



UNIVERSITY OF GENOVA

PHD PROGRAM IN BIOENGINEERING AND ROBOTICS

# Space, time and motion in a multisensory world

by

**Giorgia Bertonati**

Thesis submitted for the degree of *Doctor of Philosophy* (35° cycle)

June 2023

Dr. Monica Gori

Supervisor

Dr. Maria Bianca Amadeo

Co-Supervisor

Dr. Claudio Campus

Co-Supervisor

Prof. Paolo Massobrio

Head of the PhD program

***Thesis Jury:***

Dr. Irene Senna, *Liverpool Hope University*

External examiner

Prof. Sophie Molholm, *Albert Einstein College of Medicine*

External examiner

Prof. Silvio Sabatini, *University of Genova*

Internal examiner

**Dibris**

Department of Informatics, Bioengineering, Robotics and Systems Engineering

## **Declaration**

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Giorgia Bertonati

June 2023

## Abstract

When interacting with environmental events, humans acquire information from different senses and combine these inputs within a coherent representation of the world. The present doctoral thesis aims at investigating how humans represent space, time, and motion through auditory and visual sensory modalities. It has been widely demonstrated a predisposition of different sensory systems towards the processing of different domains of representation, with hearing that prevails in representing the time domain and vision that is the most reliable sense for processing the space domain. Given this strong link between sensory modality and domain of representation, one objective of this thesis is to deepen the knowledge of the neural organization of multisensory spatial and temporal skills in healthy adults. In addition, by using blindness as a model to unravel the role of vision in the development of spatio-temporal abilities, this thesis explores the interaction of the spatial and temporal domains in the acoustic motion perception of early blind individuals. The interplay between space and time has also been explained as the result of humans performing actions in the surrounding environment since to carry out goal-directed motor behaviors it is useful for a person to associate the spatial and temporal information of one's target into a shared mental map. In this regard, the present project also questions how the brain processes spatio-temporal cues of external events when it comes to manually intercepting moving objects with one hand. Finally, in light of the above results, this dissertation incorporates the development of a novel portable device, named MultiTab, for the behavioral evaluation of the processing of space, time, and motor responses, through the visual and acoustic sensory modality.

For the purposes of this thesis, four methodological approaches have been employed: *i*) electroencephalogram (EEG) technique, to explore the cortical activation associated with multisensory spatial and temporal tasks; *ii*) psychophysical methods, to measure the relationship between stimuli in motion and the acoustic speed perception of blind and sighted individuals; *iii*) motion capture techniques, to measure indices of movements during an object's interception task; *iv*) design and technical-behavioral validation of a new portable device.

Studies of the present dissertation indicate the following results. First, this thesis highlights an early cortical gain modulation of sensory areas that depends on the domain of representation to process, with auditory areas mainly involved in the multisensory processing of temporal inputs, and visual areas of spatial inputs. Moreover, for the spatial domain specifically, the neural modulation of visual areas is also influenced by the kind of spatial layout representing multisensory stimuli. Second, this project shows that lack of vision influences the ability to process the speed of moving sounds by altering how blind individuals make use of the sounds' temporal features. This result suggests that visual experience in the first years of life is a crucial factor when dealing with combined spatio-temporal information. Third, data of this thesis demonstrate that typically developing individuals manually intercepting a moving object with one hand take into consideration the item's spatio-temporal cues, by adjusting their interceptive movements according to the object's speed. Finally, the design and validation of MultiTab show its utility in the evaluation of multisensory processing such as the manual localization of audiovisual spatialized stimuli. Overall, findings from this thesis contribute to a more in-depth picture of how the human brain represents space, time, and motion through different senses. Moreover, they provide promising implications in exploring novel technological methods for the assessment and training of these dimensions in typical and atypical populations.

**Key-words:** space; time; motion; multisensory perception; blindness; EEG.

# Table of contents

<b>List of figures</b>	<b>vi</b>
<b>List of tables</b>	<b>viii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Spatial and temporal processing through the senses . . . . .	2
1.2 Spatio-temporal interplay in motion perception and in motor interaction with moving objects . . . . .	7
1.3 Objectives of the thesis . . . . .	11
<b>2 Cortical organization of multisensory spatial and temporal representations</b>	<b>13</b>
2.1 Study 1: the domain-specific neural organization of multisensory spatial and temporal processing . . . . .	16
2.1.1 Methods . . . . .	17
2.1.2 Results . . . . .	20
2.1.3 Discussion of Study 1 . . . . .	23
2.2 Study 2: the influence of the kind of spatial representation on the recruitment of occipital areas . . . . .	26
2.2.1 Methods . . . . .	27
2.2.2 Results . . . . .	31
2.2.3 Discussion of Study 2 . . . . .	35
<b>3 Object motion perception and interception</b>	<b>37</b>
3.1 Study 3: the role of visual experience in the use of spatio-temporal cues in the acoustic motion processing . . . . .	39
3.1.1 Methods . . . . .	41
3.1.2 Results . . . . .	47
3.1.3 Discussion of Study 3 . . . . .	50

---

3.2	Study 4: the use of visual spatio-temporal information for the manual interception of moving objects . . . . .	52
3.2.1	Methods . . . . .	54
3.2.2	Results . . . . .	59
3.2.3	Discussion of Study 4 . . . . .	61
<b>4</b>	<b>MultiTab: a novel technological device to investigate multisensory skills</b>	<b>66</b>
4.1	MultiTab system . . . . .	68
4.1.1	Hardware . . . . .	68
4.1.2	Software . . . . .	70
4.2	Technical validation . . . . .	71
4.2.1	Methods . . . . .	71
4.2.2	Results . . . . .	71
4.3	Behavioral validation . . . . .	72
4.3.1	Methods . . . . .	72
4.3.2	Results . . . . .	75
4.4	General discussion . . . . .	78
<b>5</b>	<b>General conclusions</b>	<b>81</b>
5.1	The contextual nature of multisensory processing . . . . .	81
5.1.1	Domain-specificity of multisensory representations at the cortical level	82
5.1.2	Adaptive and maladaptive use of spatio-temporal sensory cues . . .	85
5.2	Concluding remarks . . . . .	89
	<b>Scientific production</b>	<b>91</b>
	<b>References</b>	<b>94</b>

# List of figures

1.1	Processing of space, time and quantity . . . . .	8
2.1	Experimental setup of Study 1 . . . . .	18
2.2	Experimental conditions of Study 1 . . . . .	18
2.3	Cortical activity in the 50–90 ms time window for the spatial and temporal bisection . . . . .	21
2.4	Lateralized mean ERP in the 50-90 ms time window for the spatial and temporal bisection . . . . .	22
2.5	Occipital and temporal ERPs elicited by S2 during the spatial and temporal bisection . . . . .	23
2.6	Source level analysis . . . . .	24
2.7	Experimental conditions of Study 2 . . . . .	28
2.8	Occipital ERPs during the spatial bisection and the spatial localization . . . . .	32
2.9	Temporal and central ERPs during the spatial bisection and the spatial localization . . . . .	33
2.10	Cortical activity in the three time windows for the spatial bisection and the spatial localization . . . . .	33
2.11	Source level analysis . . . . .	34
3.1	Illustration of a trial . . . . .	42
3.2	Distance-duration plane of sound speeds . . . . .	43
3.3	Exemplar ellipse’s orientations . . . . .	45
3.4	Inverted psychometric curves . . . . .	46
3.5	Ellipses of early blind and sighted participants . . . . .	47
3.6	“Faster” proportion graph . . . . .	48
3.7	Group performance in the speed discrimination task . . . . .	49
3.8	Experimental procedure of Study 4 . . . . .	56

