldi, S., Rastelli, V., Finisguerra, A., Bassolino, M., & Serino, A. (2013). Tool-use reshapes the boundaries of body and peripersonal space representations. *Experimental Brain Research*, *228*(1), 25–42.

Caplan, B. (1985). Stimulus effects in unilateral neglect. *Cortex*, *21*(1), 69–80.

Cardellicchio, P., Sinigaglia, C., & Costantini, M. (2011). The space of affordances: A TMS study. *Neuropsychologia*, *49*(5), 1369–1372.

Carrasco, M. (2011). Visual attention: the past 25 years. *Vision Research*, *51*(13), 1484–1525.

Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. (2000). Spatial cover attention increases contrast sensitivity across CSF: Support for signal enhancement. *Vision Research*, *40*(10), 1203–1215.

Cassidy, T. P., Lewis, S., & Gray, C. S. (1998). Recovery from visuospatial neglect in stroke patients. *Journal of Neurology, Neurosurgery & Psychiatry*, *64*(4), 555–557.

Chen, Q., Weidner, R., Vossel, S., Weiss, P. H., & Fink, G. R. (2012). Neural mechanisms of attentional reorienting in three-dimensional space. *Journal of Neuroscience*, *32*(39), 13352–13362.

Cherney, L. R., Halper, A. S., Kwasnica, C. M., Harvey, R. L., & Zhang, M. (2001). Recovery of functional status after right hemisphere stroke: Relationship with unilateral neglect. *Archives of Physical Medicine and Rehabilitation*, *82*(3), 322–328.

Cohen, Y. E., & Andersen, R. A. (2002). A common reference frame for movement plans in the posterior parietal cortex. *Nature Reviews Neuroscience*, *3*(7), 553–562.

Colonius, H., & Arndt, P. A. (2001). A two-stage model for visual-auditory interaction in saccadic latencies. *Perception & Psychophysics*, *63*(1), 126– 147.

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* (62), 1-10.

Committeri, G., Pitzalis, S., Galati, G., Patria, F., Pelle, G., Sabatini, U., ... & Pizzamiglio, L. (2007). Neural bases of personal and extrapersonal neglect in humans. *Brain*, *130*(2), 431-441.

Costantini, M., Ambrosini, E., Tieri, G., Sinigaglia, C., & Committeri, G. (2010). Where does an object trigger an action? An investigation about affordances in space. *Experimental Brain Research*, *207*, 95-103.

Cousineau, D. (2005). Confidence intervals in within-subject designs: a simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45.

Couyoumdjian, A., Di Nocera, F., & Ferlazzo, F. (2003). Functional representation of 3D space in endogenous attention shifts. *Quarterly Journal of Experimental Psychology: Section A*, *56*(1), 155–183.

Cowey, A., Small, M., & Ellis, S. (1994). Left visuo-spatial neglect can be worse in far than in near space. *Neuropsychologia*, *32*(9), 1059–1066.

Cowey, A., Small, M., & Ellis, S. (1999). No abrupt change in visual hemineglect from near to far space. *Neuropsychologia*, *37*(1), 1–6.

de Gonzaga Gawryszewski, L., Riggio, L., Rizzolatti, G., & Umiltá, C. (1987) Movements of attention in the three spatial dimensions and the meaning of "neutral" cues. *Neuropsychologia*, *25*(1), 19–29.

De Paepe, A.L., Crombez, G., Spence, C., & Legrain, V. (2014). Mapping nociceptive stimuli in a peripersonal frame of reference: evidence from a temporal order judgment task. *Neuropsychologia*, *56*, 219–228.

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.

Doricchi, F., Onida, A., & Guariglia, P. (2002). Horizontal space misrepresentation in unilateral brain damage: II. Eye-head centered modulation of visual misrepresentation in hemianopia without neglect. *Neuropsychologia*, *40*(8), 1118–1128.

Downing, C. J., & Pinker, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 171–187). Hillsdale, NJ: Erlbaum.

Doyle, M. C., & Walker, R. (2002). Multisensory interactions in saccade target selection: curved saccade trajectories. *Experimental Brain Research*, *142*(1), 116–130.

Driver, J., & Spence, C. (1998). Cross-modal links in spatial attention. *Philosophical Transactions of the Royal Society of London Series B*, *353*(1373), 1319-1331.

Eimer, M., & Van Velzen, J. (2002). Crossmodal links in spatial attention are mediated by supramodal control processes: Evidence from event- related potentials. *Psychophysiology*, *39*(4), 437–449.

Engel, G. R., & Dougherty, W.G. (1971). Visual-auditory distance constancy. *Nature*, *234*, 308.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*(2), 336-353.

Fairhall, S. L., & Macaluso, E. (2009). Spatial attention can modulate audiovisual integration at multiple cortical and subcortical sites. *European Journal of Neuroscience*, *29*(6), 1247–1257.

Farah, M. J., Wong, A. B., Monheit, M. A., & Morrow, L. A. (1989). Parietal lobe mechanisms of spatial attention: Modality-specific or supramodal? *Neuropsychologia, 27*(4), 461–470.

Farnè, A., Bonifazi, S., & Làdavas, E. (2005). The role played by tool-use and tool-length on the plastic elongation of peri-hand space: A single case study. *Cognitive Neuropsychology*, *22*(3-4), 408-418.

Farnè, A., Dematte, M. L., & Làdavas, E. (2003). Beyond the window: Multisensory representation of peripersonal space across a transparent barrier. *International Journal of Psychophysiology*, *50*(1), 51-61.

Farnè, A., & Làdavas, E. (2002). Auditory peripersonal space in humans. *Journal of cognitive neuroscience*, *14*(7), 1030-1043.

Felipe, N. J., & Sommer, R. (1966). Invasions of personal space. *Social problems*, 206-214.

Findlay, J. M. (1982). Global visual processing for saccadic eye movements. *Vision research*, *22*(8), 1033-1045.

Fogassi, L., Gallese, V., Fadiga, L., Luppino, G., Matelli, M., & Rizzolatti, G. (1996). Coding of peripersonal space in inferior premotor cortex (area F4). *Journal of Neurophysiolology*, *76*(1), 141–157.

Frassinetti, F., Bolognini, N., & Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research*, *147*(3), 332–343.

Frassinetti, F., Bolognini, N., Bottari, D., Bonora, A., & Làdavas, E. (2005). Audiovisual integration in patients with visual deficit. *Journal of Cognitive Neuroscience*, *17*(9), 1442–1452.

Frassinetti, F., Pavani, F., & Làdavas, E. (2002). Acoustical vision of neglected stimuli: Interaction among spatially converging audiovisual inputs in neglect patients. *Journal of Cognitive Neuroscience*, *14*(1), 62–69.

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

Frens, M. A., & Van Opstal, A. J. (1995). A quantitative study of auditory-evoked saccadic eye movements in two dimensions. *Experimental Brain Research*, *107*(1)*,* 103–117.

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

Gabbard, C., Cordova, A., & Ammar, D. (2007). Estimation of reach in peripersonal and extrapersonal space: A developmental view. *Developmental Neuropsychology, 32*(3), 749–756.

Gallace, A., & Spence, C. (2014). *In Touch with the Future: The Sense of Touch from Cognitive Neuroscience to Virtual Reality*. Oxford, UK: Oxford University Press.

Gardner, J. L., Merriam, E. P., Movshon, J. A., & Heeger, D. J. (2008). Maps of visual space in human occipital cortex are retinotopic, not spatiotopic. *Journal of Neuroscience*, *28*(15), 3988–3999.

Gardner, M. B. (1968). Proximity image effect in sound localization. *The Journal of the Acoustical Society of America*, *43*(1), 163-163.

Gielen, S. C., Schmidt, R. A., & Van Den Heuvel, P. J. (1983). On the nature of intersensory facilitation of reaction time. *Perception & Psychophysics*, *34*(2), 161-168.

Girard, S, Pelland, M., Lepore, F., & Collignon, O. (2013). Impact of the spatial congruence of redundant targets on within-modal and cross-modal integration. *Experimental Brain Research*, *224*(2), 275–285.

Girard, Simon, Collignon, O., & Lepore, F. (2011). Multisensory gain within and across hemispaces in simple and choice reaction time paradigms. *Experimental Brain Research*, *214*(1), 1–8.

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

Gray, R., & Tan, H. Z. (2002). Dynamic and predictive links between touch and vision. *Experimental Brain Research*, *145*(1), 50-55.

Graziano M. S., & Gross, C. G. (1994). The representation of extrapersonal space: A possible role for bimodal visual-tactile neurons. In: Gazzaniga MS (Ed.) *The cognitive neurosciences* (1021–1034). Cambridge: MIT Press.

Graziano, M. S., Hu, X. T., & Gross, C. G. (1997). Visuospatial properties of ventral premotor cortex. *Journal of Neurophysiology*, *77*(5), 2268–2292.

Graziano, M. S., Reiss, L. A., & Gross, C. G. (1999). A neuronal representation of the location of nearby sounds. *Nature*, *397*(6718), 428–430.

Gross, C. G., & Graziano, M. (1995). Multiple Representations of Space in the Brain. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *1*(1), 43–50.