---

3.9	Graphical explanation of possible responses to the target jump . . . . .	64
3.10	Position and timing adjustment for the two target speeds . . . . .	65
4.1	MultiTab system . . . . .	69
4.2	CAD view of the device . . . . .	69
4.3	MultiTab block scheme . . . . .	70
4.4	Simultaneity of the auditory and visual stimulation . . . . .	72
4.5	Behavioral performance in the spatial localization task . . . . .	76
4.6	MREs in the audiovisual conditions . . . . .	77
4.7	Race model analysis . . . . .	78



# List of tables

3.1 Clinical details of early blind participants . . . . . 42

# Chapter 1

## Introduction

In many daily-life activities, humans encode their knowledge of the external world within the concepts of space and time. The brain attaches the observations of the where and when of external stimuli to these forms of representation in order to map the complexity of the surroundings. Despite debate on how space and time dimensions take shape, and whether these concepts are innate or experiential, clearly, they are constitutive parts of our functioning and of how we define the world. Just consider the central contribution of modern measuring instruments, which are made by humans to describe the notions of space and time with more definable variables, such as duration, timing, interval, distance, displacement, etc. Sensory modalities are the means by which we acquire the information necessary to form and manipulate space and time representations. Since humans live in a multisensory environment, all the senses contribute to the development of spatial and temporal representations. Multisensory mechanisms combine with each other inputs coming from different senses to provide redundant knowledge of the environment or assign each sense a specific domain of expertise, in particular, spatial skills to vision (Alais and Burr, 2004) and temporal skills to hearing (Burr et al., 2009). Moreover, space and time representations are shaped based on our actions in the external world. Through actions, such as goal-directed motor behaviors, the brain learns associations between space and time domains and defines a shared representational system of magnitudes (Binetti et al., 2015; Walsh, 2003). However, how the sensory modalities participate in processing these dimensions and their interaction is still an open issue.

In light of these considerations, the main purpose of this doctoral thesis is to investigate how visual and auditory sensory systems shape the representations of space, time, and motion. Chapter 1 of this thesis provides a theoretical background of the representation of spatio-temporal information through multiple senses. Chapter 2 deepens the cortical organi-

zation associated with multisensory spatial and temporal processing within a domain-specific framework. Specifically, in a first study (Section 2.1) whether the cortical gain modulation of sensory areas is domain-specific (i.e., it depends on the domain of representation, spatial or temporal, of multisensory stimuli) is tested. Subsequently, in a second study (Section 2.2), specific attention is given to the spatial domain, to investigate whether the domain-specific multisensory processing of occipital areas is also influenced by the kind of spatial representation evoked by the stimuli. Chapter 3 focuses on the interplay between spatial and temporal domains in the processing of motion, intended as the perception of external moving objects in Section 3.1, and as the implementation of motor behaviors in Section 3.2. Given the strong relationship between the visual modality and the spatial dimension, Section 3.1 explores the role of the visual experience in the first years of life in the use of spatio-temporal cues during acoustic motion perception, by testing sighted and early blind individuals. Section 3.2 deepens the interaction between space and time in planning goal-oriented motor behaviors, by analyzing the spatio-temporal components of the manual interception of moving objects. Chapter 4 of this thesis illustrates the development of a novel technological solution (Multi-Tab) for the behavioral evaluation of the processing of space, time, and motor responses with a multisensory approach.

## **1.1 Spatial and temporal processing through the senses**

The ability to evaluate the positions of one own body and objects in space is a process that characterizes human functioning since birth and develops over the years. For example, infants start to form the categories of “above” and “below” at three months of age (Quinn et al., 1996) and to process dimensions such as distance location of objects (Newcombe et al., 2005), angles (Lourenco and Huttenlocher, 2008) and height (Baillargeon and DeVos, 1991; Baillargeon et al., 1985) at five months of age. In adulthood, spatial skills reach their maximum development with complex competencies such as spatial reasoning, mental rotation, and spatial memory. Information about the spatial content of an event can be acquired through different sensory modalities, but vision seems to be the preferential sense to encode spatial inputs (Alais and Burr, 2004). To depict the world, humans also develop an internal representation of time. Newborns gradually learn to discriminate temporal intervals between inputs (Brackbill and Fitzgerald, 1972), detect temporal irregularities in an acoustic repetitive sequence (Brannon et al., 2004) and distinguish between different rhythmic attributes of auditory linguistic stimuli (Nazzi and Ramus, 2003). Later, at six and ten months of age, infants are also able to estimate the duration per se of an event (Brannon

et al., 2007; VanMarle and Wynn, 2006), and, during childhood, an age-related increase in the accuracy of this temporal estimation has been found (Droit-Volet, 2013). Among the sensory modalities conveying inputs, hearing is widely considered the most accurate sense for treating temporal information and forming a time representation of an event (Barakat et al., 2015; Bresciani and Ernst, 2007; Burr et al., 2009; Guttman et al., 2005).

Despite the dominance of specific senses for certain domains (e.g. vision for the space domain and audition for the time domain), humans get to know and interact with the world through all available senses at a time. The same environmental information is detected by more than one sense, thus the brain has the task of integrating these redundant inputs. Multisensory stimulation provides complementary information about the environment and, consequentially, results in a more precise estimate of the perceived stimulus than unisensory estimates (for an exhaustive review see Murray and Wallace, 2011). Specifically, signals coming from different sensory modalities and congruently occurring in space and time, increase the accuracy and precision of stimulus processing, more than a signal coming from only one sense. This effect of Multisensory Response Enhancement (MRE) is, for instance, revealed in shorter reaction times to multisensory stimuli as compared to the shortest reaction time to unimodal stimuli (Van der Stoep et al., 2015). Multisensory mechanisms are explained by a model of optimal combination of visual, auditory, and tactile information, for which the nervous system estimates the variances of individuals' observations and weights sensory cues accordingly (Alais and Burr, 2004; Ernst and Banks, 2002).

At the neurophysiological level, many studies have demonstrated that the midbrain structure superior colliculus is significantly involved in the integration of information among different sensory modalities (Stein and Meredith, 1993). Specifically, this subcortical structure is organized in layers, where superficial layers are unisensory, while deeper layers encode inputs coming from multiple senses (Stein et al., 2009). The work of the superior colliculus is strictly linked to the activity of cortical areas, such that the inactivation of the superior colliculus neurons by the cortex leads to multisensory responses being inhibited (Jiang and Stein, 2003). Higher-order association cortices, such as the superior temporal sulcus (Beauchamp, 2005), the intraparietal area (Andersen et al., 1997), and the frontal cortex (Fuster et al., 2000) are the cortical regions traditionally considered involved in the processing of multisensory stimuli. However, a body of research revealed that also primary visual and auditory areas could support the encoding of inputs coming from different senses (Buetti and Macaluso, 2010; Fort et al., 2002; Foxe et al., 2000; Giard and Peronnet, 1999; Molholm et al., 2002; Murray et al., 2005; van Wassenhove and Grzeczkowski, 2015). These regions own anatomical projections linking the unimodal areas themselves, which

sustain multisensory mechanisms at low levels of cortical processing (Cappe and Barone, 2005; Falchier et al., 2002; Rockland and Ojima, 2003). In this regard, results also revealed that the encoding of multiple sensory information extended over a wide range of time latencies, and that multisensory processes occur also within the first 100 ms post-stimulus onset (early-latency multisensory interactions, eMSI; Bolognini et al., 2010b; Cappe et al., 2010; De Meo et al., 2015; Gondan and Röder, 2006; Molholm et al., 2002; Raij et al., 2010; Teder-Sälejärvi et al., 2002). Overall, with these findings, research started to consider multisensory influences taking place on all levels of cortical processing, with the neocortex that is essentially multisensory (Ghazanfar and Schroeder, 2006).

Despite some basic forms of multisensory integration, such as reflexive orienting towards an audiovisual source (Neil et al., 2006), are present almost at birth, multisensory skills improve over time, as result of the infant's active interaction with a multisensory environment (Murray et al., 2016a; Murray and Wallace, 2011). Coming back to the space domain, children gradually learn to create coherent spatial representations from multiple sensory sources and reach optimal integration in all spatial skills only late in their development (Gori et al., 2008, 2012; Nardini et al., 2008). For example, Nardini et al. (2008) in a study on the use of multiple spatial cues for navigation, revealed that adults could integrate optimally multiple sources of information for navigating in the space, while young children could not (from 4 to 8 years old). This delay in the development of optimal spatial multisensory skills was related to the above-mentioned visual dominance in the processing of spatial information (Alais and Burr, 2004) that slows down the integration of the visual inputs (which are already very informative for spatial processing) with inputs coming from other sensory modalities. That vision is the preferential sense to encode spatial inputs is explained by different factors. First, the brain receives high-resolution spatial information directly from the retina, followed by characteristic retinotopic maps of the sensory inputs in the primary visual cortex (Wandell et al., 2007). Secondly, vision provides a complete representation of space as a whole, since, unlike tactile and acoustic modalities, it allows the immediate and simultaneous perception of the environmental inputs (Tinti et al., 2006). Finally, when inferring the spatial characteristics of multimodal events, visual spatialized inputs were found to bias auditory and tactile inputs presented simultaneously to the visual stimuli (Anderson and Zahorik, 2011; Flanagan and Beltzner, 2000; Pick et al., 1969; Zahorik, 2001). For example, in the so-called "ventriloquist effect" (Chen and Vroomen, 2013; Mateeff et al., 1985; Warren et al., 1981), when sensorial conflict occurs between visual and auditory/tactile stimuli, this latter is localized toward the location of the former. For what concerns multisensory temporal abilities, starting from birth infants gradually become able to match multisensory stimuli in

space and time (at six months of age; Lewkowicz, 1986), and to detect the synchrony between the auditory and visual components of vocalizing faces (Lewkowicz, 2000; Lewkowicz and Ghazanfar, 2006; Lewkowicz et al., 2010). However, for some temporal tasks, optimal multisensory integration of temporal cues does not even emerge in adulthood (Burr et al., 2009; Gori et al., 2012; Tomassini et al., 2011). Similarly to the space-vision relationship described above, the delay in the development of some multisensory temporal skills was related to audition that dominates the processing of temporal information (Barakat et al., 2015; Bresciani and Ernst, 2007; Burr et al., 2009; Guttman et al., 2005). Hearing prevails in different audiovisual temporal tasks: a visual flicker tends to be perceptually interpreted as synchronous with an auditory stimulus that is repeated at a different rate (Gebhard and Mowbray, 1959; Recanzone, 2003; Shipley, 1964). Moreover, in the so-called “temporal ventriloquism”, the perceived time of presentation of a visual stimulus is influenced by the asynchronous occurrence of an auditory event (Bertelson and Aschersleben, 2003). Finally, audition can alter the perception of a visual stimulus by making the subjects perceive multiple flashes when a single visual flash is accompanied by multiple auditory beeps (Shams et al., 2000).

That sensory systems are not mature at birth and multisensory skills need a certain degree of refinement during development is also evident from the following considerations. First, different sensory modalities develop at different rates, starting from touch, followed by vestibular, hearing, and finally vision (Gottlieb, 1971). Second, at the cortical level, protracted maturation of cortical multisensory integration was observed in typically developing people (reaching mature levels at approximately 14 years of age), which has been related to the gradual maturation of the unisensory systems underlying multisensory mechanisms and to the need for prolonged exposure to statistical relationships among multiple sensory inputs in daily life (Brandwein et al., 2011). Third, there is a discrepancy in the development of different aspects within each specific sense, since some skills go through an experience-dependent development (e.g., in the visual sensory modality, the development of spatial and temporal skills reaches different levels of encoding precision). According to the cross-sensory calibration theory (Gori, 2015), a possible way to overcome these potential obstacles to the development of cue integration is calibration. During childhood, when the body rapidly changes and the sensory systems develop at different rates, the most accurate sense for a specific task calibrates the other senses in the same task. With special reference to spatial and temporal skills development, vision seems to be the main sense to calibrate audition and touch for spatial tasks, and audition to calibrate vision and touch for temporal tasks. In support of this hypothesis, children and adults show auditory dominance in performing an

audiovisual temporal task, by relying on auditory inputs to process multisensory temporal information (Gori et al., 2012). On the contrary, children younger than 12 years of age report a clear visual dominance in a complex spatial task, which precedes multisensory integration skills (Gori et al., 2012). This mechanism is particularly important in the early years of life when the body is subjected to the highest degree of changes and needs continuous recalibration between the senses. Therefore, a priori for optimal sensory cue integration, there is the calibration of one sense on the others. Conditions of sensory deprivation (e.g., blindness and deafness) were used as models to provide evidence of this calibration. In fact, early impairment in one sense should affect some functions of the remaining senses which need calibration. For what concerns blindness, visually impaired children and adults were found to be affected in performing complex spatial tasks, compared to temporal abilities which remained intact (Gori et al., 2020a, 2014; Vercillo et al., 2016). In addition, sight restoration was found to improve auditory spatial perception in cataract-treated people, compared to untreated cataract individuals, suggesting that vision can calibrate audition also later in life and after prolonged visual impairment (Senna et al., 2022). Conversely, children with restored hearing did not show auditory dominance in processing audiovisual temporal information as typically hearing children did (Gori et al., 2017). In this thesis, I elaborate on the aspect of calibration under conditions of visual deprivation in Chapter 3.