Halligan, P. W., & Marshall, J. C. (1989). Is neglect (only) lateral? A quadrant analysis of line cancellation. *Journal of Clinical and Experimental Neuropsychology*, *11*(6), 793–798.

Halligan, P. W., & Marshall, J. C. (1991). Left neglect for near but not far space in man. *Nature*, *350*(6318), 498–500.

Halligan, P. W., Fink, G. R., Marshall, J. C., & Vallar, G. (2003). Spatial cognition: Evidence from visual neglect. *Trends in Cognitive Sciences*, *7*(3), 125–133.

Harris, L., Harrar, V., Jaekl, P., Kopinska, A., 2010. Mechanisms of simultaneity constancy. In: Nijhawan, R., Khurana, B. (Eds.), *Space and Time in Perception and Action* (232–253). Cambridge: Cambridge University Press.

Hecht, D., Reiner, M., & Karni, A. (2008). Multisensory enhancement: Gains in choice and in simple response times. *Experimental Brain Research*, *189*(2), 133–143.

Hediger, H. (1955). *Studies of the Psychology and Behavior of Animals in Zoos and Circuses*. New York: Criterion.

Heed, T., Habets, B., Sebanz, N., & Knoblich, G. (2010). Others' actions reduce crossmodal integration in peripersonal space. *Current Biology*, *20*(15), 1345-1349.

Heilman, K. M., Watson, R. T., & Valenstein, E. (2003). Neglect and related disorders. In K. M. Heilman & E. Valenstein (Eds.), *Clinical Neuropsychology* (4th ed., 296–346). New York, NY: Oxford University Press.

Heron, J., Whitaker, D., McGraw, P. V., & Horoshenkov, K. V. (2007). Adaptation minimizes distance-related audiovisual delays. *Journal of Vision*, *7*(13), 5.

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*(4), 1603–1616.

Hirsh, I. J., & Sherrick, C. E., Jr. (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology, 62*(5), 423-432.

Ho, C., & Spence, C. (2005). Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. *Journal of Experimental Psychology: Applied*, *11*(3), 157–174.

Ho, C., & Spence, C. (2006). Verbal interface design: Do verbal directional cues automatically orient visual spatial attention? *Computers in Human Behavior*, *22*(4), 733–748.

Ho, C., & Spence, C. (2009). Using peripersonal warning signals to orient a driver's gaze. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 51*, 539–556.

Ho, C., & Spence, C. (2014). Effectively responding to tactile stimulation: Do homologous cue and effector locations really matter? *Acta Psychologica*, *151*, 32–39.

Ho, C., Tan, H. Z., & Spence, C. (2005). Using spatial vibrotactile cues to direct visual attention in driving scenes. *Transportation Research. Part F: Traffic Psychology and Behaviour*, *8*(6), 397–412.

Ho, C., Tan, H. Z., & Spence, C. (2006). The differential effect of vibrotactile and auditory cues on visual spatial attention. *Ergonomics*, *49*(7), 724–738.

Holmes, N. P. (2007). The law of inverse effectiveness in neurons and behaviour: multisensory integration versus normal variability. *Neuropsychologia*, *45*(14), 3340-3345.

Holmes, N. P. (2009). The principle of inverse effectiveness in multisensory integration: some statistical considerations. *Brain Topography*, *21*(3-4), 168–176.

Holmes, N. P., & Spence, C. (2004). The body schema and multisensory representation (s) of peripersonal space. *Cognitive Processing*, *5*(2), 94-105.

Holmes, N. P., Calvert, G. A., & Spence, C. (2004). Extending or projecting peripersonal space with tools? Multisensory interactions highlight only the distal and proximal ends of tools. *Neuroscience Letters*, *372*(1), 62-67.

Holmes, N. P., Calvert, G. A., & Spence, C. (2007). Tool use changes multisensory interactions in seconds: Evidence from the crossmodal congruency task. *Experimental Brain Research*, *183*(4), 465-476.

Holmes, N. P., & Spence, C. (2005). Multisensory integration: space, time and superadditivity. *Current Biology*, *15*(18), R762-R764.

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.

Hughes, H. C., Reuter-Lorenz, P. A., Nozawa, G., & Fendrich, R. (1994). Visual-auditory interactions in sensorimotor processing: saccades versus manual responses. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(1), 131.

Ichikawa, M. (2009). Illusory temporal order for stimuli at different depth positions. *Attention, Perception, & Psychophysics*, *71*(3), 578-593.

Iriki, A., Tanaka, M., & Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport*, *7*(14), 2325-2330.

Jehkonen, M., Laihasalo, M., Koivisto, A.-M., Dastidar, P., & Ahonen, J.-P. (2007). Fluctuation in spontaneous recovery of left visual neglect: A 1-year follow-up. *European Neurology*, *58*, 210–214.

Jehkonen, M., Laihosalo, M., & Kettunen, J. E. (2006). Impact of neglect on functional outcome after stroke: A review of methodological issues and recent research findings. *Restorative Neurology and Neuroscience*, *24*(4), 209–215.

Johansson, B. B. (2012). Multisensory stimulation in stroke rehabilitation. *Frontiers in Human Neuroscience*, *6*(60).

Kandula, M., Hofman, D., & Dijkerman, H.C. (2015). Visuo-tactile interactions are dependent on the predictive value of the visual stimulus. *Neuropsychologia, 70*, 358–366.

Keetels, M., & Vroomen, J. (2005). The role of spatial disparity and hemifields in audio-visual temporal order judgments. *Experimental Brain Research*, *167*(4), 635–640.

Keller, I., Schindler, I., Kerkhoff, G., von Rosen, F., & Golz, D. (2005). Visuospatial neglect in near and far space: Dissociation between line bisection and letter cancellation. *Neuropsychologia*, *43*(5), 724–731.

Kennett, S., & Driver, J. (2014). Within-hemifield posture changes affect tactile–visual exogenous spatial cueing without spatial precision, especially in the dark. *Attention, Perception, & Psychophysics*, *76*(4), 1121-1135.

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.

Kim, M., Na, D. L., Kim, G. M., Adair, J. C., Lee, K. H., & Heilman, K. M. (1999). Ipsilesional neglect: Behavioural and anatomical features. *Journal of Neurology, Neurosurgery & Psychiatry*, *67*, 35–38.

Kim, Y., Gitelman, D. R., Nobre, A. C., Parrish, T. B., LaBar, K. S., & Mesulam, M. M. (1999). The large-scale neural network for spatial attention displays multifunctional overlap but differential asymmetry. *NeuroImage*, *9*, 269–277.

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.

Kingstone, A., & Pratt, J. (1999). Inhibition of return is composed of attentional and oculomotor processes. *Perception & Psychophysics*, *61*(6), 1046–1054.

Kitagawa, N., & Ichihara, S. (2002). Hearing visual motion in depth. *Nature*, *416*(6877), 172-174.

Kitagawa, N., & Spence, C. (2005). Investigating the effect of a transparent barrier on the crossmodal congruency effect. *Experimental Brain Research*, *161*(1), 62-71.

Kitagawa, N., Zampini, M., & Spence, C. (2005). Audiotactile interactions in near and far space. *Experimental Brain Research*, *166*(3-4), 528-537.

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

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

Kopinska, A., & Harris, L. R. (2004). Simultaneity constancy. *Perception*, *33*(9), 1049-1060.

Làdavas, E., Di Pellegrino, G., Farnè, A., & Zeloni, G. (1998). Neuropsychological evidence of an integrated visuotactile representation of peripersonal space in humans. *Journal of Cognitive Neuroscience*, *10*(5), 581-589.

Laplane, D., & Degos, J. D. (1983). Motor neglect. *Journal of Neurology, Neurosurgery & Psychiatry*, *46*(2), 152–158.

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.

Lee, C., Rohrer, W. H., & Sparks, D. L. (1988). Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature*, *332*(6162), 357–360.

Leibovitch, F. S., Black, S. E., Caldwell, C. B., Ebert, P. L., Ehrlich, L. E., & Szalai, J. P. (1998). Brain-behavior correlations in hemispatial neglect using CT and SPECT: The Sunnybrook stroke study. *Neurology*, *50*(4), 901–908.

Leicester, J., Sidman, M., Stoddard, L. T., & Mohr, J. P. (1969). Some determinants of visual neglect. *Journal of Neurology, Neurosurgery & Psychiatry*, *32*(6), 580–587.

Leo, F., Bolognini, N., Passamonit, C., Stein, B. E., & Ladavas, E. (2008). Cross-modal localization in hemianopia: New insights on multisensory integration. *Brain*, *131*(3), 855–865.

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

Lewald, J., & Guski, R. (2004). Auditory-visual temporal integration as a function of distance: no compensation for sound-transmission time in human perception. *Neuroscience letters*, *357*(2), 119-122.

Ling, S., & Carrasco, M. (2006). Sustained and transient covert attention enhance the signal via different contrast response functions. *Vision Research*, *46*(8-9), 1210–1220.

Ling, S., Liu, T., & Carrasco, M. (2009). How spatial and feature-based attention affect the gain and tuning of population responses. *Vision Research*, *49*(10), 1194–1204.