To sum up, research shows that spatial and temporal skills develop over time thanks to the interaction of the person with a multisensory environment, with vision calibrating the other senses for space perception and audition for time perception. This strong link between visual and auditory modalities and space and time domains, respectively, suggests that the recruitment of the visual and auditory cortices might be crucial for building high-resolution spatial and temporal representations. In this regard, studies revealed a contribution of visual occipital areas to the spatial processing of information coming from different sensory modalities (Campus et al., 2019, 2017; Lewald et al., 2004; Zimmer et al., 2004). Similarly, several neuroimaging studies suggest a crucial role of the auditory cortex in temporal processing involving visual, auditory, and tactile modalities (Bolognini et al., 2010a; Coull et al., 2004; Ferrandez et al., 2003; Hyde et al., 2008; Kanai et al., 2011; Kuśmierk and Rauschecker, 2014; Lewis and Miall, 2003; Obleser et al., 2007). Overall, these findings suggest a supramodal organization of the brain, for which the domains of representation (i.e., space and time) rather than the sensory modalities are the primary design principle of human perception (Amedi et al., 2017; Cecchetti et al., 2016; Heimler and Amedi, 2020; Heimler et al., 2015; Ricciardi et al., 2014; Rosenblum et al., 2017). However, it is still not clear whether or not the domain-specificity implicit in sensory areas is present also at a

multisensory level. To answer this question, Chapter 2 of this thesis investigates the neural modulation of occipital and temporal areas during multisensory spatial and temporal tasks.

## **1.2 Spatio-temporal interplay in motion perception and in motor interaction with moving objects**

The mental representations of space and time are largely independently studied. However, the similarities between these domains and their interaction in many everyday activities suggest that common mechanisms may exist between space and time representations. The several interactions found between these components have led to the proposal of a Theory of Magnitude (ATOM) for which space and time (together with other aspects such as number, size, speed, etc.) are represented in the brain by a common magnitude system and are symmetrically interrelated (Buetti and Walsh, 2009; Burr et al., 2009; Lambrechts et al., 2013; Walsh, 2003). Information about different magnitudes, instead of being analyzed separately and then compared, are directly computed according to a common metric (Figure 1.1). In this regard, different studies highlighted interference between all these dimensions. For instance, the perception of stimuli duration is biased by the size information of an object (Xuan et al., 2007), perceiving numbers (Oliveri et al., 2008), and the side of space where stimuli are presented (Vicario et al., 2008). Similarly, research revealed the influences of time perception on space representation (Best et al., 2005; Nijhawan and Khurana, 2010; Surkys, 2021). From a developmental point of view, correspondences between space and time domains develop over the years. For example, children equate distances and durations, considering that farther in distance is equal to longer in time (Fraisse, 1963). They learn these associations according to the statistics of the environment, for which temporal and spatial features are often correlated (Kirkham et al., 2002). Finally, ATOM states that magnitude processing occurs with the activation of overlapping areas in the parietal lobe. In this regard, brain imaging evidence shows that the parietal cortex is activated in temporal, spatial, and number processes (Buetti and Walsh, 2009; Dehaene et al., 1999; Kaufmann et al., 2008; Simon et al., 2002). Moreover, damages to the parietal cortex were found to be accompanied by deficits in the temporal (Danckert et al., 2007) and spatial perception (Bjoertomt et al., 2002). However, it is worth noting that magnitude processing also overlaps in the prefrontal cortex (Burbaud et al., 1995; Ferrandez et al., 2003; Kansaku et al., 2007; Pochon et al., 2001; Rao et al., 2001; Vallesi et al., 2008).



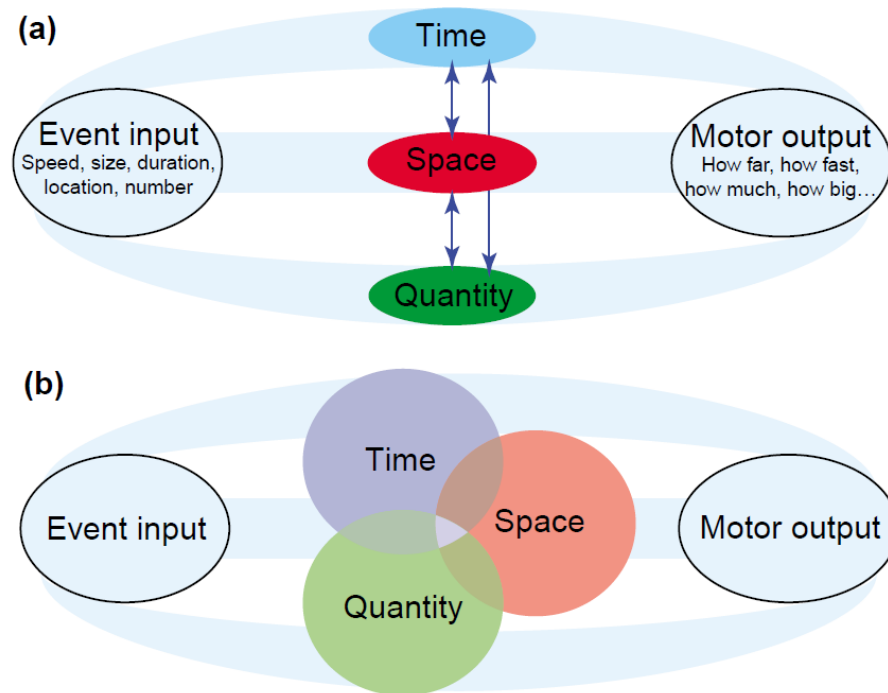


Figure 1.1 - Figure adapted from Walsh (2003). *A theory of magnitude: common cortical metrics of time, space and quantity. Trends in cognitive sciences, 7(11), 483-488. Reproduced with permission from Trends in Cognitive Sciences.*

The authors proposed two possible schemas: a) the magnitudes are analyzed separately in their individual metrics and then interrelated; b) a generalized magnitude system computes the three magnitudes according to a common metric. The authors supported the second schema.

An event in which space and time domains particularly interact is the perception of moving objects. Indeed, object motion may be described as a change of object location over time. When interacting with moving stimuli, the brain extracts both spatial (i.e., the positions of the object in space) and temporal (i.e., the arrival time of the object at each specific position) components of the item and, with this information, it also computes the object speed. Humans are able to encode the motion of an item and determine its speed either using one sense at a time or combining inputs coming from different sensory modalities. For what concerns unisensory motion perception, spatial and temporal components of motion are preferentially computed in relation to the sensory modality involved. For instance, Freeman et al. (2014) showed that temporal cues of a stimulus (i.e., its duration) dominate the auditory motion perception since listeners are more sensitive to changes in the sound duration than changes in spatial (i.e., its displacement in space) and speed (i.e., the ratio between displacement and duration) components. This preference of the auditory modality towards the temporal