Lloyd, D. M. (2007). Spatial limits on referred touch to an alien limb may reflect boundaries of visuo-tactile peripersonal space surrounding the hand. *Brain and Cognition*, *64*(1), 104-109.

Longo, M. R., & Lourenco, S. F. (2006). On the nature of near space: Effects of tool use and the transition to far space. *Neuropsychologia*, *44*(6), 977-981.

Longo, M. R., & Lourenco, S. F. (2007). Space perception and body morphology: extent of near space scales with arm length. *Experimental Brain Research*, *177*(2), 285-290.

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.

Lourenco, S. F., Longo, M. R., & Pathman, T. (2011). Near space and its relation to claustrophobic fear. *Cognition*, *119*(3), 448-453.

Lueck, C. J., Crawford, T. J., Savage, C. J., & Kennard, C. (1990). Auditory-visual interaction in the generation of saccades in man. *Experimental Brain Research*, *82*(1), 149-157.

Luo, B., Shan, C., Zhu, R., Weng, X., & He, S. (2011). Functional foveal splitting: evidence from neuropsychological and multimodal MRI investigations in a Chinese patient with a splenium lesion. *PLoS ONE*, *6*(8), e23997.

Lupianez, J., Klein, R. M., & Bartolomeo, P. (2006). Inhibition of return: Twenty years after. *Cognitive Neuropsychology*, *23*(7), 1003–1014.

Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: A window onto functional integration in the human brain. *Trends in Neurosciences*, *28*, 264–271.

Macaluso, E., Frith, C. D., & Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, *289*(5482), 1206–1208.

Macaluso, E., Frith, C. D., & Driver, J. (2002). Supramodal effects of cover spatial orienting triggered by visual or tactile events. *Journal of Cognitive Neuroscience*, *14*, 389–401.

Machulla, T.-K., Di Luca, M., Froehlich, E., & Ernst, M. O. (2012). Multisensory simultaneity recalibration: storage of the aftereffect in the absence of counterevidence. *Experimental Brain* Research, *217*(1), 89–97.

Makin, T. R., Holmes, N. P., Brozzoli, C., Rossetti, Y., & Farnè, A. (2009). Coding of visual space during motor preparation: Approaching objects rapidly modulate corticospinal excitability in hand-centered coordinates. *The Journal of Neuroscience*, *29*(38), 11841-11851.

Makin, T. R., Holmes, N. P., & Ehrsson, H. H. (2008). On the other hand: dummy hands and peripersonal space. *Behavioural Brain Research*, *191*(1), 1-10.

Makin, T. R., Holmes, N. P., & Zohary, E. (2007). Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. *Journal of Neuroscience*, 27(3), 731–740.

Maravita, A., Husain, M., Clarke, K., & Driver, J. (2001). Reaching with a tool extends visual–tactile interactions into far space: Evidence from cross-modal extinction. *Neuropsychologia*, *39*(6), 580-585.

Maravita, A., Spence, C., Sergent, C., & Driver, J. (2002). Seeing your own touched hands in a mirror modulates cross-modal interactions. *Psychological Science*, *13*(4), 350-355.

Mays, L. E., & Sparks, D. L. (1980). Dissociation of visual and saccade-related responses in superior colliculus neurons. *Journal of Neurophysiology*, *43*(1), 207-232.

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.

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., & Ward, L. M. (2001). Multisensory integration and crossmodal attention effects in the human brain. Science, 292(5523), 1791.

Meador, K. J., Loring, D. W., Bowers, D., & Heilman, K. M. (1987). Remote memory and neglect syndrome. *Neurology*, *37*(3), 522.

Meng, F., & Spence, C. (2015). Tactile warning signals for in-vehicle systems. *Accident; Analysis and Prevention*, *75*, 333–346.

Mennemeier, M., Wertman, E., & Heilman, K. M. (1992). Neglect of near peripersonal space: Evidence for multidirectional systems in humans. *Brain*, *115*(1), 37–50.

Meredith, M. A., & Stein, B. E. (1983). Interactions among Converging Sensory Inputs in the Superior Colliculus. *Science*, *221*(4608), 389–391.

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

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.

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.

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

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

Miller, R. L., Pluta, S. R., Stein, B. E., & Rowland, B. A. (2015). Relative unisensory strength and timing predict their multisensory product. *Journal of Neuroscience*, *35*(13), 5213–5220.

Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia*, *46*, 774–785.

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

Montagna, B., Pestilli, F., & Carrasco, M. (2009). Attention trades off spatial acuity. *Vision Research*, *49*(7), 735–745.

Motulsky, H. J. (1995). *Intuitive biostatistics*. Oxford, UK: Oxford University Press.

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.

Murray, M., Spence, C., & Harris, L. (2013). International Multisensory Research Forum 2012 Meeting Special Issue. *Multisensory research*, *26*(3), 287-289.

Navarra, J., Alsius, A., Soto-Faraco, S., & Spence, C. (2010). Assessing the role of attention in the audiovisual integration of speech. *Information Fusion*, *11*(1), 4–11.

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

Nijboer, T. C., Kollen, B. J., & Kwakkel, G. (2013). Time course of visuospatial neglect early after stroke: a longitudinal cohort study. *Cortex*, *49*(8), 2021-2027.

Nijboer, T. C. W., Van de Port, I., Schepers, V., Post, M., & Visser-Meiley, A. (2013). Predicting functional outcome after stroke: The influence of neglect on basic activities in daily living. *Frontiers in Human Neuroscience, 7*(182), 1–6.

Occelli, V., Hartcher-O'Brien, J. H., Spence, C., & Zampini, M. (2010). Assessing the audiotactile Colavita effect in near and rear space. *Experimental Brain Research*, *203*(3), 517-532.

Occelli, V., Spence, C., & Zampini, M. (2011). Audiotactile interactions in front and rear space. *Neuroscience & Biobehavioral Reviews*, *35*(3), 589-598.

Ottes, F. P., Van Gisbergen, J. A., & Eggermont, J. J. (1985). Latency dependence of colour-based target vs nontarget discrimination by the saccadic system. *Vision research*, *25*(6), 849-862.

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. *The Journal of Neuroscience*, *33*(17), 7463–7474.

Pegna, A. J., Petit, L., Caldara-Schnetzer, A. S., Khateb, A., Annoni, J. M., Sztajzel, R., & Landis, T. (2001). So near yet so far: neglect in far or near space depends on tool use. *Annals of Neurology*, *50*(6), 820–822.

Pierson-Savage, J. M., Bradshaw, J. L., Bradshaw, J. L., & Nettleton, N. C. (1988). Vibrotactile reaction times in unilateral neglect. *Brain*, *111*(6), 1531–1545.

Pitts, W., & McCulloch, W. S. (1947). How we know universals the perception of auditory and visual forms. *The Bulletin of Mathematical Biophysics*, *9*(3), 127-147.

Pitzalis, S., Di Russo, F., Spinelli, D., & Zoccolotti, P. (2001). Influence of the radial and vertical dimensions on lateral neglect. *Experimental Brain Research*, *136*(3), 281–294.

Pitzalis, S., Spinelli, D., & Zoccolotti, P. (1997). Vertical neglect: Behavioral and electrophysiological data. *Cortex*, *33*(4), 679–688.

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

Previc, F. H. (1998). The neuropsychology of 3-D space. *Psychological Bulletin*, *124*(2), 132–164.

Previc, F. H. (1990). Functional specialization in the lower and upper visual fields in humans: Its ecological origins and neurophysiological implications. *Behavioral and Brain Sciences*, *13*(3), 519–575.

Previc, F. H. (2000). Neuropsychological guidelines for aircraft control stations. *Engineering in Medicine and Biology Magazine, IEEE*, *19*(2), 81-88.

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

Ramachandran, V. S., Altschuler, E. L., Stone, L., Al-Aboudi, M., Schwartz, E., & Siva, N. (1999). Can mirrors alleviate visual hemineglect? *Medical hypotheses*, *52*(4), 303-305.

Rapcsak, S. Z., Cimino, C. R., & Heilman, K. M. (1988). Altitudinal neglect. *Neurology*, *38*(2), 277–281.

Reches, A., Netser, S., & Gutfreund, Y. (2010). Interactions between stimulus-specific adaptation and visual auditory integration in the forebrain of the barn owl. *The Journal of Neuroscience*, *30*(20), 6991–6998.

Ringman, J. M., Saver, J. L., Woolson, R. F., Clarke, W. R., & Adams, H. P. (2004). Frequency, risk factors, anatomy, and course of unilateral neglect in an acute stroke cohort. *Neurology*, *63*(3), 468–474.

Rizzolatti, G., Scandolara, C., Matelli, M., & Gentilucci, M. (1981). Afferent properties of periarcuate neurons in macaque monkeys. II. Visual responses. *Behavioural Brain Research*, *2*(2), 147-163.

Robertson, I. H., & Halligan, P. W. (1999). *Spatial neglect: A clinical handbook for diagnosis and treatment*. Hove: Psychology Press.

Rorden, C., & Karnath, H. O. (2010). A simple measure of neglect severity. *Neuropsychologia*, *48*(9), 2758–2763.

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex*, *17*(5), 1147-1153.

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.

Sambo, C. F., & Forster, B. (2011). When far is near: ERP correlates of crossmodal spatial interactions between tactile and mirror-reflected visual stimuli. *Neuroscience letters*, *500*(1), 10-15.

Sambo, C. F., & Iannetti, G. D. (2013). Better safe than sorry? The safety margin surrounding the body is increased by anxiety. *The Journal of Neuroscience*, *33*(35), 14225-14230.

Santangelo, V., & Spence, C. (2009). Crossmodal exogenous orienting improves the accuracy of temporal order judgments. *Experimental Brain Research*, *194*(4), 577–586.

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

Santangelo, V., & Spence, C. (2007). Multisensory cues capture spatial attention regardless of perceptual load. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(6), 1311.

Santangelo, V., Van der Lubbe, R. H., Belardinelli, M. O., & Postma, A. (2006). Spatial attention triggered by unimodal, crossmodal, and bimodal exogenous cues: A comparison of reflexive orienting mechanisms. *Experimental Brain Research*, *173*(1), 40–48.

Santangelo, V., Van der Lubbe, R. H., Belardinelli, M. O., & Postma, A. (2008). Multisensory integration affects ERP components elicited by exogenous cues. *Experimental Brain Research*, *185*(2), 269–277.

Schicke, T., Bauer, F., & Röder, B. (2009). Interactions of different body parts in peripersonal space: how vision of the foot influences tactile perception at the hand. *Experimental Brain Research*, *192*(4), 703-715.

Schubert, F., & Spatt, J. (2001). Double dissociations between neglect tasks: Possible relation to lesion site. *European Neurology*, *45*(3), 160–164.

Serino, A., Canzoneri, E., & Avenanti, A. (2011). Fronto-parietal areas necessary for a multisensory representation of peripersonal space in humans: an rTMS study. *Journal of Cognitive Neuroscience*, *23*(10), 2956-2967.

Shulman, G. L., Pope, D. L. W., Astafiev, S. V., McAvoy, M. P., Snyder, A. Z., & Corbetta, M. (2010). Right hemisphere dominance, during spatial selective attention and target detection occurs outside the dorsal frontroparietal network. *Journal of Neuroscience*, *30*(10), 3640–3651.

Silva, C. C., Mendonça, C., Mouta, S., Silva, R., Campos, J. C., & Santos, J. (2013). Depth cues and perceived audiovisual synchrony of biological motion. *PLoS ONE*, *8*(11), e80096.

Sommer, R., (1959). Studies in personal space. *Sociometry*, *22*, 247–260.

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.

Sparks, D. L., & Nelson, I. S. (1987). Sensory and motor maps in the mammalian superior colliculus. *Trends in Neurosciences*, *10*(8), 312-317.

Spence, C. (2007). Audiovisual multisensory integration. *Acoustical Science and Technology*, *28*(2), 61-70.

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

Spence, C. (2011). Assessing the consequences of tool-use for the representation of peripersonal space in humans. In: McCormack, T., Hoerl, C., Butterfill, S. (Eds.), *Tool Use and Causal Cognition* (220-247). Oxford, UK: Oxford University Press.

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. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(3), 555-574.

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, *59*(1), 1–22.

Spence, C, & Driver, J. (1998a). Auditory and audiovisual inhibition of return. *Perception & Psychophysics*, *60*(1), 125–139.

Spence, C, & Driver, J. (1998b). Inhibition of return following an auditory cue: The role of central reorienting events. *Experimental Brain Research*, *118*(3), 352–360.

Spence, C, & Driver, J. (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., & Ho, C. (2008). Multisensory warning signals for event perception and safe driving. *Theoretical Issues in Ergonomics Science*, *9*(6), 523–554.

Spence, C., Lloyd, D., McGlone, F., Nicholls, M. E., & Driver, J. (2000). Inhibition of return is supramodal: a demonstration between all possible pairings of vision, touch, and audition. *Experimental Brain Research*, *134*(1), 42-48.

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., Nicholls, M. E., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, *63*(2), 330–336.

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*(4), 544–557.

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

Spence, C., & Squire, S. (2003). Multisensory integration: Maintaining the perception of synchrony. *Current Biology*, *13*(13), R519-R521.

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

Spence, C., Parise, C., & Chen, Y.-C. (2011). The Colavita visual dominance effect. In: Murray, M.M., Wallace, M. (Eds.), *Frontiers in the Neural Bases of Multisensory Processes* (523-550). Boca Raton, FL: CRC Press.

Spence, C., & Santangelo, V. (2009). Capturing spatial attention with multisensory cues: A review. *Hearing Research*, *258*(1-2), 134–142.

Spence, C., & Squire, S.B. (2003). Multisensory integration: Maintaining the perception of synchrony. *Current Biology*, *13*(13), R519–R521.

Stein, B. E., Burr, D., Constantinidis, C., Laurienti, P. J., Alex Meredith, M., Perrault, T. J., ... & Lewkowicz, D. J. (2010). Semantic confusion regarding the development of multisensory integration: A practical solution. *European Journal of Neuroscience*, *31*(10), 1713-1720.

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: The 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.

Stein, B. E., Stanford, T. R., Wallace, M. T., Vaughan, J. W., & Jiang, W. (2004). Crossmodal spatial interactions in subcortical and cortical circuits. In: Spence, C., Driver, J. (Eds.), *Crossmodal Space and Crossmodal Attention* (25-50). Oxford, UK: Oxford University Press.

Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: Task and stimulus dependencies. *Experimental Brain Research*, *227*(2), 249–261.

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.

Stevenson, R. A., Krueger Fister, J., 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.

Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., ... & Porter, N. R. (2001). When is now? Perception of simultaneity. *Proceedings of the Royal Society of London B*, *268*(1462), 31-38.

Stone, S. P., Wilson, B., Wroot, A., Halligan, P. W., Lange, L. S., Marshall, J. C., & Greenwood, R. J. (1991). The assessment of visuo-spatial neglect after acute stroke. *Journal of Neurology, Neurosurgery & Psychiatry*, *54*, 345–350.

Suchoff, I. B., Kapoor, N., Ciuffreda, K. J., Rutner, D., Han, E., & Craig, S. (2008). The frequency of occur- rence, types, and characteristics of visual field defects in acquired brain injury: A retrospective analysis. *Optometry*, *79*(5), 259–256.

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature*, *421*(6926), 911.

Tajadura-Jiménez, A., Kitagawa, N., Väljamäe, A., Zampini, M., Murray, M. M., & Spence, C. (2009). Auditory–somatosensory multisensory interactions are spatially modulated by stimulated body surface and acoustic spectra. *Neuropsychologia*, *47*(1), 195-203.

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

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

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

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

Taylor, T. L., Klein, R. M., & Munoz, D. P. (1999). Saccadic performance as a function of the presence and disappearance of auditory and visual fixation stimuli. *Journal of Cognitive Neuroscience*, *11*(2), 206-213.

Ten Brink, A. F., Nijboer, T. C. W., Van der Stoep, N., & Van der Stigchel, S. (2014). The influence of vertically and horizontally aligned visual distractors on aurally guided saccadic eye movements. *Experimental Brain Research*, *232*(4), 1357-1366.

Teneggi, C., Canzoneri, E., Di Pellegrino, G., & Serino, A. (2013). Social modulation of peripersonal space boundaries. *Current Biology*, *23*(5), 406-411.

Teramoto, W., Nozoe, Y., & Sekiyama, K. (2013). Audiotactile interactions beyond the space and body parts around the head. *Experimental Brain Research*, *228*(4), 427-436.

Theeuwes, J., & Pratt, J. (2003). Inhibition of return spreads across 3-D space. *Psychonomic Bulletin & Review*, *10*(3), 616-620.

Tipper, S. P., Howard, L. A., & Jackson, S. R. (1997). Selective Reaching to Grasp: Evidence for Distractor Interference Effects. *Visual Cognition*, *4*(1), 1–38.

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

Van der Stigchel, S. (2010). Recent advances in the study of saccade trajectory deviations. *Vision Research*, *50*(17), 1619–1627.

Van der Stigchel, S., Heeman, J., & Nijboer, T. C. W. (2012). Averaging is not everything: The saccade global effect weakens with increasing stimulus size. *Vision Research*, *62*, 108-115.

Van der Stigchel, S., & Nijboer, T. C. W. (2011). The global effect: What determines where the eyes land. *Journal of Eye Movement Research*, *4*(2), 1-13.

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.

Van der Stoep, N., Van der Stigchel, S., Nijboer, T. C. W., & Van der Smagt, M. J. (in press). Audiovisual integration in near and far space: Effects of changes in distance and stimulus effectiveness. *Experimental Brain Research*. DOI: 10.1007/s00221-015-4248-2

Van der Stoep, N., Van Engelen, R. C., Van der Stigchel, S., & Nijboer, T. C. W. (in prep.) Differences in multisensory integration in the left and right visual field after temporal-parietal lesions.