cues of motion may be explained by the fact that hearing is considered the most accurate sense to compute time (Burr et al., 2009) and no motion-sensitive detectors seem to exist in the acoustic system (Carlile and Best, 2002; Carlile and Leung, 2016; Freeman et al., 2014; Locke et al., 2016). On the contrary, in perceiving visual moving stimuli, observers are sensitive to speed per se. Specifically, motion discrimination between two moving gratings is more precise when the two stimuli differ in velocity (i.e., in the ratio between displacement and duration of the stimulus) than in displacement and duration separately (Gegenfurtner and Hawken, 1995; Reisbeck and Gegenfurtner, 1999; Wardle and Alais, 2013). These findings suggest that for moving objects detected through vision, there may be velocity-tuned mechanisms rather than separable mechanisms assigned to spatial and temporal components. In this regard, extrastriate visual area V5/MT has been identified as a crucial substrate for motion processing since neurons in this area are sensitive to aspects such as the direction and speed of visual moving stimuli (Culham et al., 2001; Riecanský, 2004). For what concerns multisensory motion perception, a variety of psychophysical studies investigate the contribution of multiple senses to motion perception (Kitagawa and Ichihara, 2002; Kitajima and Yamashita, 1999; Meyer and Wuerger, 2001; Shams et al., 2002; Wuerger et al., 2003a). First, behavioral evidence shows that stimuli in one sensory modality can affect the motion perception of stimuli in another modality (Kitajima and Yamashita, 1999; Manabe and Riquimaroux, 2000; Mateeff et al., 1985; Soto-Faraco et al., 2003). However, some studies also revealed that cross-modal interactions in the processing of motion are asymmetrical, with effects of visual motion on auditory motion that are stronger than the other way around (Meyer and Wuerger, 2001; Soto-Faraco et al., 2004). These results fit the “modality appropriateness hypothesis” (Welch and Warren, 1980) according to which the sensory modality that is the most reliable for a particular task (for motion perception, the visual system which is characterized by motion-tuned mechanisms) dominates the performance in this task (Wada et al., 2003; Wuerger et al., 2010). Secondly, sensitivity to motion was found to be improved when motion cues were provided in both visual and auditory modalities, with lower variances of the perceived arrival time of the stimulus when this was bimodal than unimodal (Wuerger et al., 2003b, 2010). Finally, as with multisensory static stimuli, multisensory enhancement with moving stimuli decreases with increasing spatial (Meyer and Wuerger, 2001; Soto-Faraco et al., 2002; Wuerger et al., 2003b) and temporal distance between stimuli presented to different modalities (Ohmura, 1987; Soto-Faraco et al., 2002).

Accurate estimation of space and time, and their interaction, becomes particularly important also when we have to perform actions. In this regard, ATOM hypothesizes that space and time domains originate from the same magnitude system and are represented in coupled

metrics because they concurrently address performing actions (Binetti et al., 2015; Walsh, 2003). Through actions we learn association across different magnitudes, just experiencing, for instance, that actions that cover larger distances (e.g., travel a certain route) generally require longer execution time (e.g., route time). In this regard, behavioral studies revealed the interaction between action and time (Gavazzi et al., 2013; Hagura et al., 2012; Morrone et al., 2005; Wenke and Haggard, 2009), and action and space (Witt et al., 2004, 2005). For example, magnitude information primes grasping actions (Lindemann et al., 2007; Moretto and Di Pellegrino, 2008), and temporal estimation is modulated by whether the stimuli are within or out of the action space (Zäch and Brugger, 2008). Actions also modulate the interaction between space and time: Binetti et al. (2015) showed that the perceived rate of a tone (temporal information) was biased by changes in a bar height (spatial information), but especially when the bar height was actively controlled by the subject (and also accordingly to the magnitude of the subject's action). Moreover, the fact that the parietal lobe is the cortical area mainly ascribed for processing space and time within a common magnitude system (Buetti and Walsh, 2009; Walsh, 2003) corroborates that this interaction aims at performing actions since the parietal cortex is also commonly considered the area assigned to process goal-directed motor behaviors (Freund, 2001; Pisella et al., 2000).

To successfully perform actions, sensory information are essential guides. For example, to stop an object moving in front, one would give a look at the item in motion in order to accurately estimate where the object is, millisecond by millisecond, in relation to one's own hand (Land et al., 1999). The interception of a moving target is successful if the hand (or alternative intercepting effectors) and the target meet at the same spatial position and at the same time. Thus, interceptive movements are always constrained in space and time and require fine spatio-temporal accuracy and precision by the subject (Tresilian, 2004; Tresilian et al., 2003). For what concerns temporal precision, studies have shown that movement time of interception (MT) varies in relation to the task's temporal demands, with lower MTs when the time window for interception is shorter (Tresilian and Plooy, 2006; Tresilian et al., 2009). Also, other target factors influence people's estimates of temporal motion information. For example, temporal precision is higher when intercepting a fast target, and lower when a region of interception is specified in advance (Brenner and Smeets, 2015). Regarding spatial precision, less experimental evidence showed how humans make precise estimates of where to hit a target in space (i.e., endpoint precision). However, it is demonstrated that to meet great spatial requests (e.g., hitting small targets), people generally show higher MTs (Tresilian et al., 2009), especially if the object is moving fast (Schmidt et al., 1979). Despite the spatial and temporal components of interception are often evaluated

separately by experimental investigations, these aspects are interrelated parts of the same construct in real-life interceptive behaviors. In various experiments, Tresilian et al. (2009) showed that spatial and temporal demands of interception are interdependent: the spatial accuracy achieved influences the temporal accuracy and vice versa, with, for example, small spatial errors that are accompanied by high temporal errors. Finally, in order to meet the spatio-temporal accuracy demands of the task, humans are also able to adjust the interceptive movements based on errors in previous attempts (López-Moliner et al., 2019). Since the way one repeatedly hits some moving objects is fairly constant over time, a person can learn from these errors and plan future movements more efficiently. Past studies investigated how people correct spatial and temporal errors in response to external perturbations (Burge et al., 2008; Thoroughman and Shadmehr, 2000; van Beers, 2009) and revealed, for example, that the way one hits a moving target (e.g., the hand's initial movement or the final hitting position) depends on the speed of preceding targets (De Lussanet et al., 2001). Thus, over the last few years, scientific research increasingly agrees that different sensorimotor strategies of interception can be implemented depending on the specific task and context (DeLucia, 2004; Hecht and Savelsbergh, 2004; Regan and Gray, 2000; Tresilian, 1999; Zago et al., 2009). Indeed, many theories of motor behaviors suggest an “optimal performance” model, for which the brain calculates the cost-benefit ratio of applying specific sensorimotor programs to achieve task goals and, from this computation, it obtains detailed behavioral prediction for optimal performance (Todorov, 2004).

To sum up, scientific literature shows that humans are able to extract spatio-temporal information from objects traveling in the surroundings and process these motion cues to define items' characteristics (e.g., speed, direction, etc.) and program motor actions on them. Sensory systems differently contribute to these motion processes (Albright and Stoner, 1995; Braddick et al., 2003; Carlile and Leung, 2016), which are increasingly considered context-dependent (DeLucia, 2004; Hecht and Savelsbergh, 2004; Regan and Gray, 2000; Tresilian, 1999). However, these findings raise the questions addressed in Chapter 3 of how the domain-specific processes of motion perception manifest themselves in the absence of vision, and whether the context-dependence of interception of moving targets persists in sensorimotor error correction.