Van der Stoep, N., Visser-Meily, J. M., Kappelle, L. J., de Kort, P. L., Huisman, K. D., Eijsackers, A. L., ... & Nijboer, T. C. (2013). Exploring near and far regions of space: distance-specific visuospatial neglect after stroke. *Journal of Clinical and Experimental Neuropsychology*, *35*(8), 799-811.

Van Zoest, W., Donk, M., & Theeuwes, J. (2004). The role of stimulus-driven and goal-driven control in saccadic visual selection. *Journal of Experimental Psychology: Human perception and performance*, *30*(4), 746.

Verdon, V., Schwartz, S., Lovblad, K. O., Hauert, C. A., & Vuilleumier, P. (2010). Neuroanatomy of hemispatial neglect and its functional components: A study using voxel-based lesion–symptom mapping. *Brain*, *133*(3), 880–894.

Viaud‐Delmon, I., Brugger, P., & Landis, T. (2007). Hemineglect: Take a look at the back space. *Annals of neurology*, *62*(4), 418-422.

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., & De Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1583.

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

Vroomen, J., Keetels, M., De Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive brain research*, *22*(1), 32-35.

Vuilleumier, P., Schwartz, S., Verdon, V., Maravita, A., Hutton, C., Husain, M., & Driver, J. (2008). Abnormal attentional modulation of retinotopic cor- tex in parietal patients with spatial neglect. *Current Biology*, *18*(19), 1525–1529.

Vuilleumier, P., Valenza, N., Mayer, E., Reverdin, A., & Landis, T. (1997). Near and far visual space in unilateral neglect. *Annals of Neurology*, *43*(3), 406–410.

Wallace, M. T., & Stein, B. E. (1997). Development of multisensory neurons and multisensory integration in cat superior colliculus. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *17*(7), 2429–2444.

Wallace, M. T., & Stein, B. E. (2001). Sensory and multisensory responses in the newborn monkey superior colliculus. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *21*(22), 8886–8894.

Wallace, M. T., & Stein, B. E. (2007). Early experience determines how the senses will interact. *Journal of Neurophysiology*, *97*(1), 921–926.

Wan, X., Spence, C., Mu, B., Zhou, X., & Ho, C. (2014). Assessing the benefits of multisensory audiotactile stimulation for overweight individuals. *Experimental Brain Research*, *232*(4), 1085-1093.

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

Warren, R. M. (1963) Reply to S. S. Stevens. *Journal of the Acoustical Society of America*, *35*, 1663–1665 .

Weintraub, S., & Mesulam, M. M. (1988). Visual hemispatial inattention: Stimulus parameters and exploratory strategies. *Journal of Neurology, Neurosurgery & Psychiatry*, *51*(12), 1481–1488.

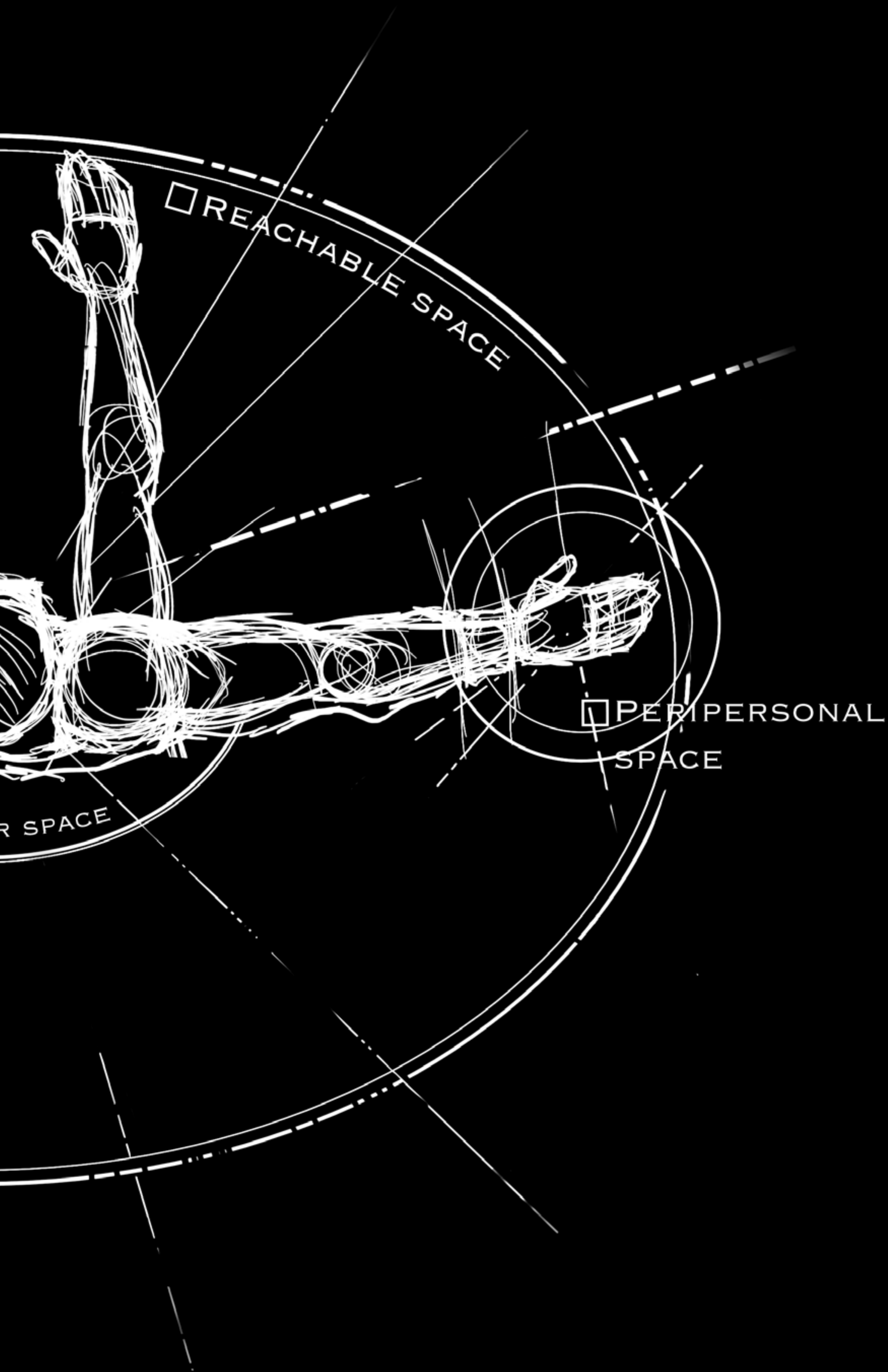
Zambarbieri, D., Schmid, R., Magenes, G., & Prablanc, C. (1982). Saccadic responses evoked by presentation of visual and auditory targets. *Experimental Brain Research*, *47*(3), 417-427.

Zampini, M., Guest, S., Shore, D. I., & Spence, C. (2005). Audio-visual simultaneity judgments. Perception & Psychophysics, 67(3), 531– 544.

Zampini, M., Torresan, D., Spence, C., & Murray, M. M. (2007). Auditory–somatosensory multisensory interactions in front and rear space. *Neuropsychologia*, *45*(8), 1869–1877.

Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika*, *70*(1), 203-212.

Zou, H., Müller, H. J., & Shi, Z. (2012). Non-spatial sounds regulate eye movements and enhance visual search. *Journal of Vision*, *12*(5), 1-18.


☐Reachable space
☐Peripersonal space
r space

Part 3

Appendix

Nederlandse samenvatting

Iedere dag moet ons brein een enorme hoeveelheid aan zintuigelijke informatie uit de omgeving verwerken. Twee processen die helpen bij het verwerken van al deze informatie en die centraal staan in dit proefschrift zijn reflexieve ruimtelijke aandachtverschuivingen tussen zintuigen en multisensorische integratie. Het doel van dit proefschrift was om de volgende vragen te beantwoorden: (1) hoe dragen reflexieve aandacht en multisensorische integratie bij aan verbeteringen in onze waarneming, en (2) hoe beïnvloedt de afstand waarop licht en/of geluid wordt waargenomen reflexieve aandacht en multisensorische integratie. Dit laatste punt is met name interessant omdat er uit verschillende onderzoeken is gebleken dat het brein informatie van ver weg en dichtbij apart lijkt te verwerken. Eerst volgt een korte uitleg van de twee processen die centraal staan in dit proefschrift. Daarna wordt er een beknopt overzicht gegeven van de bevindingen van de studies die beschreven zijn in dit proefschrift.

Reflexieve ruimtelijke aandachtverschuivingen tussen zintuigen

Aandacht is belangrijk voor het filteren van zintuigelijke informatie. Wanneer we ergens geen aandacht voor hebben, nemen we het niet of nauwelijks waar. Aandacht zorgt voor het selecteren van relevante informatie. We kunnen onze aandacht vrijwillig op iets in de omgeving richten. Dit heet endogene ruimtelijke aandacht. Geluid, licht, en aanrakingen kunnen onze aandacht trekken. Deze vorm van aandacht heet reflexieve of exogene aandacht. Onze aandacht kan bijvoorbeeld getrokken worden naar de locatie van een geluid dat plotseling van rechts komt. We zijn dan sneller in het verwerken van en gevoeliger voor, onder andere, visuele informatie die een moment later op diezelfde plek verschijnt. Iets dat automatische de aandacht trekt, bijvoorbeeld een licht of een geluid, heet een exogene cue. Het idee is dat deze exogene cue de aandacht trekt naar de locatie waar deze verschijnt en dat deze aandachtverschuiving vervolgens de (visuele) informatie verwerking op die locatie verbetert. Dit effect van ruimtelijke aandachtverschuivingen tussen zintuigen is inmiddels gedemonstreerd voor alle paren van visuele, auditieve, en tactiele informatie. Deze interactie tussen de zintuigen maakt onze waarneming efficiënter door onze aandacht naar nieuwe en opvallende gebeurtenissen te trekken en deze vervolgens beter waar te nemen. Het effect van deze reflexieve aandachtverschuivingen is het grootst wanneer er enige tijd tussen de twee signalen zit (100-300 ms). Als de tijd tussen de signalen langer is dan keert dit effect om (>300 ms): we zijn dan trager in het reageren op een lichtsignaal als daarvoor op dezelfde plek een geluid is aangeboden. Dit effect heet inhibitie van terugkeer en wordt vaak geïnterpreteerd in het kader van efficiënt zoekgedrag. Als er binnen 300 ms niets nieuws meer verschijnt op de plek waar onze aandacht naar toe is gegaan, dan zou het efficiënter kunnen zijn om de aandacht op een nieuwe plek in de ruimte te richten.

Multisensorische integratie

Wanneer licht en geluid op hetzelfde moment op dezelfde locatie verschijnen, dan worden deze signalen door het brein gecombineerd tot één veel sterker signaal. Dit proces heet multisensorische integratie en zorgt ervoor dat we de signalen sneller kunnen lokaliseren en identificeren. Deze integratie wordt mogelijk gemaakt door zenuwcellen in het brein die reageren op zowel licht als geluidsignalen. Wanneer licht en geluid samen worden aanboden op deze locatie dan is de reactie in een subgroep van deze cellen vele malen sterker dan de som van de activiteit als reactie op alleen licht of alleen geluid. Hoeveel sterker de activiteit in deze cel is, hangt af van hoe sterk de signalen zijn. Op celniveau is multisensorische integratie sterker wanneer de signalen zwak zijn dan wanneer de signalen sterk zijn. Dit heet het principe van omgekeerde effectiviteit.

Ruimtelijke aandacht en multisensorische integratie

Reflexieve aandachtverschuivingen tussen zintuigen en multisensorische integratie leiden beiden tot het sneller detecteren, lokaliseren, en identificeren van informatie in de omgeving (licht, geluid). We hebben daarom onderzocht onder welke omstandigheden elk proces bijdraagt aan een verbetering van onze waarneming. Dit bleek met name af te hangen van de tijd tussen het verschijnen van de licht en geluid signalen. Wanneer licht en geluid signalen tegelijkertijd worden aangeboden zorgt multisensorische integratie voor een verbetering van onze waarneming. Als er ongeveer 50 ms tussen de signalen zit, dragen zowel multisensorische integratie als reflexieve aandachtverschuivingen tussen zintuigen bij aan een verbetering van onze waarneming. Aandacht, maar niet multisensorische integratie, speelt een rol in de verbetering van onze waarneming als er 100-200 ms tussen de signalen zit. De grootste verbetering werd altijd gezien wanneer licht en geluid op dezelfde locatie werden aangeboden.

Onderzoekers hebben zich afgevraagd of er aandacht nodig is voor multisensorische integratie. Verschillende onderzoeken hebben laten zien dat multisensorische integratie eerder optreedt dan ruimtelijke aandacht verschuivingen tussen zintuigen, maar ook dat selectieve aandacht integratie kan versterken of kan verzwakken. Er was echter weinig bekend over hoe reflexieve aandachtverschuivingen multisensorische integratie beïnvloeden. We hebben dit onderzocht omdat beide processen erg belangrijk zijn voor snelle ruimtelijke lokalisatie. Wanneer een geluid de aandacht naar een locatie trekt, is de integratie van licht en geluid dat een moment later op die locatie verschijnt minder sterk dan wanneer deze signalen op een andere locatie verschijnen. Reflexieve ruimtelijke aandacht kan multisensorische integratie dus beïnvloeden. Een mogelijke verklaring voor dit effect is dat aandacht de signalen die op dezelfde locatie verschijnen zo versterken dat multisensorische integratie afneemt volgens het principe van omgekeerde effectiviteit.

We hebben ook onderzocht hoe inhibitie van terugkeer naar een locatie de integratie van zintuigelijke informatie beïnvloedt. We trokken de aandacht van proefpersonen naar een bepaalde locatie met een visuele cue en onderzochten hoe dit de verwerking van licht, geluid, en de combinatie van licht en geluid beïnvloedde die 300 ms later werd aangeboden. Met name *visuele* informatie werd trager verwerkt op de locatie van de cue ten opzichte van andere locaties. Dit is het bovengenoemde inhibitie van terugkeer effect. Dit effect was echter niet aanwezig wanneer proefpersonen moesten reageren op een *geluid* dat 300 ms na de cue werd aangeboden. Integendeel, proefpersonen reageerden sneller op geluiden die op dezelfde locatie als de cue werden aangeboden ten opzichte van andere locaties. De integratie van licht en geluid op de locatie waar eerder een cue verscheen was zwakker in vergelijking met andere locaties. Inhibitie van terugkeer voor visuele informatie zorgde ervoor dat de verwerking van visuele informatie werd vertraagd maar de verwerking van auditieve informatie werd juist versneld. Dit zorgde ervoor dat auditieve informatie (geluid) dominant werd en integratie verminderde op de locatie van de cue.

In bovengenoemde onderzoeken verscheen er per moment maar één lichtsignaal. In het dagelijks leven kunnen er uiteraard meerdere lichtsignalen tegelijkertijd verschijnen. We weten uit eerder onderzoek dat het brein een oogbeweging programmeert naar ieder element in een scene. Er kan natuurlijk maar één oogbeweging op een moment gemaakt worden dus er is competitie tussen de geprogrammeerde oogbewegingen. Onder andere onder invloed van aandacht kan er een keuze gemaakt worden tussen de geprogrammeerde oogbewegingen, en kan er een oogbeweging uitgevoerd worden naar het meest relevante object. Echter, wanneer er meerdere elementen dicht bij elkaar verschijnen maken we vaak eerst een oogbeweging naar wat ongeveer het centrum van deze groep elementen is. Dit terwijl hier niks te zien is! Het idee hierachter is dat de competitie tussen de oogbewegingsprogramma's nog niet opgelost is en dat de oogbeweging die uitgevoerd wordt hierdoor naar het gemiddelde van alle programma's gaat. We hebben onderzocht hoe het brein in dit geval om gaat met de aanwezigheid van geluid. Wanneer er een geluid via een koptelefoon werd afgespeeld op het moment dat er twee cirkels dicht bij elkaar werden aangeboden op een scherm, landde de eerste oogbeweging van proefpersonen nog dichter bij het midden van deze cirkels dan zonder het geluid. Hieruit concludeerden we dat zolang het geluid geen duidelijke locatie aangeeft de competitie tussen de oogbewegingsprogramma's niet opgelost kan worden. Het geluid wordt dan aan beide programma's gekoppeld en deze worden dus beiden versterkt. Hierdoor wordt de gemiddelde oogbeweging ook sterker en landen er gemiddeld meer oogbewegingen precies in het midden van de twee cirkels.

Ruimtelijke aandacht en multisensorische integratie op verschillende afstanden

De studies die tot nu toe genoemd zijn hebben bijgedragen aan een beter begrip van de omstandigheden waarin informatie van meerdere zintuigen tot een verbetering van de waarneming leidt door reflexieve ruimtelijk aandacht en multisensorische integratie. De signalen die gebruikt werden in deze onderzoeken werden echter altijd op dezelfde afstand van de proefpersoon aangeboden. In het dagelijks leven nemen we echter informatie van verschillende afstanden waar. In deze sectie zullen onderzoeken besproken worden waarin werd gekeken hoe het aanbieden van informatie op verschillende afstanden de waarneming beïnvloedt.

In de afgelopen decennia is er veel onderzoek gedaan naar multisensorische integratie. Er is echter nauwelijks rekening gehouden met het feit dat de afstand waarop informatie wordt waargenomen invloed kan hebben op onze multisensorische waarneming. Uit onderzoek bij apen weten we dat er multisensorische neuronen zijn die reageren op licht, geluid, en aanraking van bijvoorbeeld de hand. Echter alleen wanneer deze signalen ook op ongeveer dezelfde afstand van het lichaam worden aangeboden worden ze geïntegreerd tot een sterker signaal. Zo zal een geluid dat dichtbij de hand wordt aangeboden geïntegreerd worden met een aanraking van de hand, maar een geluid dat ver weg van de hand wordt aangeboden niet. De ruimte rondom een lichaamsdeel waarin verschillende signalen wordt geïntegreerd, heet de peripersoonlijke ruimte. De ruimte daarbuiten heet de extrapersoonlijke ruimte. Aangezien tast gebonden is aan het lichaam is onze ruimtelijke waarneming door middel van tast relatief beperkt. Geluid en licht moeten dus altijd dichtbij het lichaam worden aangeboden om geïntegreerd te worden met aanraking. Licht en geluid signalen kunnen echter op allerlei afstanden worden waargenomen. Het waarnemen van licht en geluid is dus niet beperkt door de afstand tot het lichaam. De vraag is dan ook of de afstand waarop licht en geluid signalen worden aangeboden invloed heeft op de mate waarin ze onze waarneming beïnvloeden. Een van de ideeën over onze ruimtelijke waarneming is dat verschillende combinaties van zintuigen meer of minder van belang zijn op verschillende afstanden.