### **1.3 Objectives of the thesis**

In light of the state-of-art, the overall objective of this doctoral thesis is to study how humans represent space, time, and motion through the auditory and visual sensory modalities, by

investigating the domain-specificity associated with multisensory spatio-temporal processing and the interplay between these domains during motion perception and moving target's interception. Chapter 1 illustrates how building spatial and temporal representations is a fundamental task for humans to define the world and act on it. In Chapter 2 I shed light on the domain-specific neural organization associated with multisensory spatial and temporal processing. Thanks to the use of the electroencephalography (EEG) technique, I reveal activation of regions likely involving the visual and auditory cortices that depends on the domain of representation of multisensory stimuli, but also on the complexity behind these representations. Then, starting from the idea that "Space and time are tightly interwoven dimensions in the brain" (Binetti et al., 2015), Chapter 3 describes the interaction between space and time dimensions in motion processing. First, a psychophysical study investigating how visual deprivation shapes the use of spatio-temporal cues in acoustic motion perception shows that early blind individuals struggle in discriminating the speed of moving sounds. Second, a paradigm using a motion capture technique allows exploring the manual response of healthy subjects in the interception of a moving item, by measuring how people adjust their hand movements to correct errors in previous attempts. In Chapter 4 this thesis brings together the scientific results of the previous chapters in a final technological aim that consists in the development of a novel device for the behavioral evaluation of the processing of space, time, and motor responses. Finally, in Chapter 5, which is the final chapter of this doctoral thesis, I discuss the results in light of the existing literature, and I describe the context-dependent mechanisms that these findings emphasize.

Overall, the results obtained from this project improve knowledge of the neural mechanisms governing the spatial and temporal encoding and the perceptual and behavioral implications of the interaction between these domains of representation in motion processing. Having this kind of knowledge will allow to rethink the categories of space and time for depicting the external world not as independent dimensions but as interconnected parts that work together to successfully respond to contextual demands. Based on these results, technological and rehabilitation solutions identified to support the development of these forms of representation will be even closer to reality and to the daily needs of the healthy population and those with sensory disabilities.

## Chapter 2

# Cortical organization of multisensory spatial and temporal representations

In Chapter 1, I introduced the main research findings related to how the brain relies on sensory perception to represent the world within the domains of space and time. Overall, such results indicate a fundamental role of vision in shaping the spatial representation, and of audition in the temporal representation. In light of these considerations, Chapter 2 aims at investigating how the sense-domain association manifests itself at the cortical level.

Neurosciences have widely questioned how the human brain perceives the surrounding world through the senses and how sensory information interacts with each other. For many decades, unisensory-based segregation of brain mechanisms has been theorized, with cortical organization constrained to specific sensory modalities (Amedi et al., 2017; Heimler and Amedi, 2020). This functional specialization has been described as the result of natural selection mechanisms. For instance, brain specialization associated with visual processing was generally identified in the retinotopic mapping and in the two visual streams, ventral and dorsal (Goodale and Milner, 1992), which are organized based on the kinds of information that are particularly important to primates (e.g., faces, movements, tools, etc.). More recently, this sensory-anchored assumption of brain mechanisms has been extensively reconsidered. First, it was found that certain cortical areas, including the sensory cortices, are recruited in the processing of stimuli conveyed by different sensory modalities. The visual cortex is not solely involved in processing visual inputs since overlapping activation of these areas has been detected also during non-visual tasks (Lucan et al., 2010; Romei et al., 2009; Vetter et al., 2014). For example, Lucan et al. (2010) reported neural responses of the lateral occipital complex (LOC) during tactile shape discrimination in the same time window in which the LOC is generally recruited for visual object recognition. Similarly,

---

auditory areas were proven not to be involved exclusively in acoustic processing, but also in the encoding of visual and somatosensory inputs (Rosenblum et al., 2017). Second, multisensory functions (i.e., the processes of combining information from multiple sensory modalities) are no longer considered the exclusive task of association cortices (as described by previous studies; Andersen et al., 1997; Beauchamp, 2005; Fuster et al., 2000). Occipital, temporal, and central areas, including primary sensory cortices, were found to also support the combined encoding of multiple sensory inputs simultaneously (Bueti and Macaluso, 2010; Fort et al., 2002; Foxe et al., 2000; Giard and Peronnet, 1999; Molholm et al., 2002; Murray et al., 2005; van Wassenhove and Grzeczowski, 2015). Third, some authors preferentially defined cortical organization as supramodal (Cecchetti et al., 2016; Heimler et al., 2015; Ricciardi et al., 2014, 2020). Given that specific cortical regions were found to be activated across different sensory modalities (Pietrini et al., 2004), they hypothesized the existence of a more abstract representation of the perceived stimuli within the sensory cortices, which does not depend on the sensory modality conveying the inputs, but on stimuli content to respond. This mechanism was also defined by some authors as task-specific sensory-independent (TSSI; Amedi et al., 2017; Heimler et al., 2015) brain organization. Overall, these observations suggest that the sensory modalities no longer seem to be the main design principle of the cortical architecture and that other aspects, such as stimuli content, task, and domain, may be responsible for this operation. In this regard, some studies have hypothesized a domain-specific organization of cortical activity. Specifically, occipital areas were shown to mainly support neural responses underlying the space domain (Campus et al., 2017), while temporal regions to primarily shape the time domain (Amadeo et al., 2020a). This pattern of activation was independent of the sensory modality conveying the sensory information, with occipital and temporal areas that responded to spatial and temporal inputs, respectively, in both the visual and auditory modalities.

In addition to these considerations, if specific computations rather than sensory modalities mainly constrain cortical specialization, the sensory experience should not be a mandatory prerequisite for the brain to develop its functional organization (Ricciardi et al., 2014, 2020). Studies on sensory deprivation provide a unique opportunity to explore this assumption. For instance, overlapping activation in cortical regions traditionally considered “visual” in nature was observed in both sighted and blind individuals (Battal et al., 2022; Bonino et al., 2015; Collignon et al., 2011; Kupers and Ptito, 2014; Röder et al., 2002; Striem-Amit et al., 2012; Weeks et al., 2000). These patterns of activation should not be due to visual imagery processing since they were also revealed in individuals blind from birth or at a very early age (i.e., with no visual memories) (Pietrini et al., 2004). Similarly, but in the case of hearing

impairments, visual and somatosensory stimulations were shown to induce cortical activation in the auditory cortices of deaf individuals (Auer et al., 2007; Bavelier et al., 2006; Bottari et al., 2014; Karns et al., 2012; Levänen et al., 1998; Rosenblum et al., 2017). In both blindness and deafness, the cortical activity of the sensory-deprived brain may be interpreted in two ways: either as the result of a plastic rearrangement due to the lack of one sense (cross-modal plasticity) or as the expression of an abstract representation of sensory information which is supramodal and pre-programmed in the human brain (supramodal organization). These two views should not necessarily be seen as mutually exclusive but as two sides of the same coin. Specifically, while some degree of brain activity seems to occur despite the absence of any sensory experience, the neural responses of some other deprived cortical areas can be affected by cross-modal reorganization due to plasticity (Cecchetti et al., 2016). It is worth noting that despite the engagement of sensory regions by the intact senses of sensory-deprived people (e.g., of the visual cortices by auditory and tactile modalities of blind individuals), some studies showed that the lack of the primary and more informative sense for a specific domain of representation (e.g., the lack of vision for spatial tasks) influenced the acuity of this domain in the sensory impaired population (Amadeo et al., 2019b, 2022; Gori et al., 2020a, 2022a, 2017, 2014; Vercillo et al., 2016). Thus, as partially mentioned also in Chapter 1, the fact that there are better senses for more effectively communicating certain types of information would explain why specific sensory regions are recruited in the representation of spatial and temporal information.