Zo is tastzin en visuele informatieverwerking belangrijk voor het grijpen en interacteren met objecten in de omgeving, terwijl gehoor en zicht met name van belang zijn op grotere afstanden voor het navigeren en oriënteren.

We hebben onderzocht of diepte-informatie van belang is tijdens reflexieve ruimtelijke aandachtverschuivingen tussen zintuigen. Proefpersonen moesten zo snel mogelijk reageren op een stip die links of rechts dichtbij of ver weg werd aangeboden. Ongeveer 200 ms voor het verschijnen van de stip werd er een geluid aangeboden op dezelfde of een andere diepte (dichtbij versus ver weg) en aan dezelfde of een andere kant (links of rechts). Proefpersonen reageerden sneller op de stip wanneer deze aan dezelfde kant als het geluid verscheen (bijvoorbeeld beiden aan de linker kant) dan wanneer de stip aan de andere kant als het geluid verscheen (geluid links, stip rechts). Dit effect was echter alleen aanwezig wanneer het geluid en de stip op dezelfde afstand werden aangeboden (beiden ver weg of beiden dichtbij). Deze resultaten bieden ondersteuning voor het idee dat de reflexieve ruimtelijke aandacht verschuivingen gevoelig zijn voor de afstand waarop informatie wordt aangeboden. Een geluid ver weg trekt dus de aandacht ver weg en verbetert visuele informatieverwerking op die afstand, maar niet op andere afstanden.

Ruimtelijke aandacht kan ook aangedaan zijn na schade aan de hersenen terwijl de visuele informatieverwerking nog intact is. Bij visueel ruimtelijk neglect negeren mensen vaak de linkerkant van de ruimte en zijn ze erg gefocust op de rechterkant. In één van onze studies hebben we onderzocht of ruimtelijke aandacht ook specifiek aangedaan kan zijn voor een bepaalde afstand. Uit deze studie bleek dat ongeveer ongeveer 22% alleen neglect voor visuele informatie dichtbij hadden, 11% alleen voor visuele informatie ver weg, en 14% neglect voor informatie zowel dichtbij als ver weg. Deze resultaten bieden niet alleen ondersteuning voor de idee dat diepte-informatie een belangrijke rol speelt in ruimtelijke aandacht, maar lijkt er ook op te wijzen dat ruimtelijke aandacht voor de peripersoonlijke ruimte en de extrapersoonlijke ruimte door verschillende hersengebieden mogelijk gemaakt worden.

Zoals eerder aangegeven worden licht en geluid signalen geïntegreerd als ze van ongeveer dezelfde locatie en op ongeveer hetzelfde moment worden aangeboden. Gehoor en zicht zijn relevanter op grotere dan op kleinere afstanden vanwege hun belang voor het navigeren en lokaliseren van informatie verder weg. Als dit inderdaad zo is, dan zou multisensorische integratie van licht en geluid sterker kunnen zijn op grote afstanden dan op kleine afstanden. Wanneer we iets van een grotere afstand waarnemen is het veelal moeilijker te zien en klinkt het minder hard dan wanneer we het van dichtbij waarnemen. Er zijn echter twee factoren die veranderen als de afstand toeneemt: de effectiviteit van licht en geluid, en de afstand en dus de regio van waar de signalen worden aangeboden (peripersoonlijk vs. extrapersoonlijk). Door deze factoren individueel en samen te variëren konden we achterhalen wanneer multisensorische integratie het sterkst is. Op basis van de resultaten konden we concluderen dat multisensorische integratie van licht en geluid sterker is wanneer deze informatie buiten in plaats van binnen handbereik wordt aangeboden. Zowel een afname in de effectiviteit van de signalen als een toename in de afstand van de proefpersoon droegen bij aan een sterkere integratie van licht en geluid buiten handbereik. Er werd namelijk geen toename in integratie gevonden wanneer de signalen ver weg en dichtbij even effectief werden gemaakt. Ook was er geen toename in integratie wanneer de signalen binnen handbereik afnamen in effectiviteit.

Op basis van de studies in dit proefschrift hebben we laten zien dat onze waarneming verbeterd kan worden door twee processen waarbij wat we zien en wat we horen elkaar beïnvloeden: reflexieve ruimtelijke aandachtverschuivingen tussen zintuigen en multisensorische integratie. Afhankelijk van de tijd tussen de licht en geluid signalen draagt het een of het andere proces meer bij aan een verbetering in de waarneming. Ruimtelijke aandachtverschuivingen worden ook gemaakt in de diepte-dimensie en ruimtelijke aandacht kan aangedaan zijn in een specifieke regio in de ruimte (peripersoonlijke vs. extrapersoonlijke ruimte). Integratie van geluid of licht met tactiele stimulatie is sterker in de peripersoonlijke ruimte, terwijl integratie van geluid en licht sterker lijkt te zijn op grotere afstanden (in de extrapersoonlijke ruimte). Aangezien de resultaten aangeven onder welke omstandigheden lokalisatie het snelst is in de diepte-dimensie en de horizontale ruimte, is het mogelijk om optimale signalen te ontwerpen voor situaties waarin de snelheid van reageren erg belangrijk is. Zo kan deze informatie bijvoorbeeld worden gebruikt voor het ontwerpen van waarschuwingssignalen in auto's en cockpits. De snelheid van voortbeweging in auto's en vliegtuigen is erg hoog. Elke milliseconde is dus van belang bij het reageren op gevaar. Het is daarom cruciaal dat er snel en accuraat gehandeld kan worden als er bijvoorbeeld een obstakel op de route wordt waargenomen. Multisensorische waarschuwingssignalen die van dezelfde afstand lijken te komen als het naderende obstakel kunnen hier wellicht zeer effectief zijn.

Dankwoord

Er is veel gebeurd in de afgelopen vier jaar. Ik ben ontzettend dankbaar voor alle mensen die ik heb mogen leren kennen en waar ik van heb mogen leren. Er zijn veel mensen die een belangrijke rol gespeeld hebben tijdens mijn promoveren, in de totstandkoming van dit proefschrift, en in mijn ontwikkeling als onderzoeker en persoon.

Albert. Stille wateren, diepe gronden. Je enthousiasme voor onderzoek en je (veelal) goede humeur doen me altijd goed. Bedankt voor je adviezen en dat je me hebt aangemoedigd om naar Oxford te gaan.

Stefan en Tanja. Bedankt voor jullie begeleiding, jullie steun, betrokkenheid, sturing, werkhouding, snelheid en scherpte in denken, (snelle!) feedback, en vertrouwen in mijn kunnen. Bedankt ook voor de waardevolle kijk in jullie levens. Ik kon altijd bij jullie terecht. Fijn dat we zo open kunnen communiceren. Tanja, ik heb veel geleerd van je klinische kijk op onderzoek.

Martine, Esther, Carla. Ontzettend bedankt voor jullie begeleiding en de vele lessen in de wereld van de neuropsychologie. Ik ben blij met de ervaring die ik heb opgedaan bij jullie.

Anne. Je hebt me geleerd om de klinische relevantie van onderzoek te begrijpen, uit te leggen, en in het oog te houden. Dank daarvoor.

Charles. Thank you for the great time in the Crossmodal Research Lab in Oxford. Thanks for all your feedback, the opportunities, and the continuous support you have given me. You have made me laugh more than once when I had trouble seeing the forest through the trees. Thanks also go to all the wonderful colleagues and friends from the CRL.

Max. It was great meeting you at my first IMRF in Jerusalem in 2013. Thanks for letting me spent some time in your lab in Birmingham.

Maarten. Bedankt voor de leuke discussies, brainstorm sessies, en goede feedback. Ik heb me altijd welkom gevoeld en geen vraag was te veel.

Jean. Bedankt dat ik altijd langs mag komen om te klankborden. Bedankt voor de zinvolle discussies, het meedenken en je bemoedigingen.

Rolf. Jouw kijk op en je enthousiasme voor onderzoek hebben me jaren geleden aangestoken. Bedankt ook voor je adviezen en je meedenken.

Manasa, Haike, Miranda, Alyanne, Surya, Jim, Dennis. Bedankt voor jullie humor, relativeringsvermogen, bemoedigingen, steun, vriendschap, en enthousiasme voor onderzoek.

Siarhei. Thank you so much for all the help with my experimental setups: measuring audiovisual timing, building the ultimate stimulus box, and the nice discussions (about various topics).

Menno, Peter, Thomas. Jullie zijn geweldige vrienden. Bedankt voor jullie luisterende oor en de goede tijd met elkaar.