In light of these considerations, the studies presented in this chapter want to test the idea that the domains of representation primarily shape human perception and its associated cortical activity by investigating multisensory processes at the cortical level. Specifically, in Study 1 (Section 2.1), we asked participants to perform audiovisual spatial and temporal tasks (for which the bimodal stimuli were the same between the two tasks and the only difference was the experimental request) and we investigated whether the space and time domains modulate the cortical activation also in response to multisensory stimuli. Subsequently, in Study 2 (Section 2.2), we focused on the spatial domain, and we explored whether also other aspects of this domain, such as the kind of layout in which the spatial multisensory information is represented, influence the modulation of occipital activity.



## **2.1 Study 1: the domain-specific neural organization of multisensory spatial and temporal processing**

Vision and hearing have been widely associated with the processing of the space and time domains, respectively, and this link supported at the cortical level by sensory-independent recruitment of the visual and auditory cortices during spatial and temporal tasks. In this regard, the past literature revealed the recruitment of occipital areas underlying complex spatial representations of healthy individuals (Campus et al., 2017), and of auditory regions supporting the representation of a temporal metric (Amadeo et al., 2020a). This pattern of activation emerged, for example, in studies using the bisection paradigm. In the bisection, three stimuli are reproduced in sequence, with the first and the third stimuli delivered at fixed spatial and temporal intervals, and the second stimulus randomly reproduced at two different spatial positions and temporal lags. Participants compare the stimuli with each other and evaluate whether the second stimulus is spatially (spatial bisection) or temporally (temporal bisection) farther from the first or the third stimulus. This paradigm is particularly interesting since, by using exactly the same stimuli but changing only the task request, it can elicit either a spatial or a temporal representation of the sensory inputs. Previous studies using this paradigm revealed that the activity of visual and auditory cortices is particularly important in the processing of this kind of task (Amadeo et al., 2019a, 2020a,b; Campus et al., 2019, 2017; Martolini et al., 2020). Studies on sensory deprivation also provided evidence of this assumption. Gori and colleagues revealed that people with visual impairment were impacted when performing an auditory spatial bisection task and that this observation was complemented by a reduced occipital response for acoustic space perception of early blind individuals (Amadeo et al., 2019a, 2020b; Campus et al., 2019; Gori et al., 2020a,b, 2014; Tonelli et al., 2020; Vercillo et al., 2016). Similarly, but in the case of deafness, recent studies demonstrated impairment in a visual temporal bisection task by deaf individuals, accompanied by reduced recruitment of auditory areas compared to hearing people (Amadeo et al., 2019b, 2022; Gori et al., 2022a, 2017). Overall, these findings highlighted the role of the domain of representation in the organization of cortical sensory regions. Studies on the neural mechanisms underlying multisensory perception sustain this view by showing that also occipital, temporal, and central areas support the encoding of multiple sensory modalities (Buetti and Macaluso, 2010; Fort et al., 2002; Foxe et al., 2000; Ghazanfar and Schroeder, 2006; Giard and Peronnet, 1999; Molholm et al., 2002; Murray et al., 2005; van Wassenhove and Grzeczkowski, 2015). In addition, these multisensory processes occur over a wide range of temporal latencies, also within the first 100 ms post-stimulus onset (reviewed

in De Meo et al., 2015). For example, Bolognini et al. (2010b) observed facilitation of auditory stimulation on the perception of phosphenes (induced by Transcranial Magnetic Stimulation delivered to the occipital visual cortex) in a temporal window <50 ms from stimulus onset.

Given that visual and auditory areas play a fundamental role in the development of the spatial and temporal representation respectively, and that these cortical areas are also essentially multisensory, in the next sections we investigated whether the multisensory nature of sensory areas is modulated by the domain of representation at the cortical level (Gori et al., 2022b). To test this, Event-Related Potentials (ERPs) and psychophysical responses were recorded during multisensory (audiovisual) spatial and temporal bisection tasks in typically developing individuals.

### 2.1.1 Methods

#### Sample and experimental procedure

16 adults (9 females, mean age  $\pm$  standard deviation (SD):  $24 \pm 2.95$  years old) participated in the study. All participants reported no history of neurological, cognitive, and/or sensory deficits and they gave written informed consent prior to testing. We decided on sample size based on previous studies testing the neural correlates of spatial and temporal abilities during similar tasks (Amadeo et al., 2019a; Campus et al., 2019, 2017). A priori power analysis (two-tailed t-test, estimated effect size Cohen's  $d = 0.80$ ,  $\alpha = 0.05$ ) determined a minimum sample size of 15 participants. The study was approved by the ethics committee of the local health service (Comitato etico, ASL 3 Genova) and conducted in line with the Declaration of Helsinki.

Participants sat at a distance of 180 cm from the center of a horizontal array (Figure 2.1) composed by 23 speakers spatially aligned with 23 Light Emitting Diodes (LEDs) and spanning  $\pm 25^\circ$  of visual angle ( $0^\circ$  represented the central speaker/LED, with negative values on the left and positive values on the right). Each trial involved three audiovisual (AV) stimuli (namely S1, S2, S3) played from the horizontal array at three different spatial positions and time lags. Each AV stimulus consisted in a single sound (60 db SPL at ears' level, 500 Hz) simultaneously reproduced for 75 ms with a single red flash ( $2.3^\circ$  diameter,  $20 \text{ cd/m}^2$  luminance). S1 and S3 were always played at  $-25^\circ$  and  $+25^\circ$ , respectively, and separated by a fixed time interval of 1.5 seconds. S2 could be randomly presented from either  $-2.3^\circ$  or  $+2.3^\circ$  in space and at either  $-250 \text{ ms}$  or  $+250 \text{ ms}$  in time. S2 randomization in spatial positions and time lags led to four possible conditions (Figure 2.2): a) S1-S2 distance/interval narrow

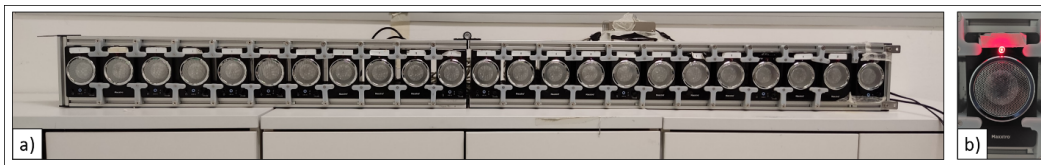


Figure 2.1 - Experimental setup. a) A horizontal array of 23 speakers and 23 LEDs. b) Detail of one speaker spatially aligned with one LED. *Reproduced with permission. Gori et al. (2022b). Multisensory representations of space and time in sensory cortices. Human Brain Mapping, 44(2), 656– 667.*