Rudmer. Roomie. Ik ben dankbaar voor de vriendschap die we hebben opgebouwd in de afgelopen jaren. Je hield me altijd met twee benen op de grond. Út it each mar net út it hert!

Marco. We kennen elkaar door en door. Bedankt voor je steun, relativeringsvermogen en goede adviezen.

Lieve familie en vrienden. Bedankt voor jullie meedenken, meeleven, en vertrouwen in mij.

Eileen. Wat kan ik zeggen. Jij maakt me beter. Bedankt voor wie je bent. Ik ben dankbaar dat we samen door het leven gaan.

Curriculum Vitae

Nathan van der Stoep werd op 22 augustus 1987 geboren in Dordrecht. In 2005 behaalde hij zijn gymnasium diploma aan het Johan de Witt Gymnasium te Dordrecht. Hij begon dat najaar aan zijn bachelor opleiding psychologie aan de Erasmus Universiteit te Rotterdam die hij afrondde in 2008. Aansluitend heeft hij de master biologische en cognitieve psychologie gevolgd aan dezelfde universiteit waar hij in 2009 Cum Laude afstudeerde. Na zijn afstuderen werd hij aangenomen voor het advanced research program van de Erasmus Universiteit te Rotterdam die hij afrondde in 2010. Na een korte tijd als docent psychologie aan de Erasmus Universiteit Rotterdam gewerkt te hebben begon hij zijn promotieonderzoek aan de Universiteit Utrecht onder begeleiding van Dr. Stefan van der Stigchel en Dr. Tanja Nijboer. Tijdens zijn promotie onderzoek werkte hij drie maanden in het Crossmodal Research Laboratory van Prof. Dr. Charles Spence in Oxford, Groot-Brittannië. Nathan werkt momenteel als universitair docent bij de afdeling Psychologische Functieleer aan de Universiteit Utrecht.

Nathan van der Stoep was born on August 22nd 1987 in Dordrecht. In 2005, he obtained his gymnasium diploma at the Johan de Witt Gymnasium in Dordrecht. In September 2005 he started the bachelor program in psychology at Erasmus University Rotterdam, which he finished in 2008. Following his graduation he enrolled in the master program in biological and cognitive psychology at the same university. He graduated with honors in 2009. Next, he was accepted to the Advanced Research Program at the Erasmus University Rotterdam, which he finished in 2010. After a period of teaching psychology at Erasmus University Rotterdam he started his PhD research at Utrecht University under supervision of Dr. Stefan van der Stigchel en Tanja Nijboer. During his PhD research he spent three months in the Crossmodal Research Laboratory of Prof. Dr. Charles Spence in Oxford, United Kingdom. Currently, Nathan works as an assistant professor at the department of Experimental Psychology at Utrecht University.

Publications

Van der Stoep, N., Spence, C. Nijboer, T. C. W., & Van der Stigchel, S. (accepted pending minor revisions). On the relative contributions of multisensory integration and crossmodal exogenous spatial attention to multisensory response enhancement. *Acta Psychologica*.

Van der Stoep, N., Van der Stigchel, S., Nijboer, T. C. W., & Van der Smagt, M. J. (in press). Audiovisual integration in near and far space: Effects of changes in distance and stimulus effectiveness. *Experimental Brain Research*. DOI: 10.1007/s00221-015-4248-2

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. DOI:10.1016/j.neuropsychologia.2014.12.007

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.

Nijboer, T. C. W., Ten Brink, A. F., Van der Stoep, N., & Visser-Meiley, M. A. (2014). Neglecting posture: Differences in balance impairments between peripersonal and extra personal neglect. *NeuroReport*, *25* (17), 1381-1385. DOI: 10.1097/WNR.0000000000000277

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*, 708-714. DOI: 10.3758/s13423-013-0532-y

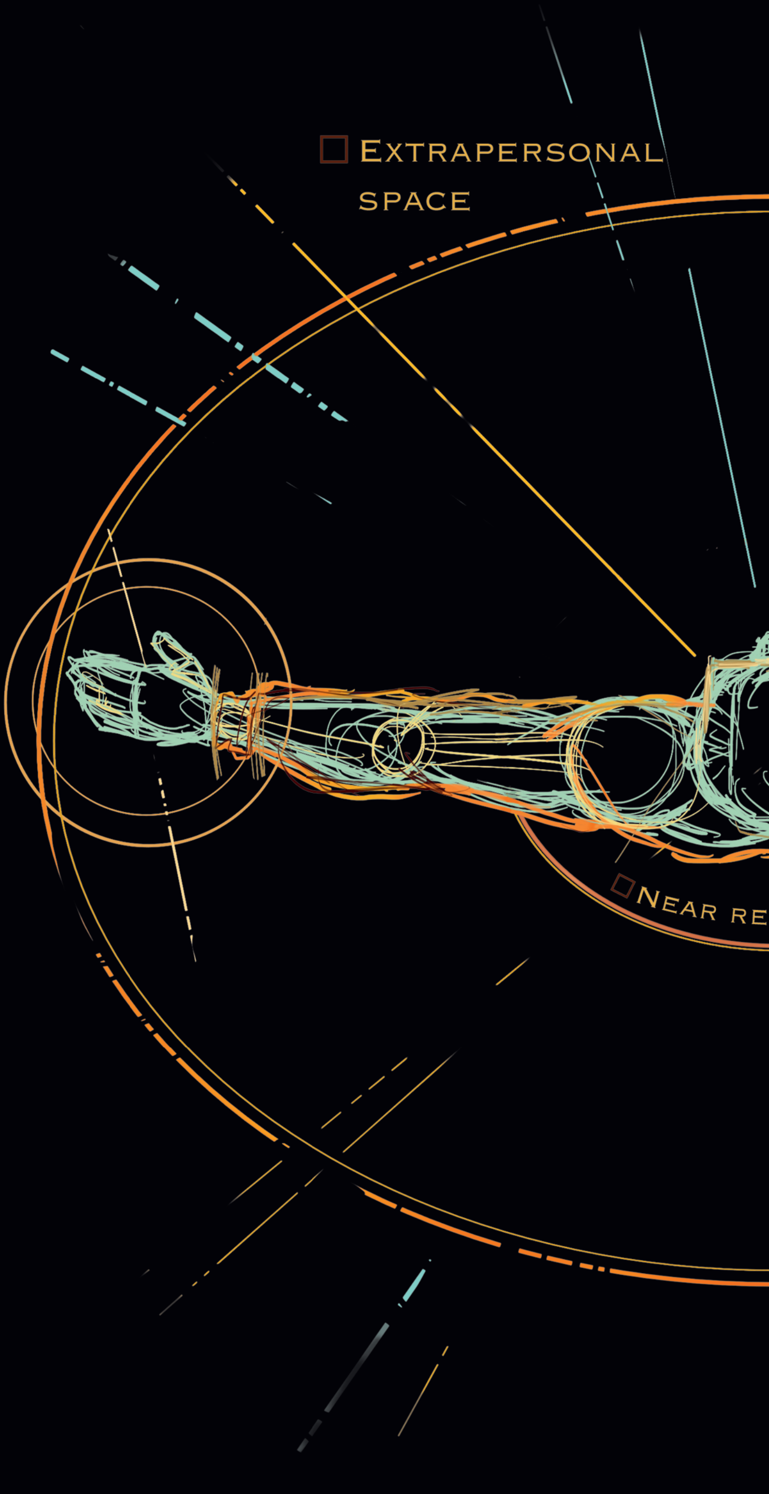
Ten Brink, A. F., Nijboer, T. C. W., Van der Stoep, N., & Van der Stigchel, S. (2014). The influence of vertically and horizontally aligned visual distractors on aurally guided saccadic eye movements. *Experimental Brain Research, 232, 1357-1366*. DOI: 10.1007/s00221-014-3854-8

Van der Stoep, N., Visser-Meily, J. M. A., Kappelle, L. J., De Kort, P. L. M., Huisman, K. D., Eijsackers, A. L. H., Kouwenhoven, M., Van der Stigchel, S., & Nijboer, T. C. W. (2013). Exploring near and far regions of space: Distance specific visuospatial neglect after stroke. *Journal of Clinical and Experimental Neuropsychology, 35*(8), 799-811. DOI: 10.1080/13803395.2013.824555

Schneider, I. K., Eerland, A., van Harreveld, F., Rotteveel, M., van der Pligt, J., Van der Stoep, N., & Zwaan, R. A. (2013). One way and the other: The bi-directional relationship between ambivalence and body movement. *Psychological Science, 24*(3), 319-325.

Van der Stoep, N., Nijboer, T. C. W., & Van der Stigchel, S. (2012). Non-lateralized auditory input enhances averaged vectors in the oculomotor system. *Experimental Brain Research, 221*(4), 377-384. DOI: 10.1007/s00221-012-3178-5

Zwaan, R. A., Van der Stoep, N., Guadalupe, T., & Bouwmeester, S. (2012). Language comprehension in the balance: The robustness of the action-compatibility effect (ACE). *PLoS ONE*. DOI: 10.1371/journal.pone.0031204


☐ Extrapersonal space
☐ Near re