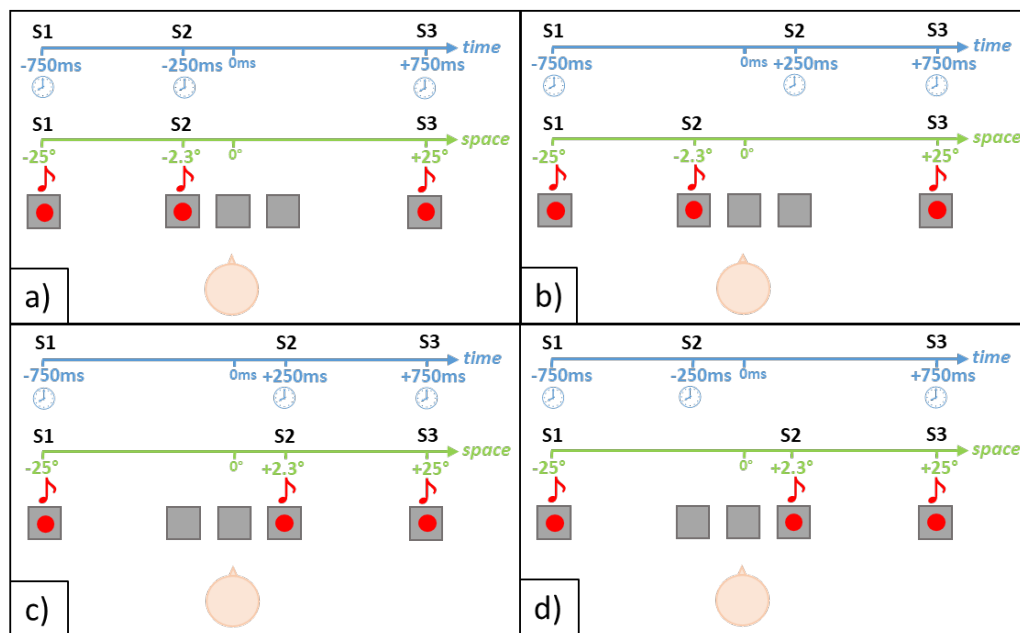


Figure 2.2 - Four experimental conditions according to S2 spatial and temporal features. *Reproduced with permission. Gori et al. (2022b). Multisensory representations of space and time in sensory cortices. Human Brain Mapping, 44(2), 656– 667.*

in space and short in time, b) S1-S2 distance/interval narrow in space and long in time, c) S1-S2 distance/interval wide in space and long in time, and d) S1-S2 distance/interval wide in space and short in time.

Participants performed a spatial bisection task and a temporal bisection task, in counter-balanced blocks across subjects. In the spatial bisection task, participants evaluated whether S2 was spatially farther from S1 or S3, while in the temporal bisection task they evaluated whether S2 was temporally farther from S1 or S3. AV stimuli were identical in both tasks, which differed only in relation to the experimental question which focused either on the spatial or the temporal features of S2 in relation to S1 and S3. Participants were asked

to maintain a stable head position that was continuously monitored by the experimenter, together with the electrooculogram (EOG) signal.

### **EEG data collection, pre-processing, and data analysis**

During the tasks, EEG and EOG were recorded with 64 active electrodes using the Biosemi ActiveTwo EEG System. Thus, within the EEGLAB plug-in in Matlab (Delorme and Makeig, 2004), EEG data have been pre-processed by applying the Artifact Subspace Reconstruction method (ASR) to remove transient stereotypical (e.g. eye blinks) and non-stereotypical (e.g. movement or muscle bursts) high-amplitude artifacts. ASR uses a sliding window technique, decomposing each window of EEG data via principal component analysis and allowing statistical comparison with data from a clean baseline EEG recording. Within each sliding window, the ASR algorithm identifies principal subspaces which significantly deviate from the baseline and then reconstructs these subspaces using a mixing matrix computed from the baseline EEG recording. Additionally, Independent Component Analysis (ICA) has been used with two EEGLAB toolboxes (SASICA, Chaumon et al., 2015; ICMARC, Frølich et al., 2015).

We compared the mean ERPs response to S2 with that to S1 for the spatial and temporal bisection tasks separately. We performed this analysis since previous studies showed that S2, contrary to S1, represents the starting point for the development of spatial and temporal metrics correlated with early contralateral activation of occipital and temporal areas, respectively (Amadeo et al., 2020a; Campus et al., 2019, 2017). To obtain the ERPs, we focused on electrodes related to visual (O1, O2 in occipital areas) and auditory (T7, T8, in temporal areas) processing, and we computed mean ERP amplitude by averaging the voltage in a time window between 50 and 90 ms after the stimulus, which was defined as a crucial early interval for multisensory processing. We then grouped ERPs recorded on the contralateral and the ipsilateral hemisphere with respect to stimulus position in space and we computed lateralized ERP responses (i.e., the difference between the contralateral and ipsilateral ERPs responses). Analysis of variance (ANOVA) was run on lateralized mean ERP responses to perform statistical comparisons, considering as factors: Area (Temporal, Occipital), Task (Spatial bisection, Temporal bisection), and AV stimulus (S1, S2). Paired two-tailed t-tests were performed as post hoc comparisons and Bonferroni corrected (alpha level 0.05).

In order to estimate the cortical generators of the ERP components, we performed a distributed source analysis using the Brainstorm software (Tadel et al., 2011). We re-referenced data to the common average, we used standard 1 mm resolution template of the Montreal Neurological Institute (non-linear average of 152 subjects, processed with FreeSurfer 5.3

ICBM152, Fonov et al., 2009), we performed forward modeling using a three-layer (head, outer and inner skull) symmetric boundary element model (BEM) generated with Open-MEEG86, and we estimated source intensities using the sLORETA approach (Gramfort et al., 2011). Brainstorm's default parameter settings have been used for both source reconstruction and BEM creation. Dipole orientations were not fixed to the cortex surface but were free to assume whichever orientation, in order to avoid misleading over-interpretation. Source activation for each subject and condition was averaged within the selected 50-90 ms time window after S2 and then the norm of the vectorial sum of the three orientations at each vertex was estimated. In the end, we performed paired t-tests to investigate pairwise comparisons (results were corrected for multiple comparisons with the False Discovery Rate (FDR) method, and a  $p = .00001$  was used as threshold), specifically by comparing the neural response after S2 between the two bisection tasks, considering S2 positions in space ( $\pm 2.3^\circ$ ) separately.

Finally, behavioral performance was computed as the % of correct responses for each task.

### 2.1.2 Results

Behavioral performance analysis showed that participants performed equally well in the two tasks ( $t_{(15)} = 1.80$ ,  $p = 0.091$ , Cohen's  $d = 0.45$ , 95% CI = [-0.08, 0.98]). At the neural level, we revealed in the 50–90 ms time window after S1 onset (Figure 2.3) activation involving the temporal and the occipital areas, that was contralateral to the AV stimulus position in space. This activation appeared similar between the temporal and the spatial bisection tasks, in both occipital and temporal areas. On the contrary, the scalp topographies in the same time window but after the S2 onset showed a more prominent positivity than S1 in occipital areas for the spatial bisection task and S1 in temporal areas for the temporal bisection task. This cortical positivity was always lateralized with respect to the AV stimulus position in space.

ANOVA on lateralized mean ERP responses corroborated these results by revealing a significant three-way interaction between Area (Temporal, Occipital), Task (Spatial bisection, Temporal bisection), and AV stimulus (S1, S2) ( $F_{(1,15)} = 123.1$ ,  $p < 0.001$ , partial  $\eta^2 = 0.89$ , 95% CI [0.75, 0.94]). We further investigated this result by splitting the analysis into two distinct hypothesis-driven follow-up ANOVAs to focus on occipital and temporal areas separately. The Task X AV Stimulus follow-up ANOVA on temporal regions showed a contralateral temporal activity that was stronger during the temporal bisection task than during the spatial bisection task ( $F_{(1,15)} = 26.76$ ,  $p < 0.001$ , partial  $\eta^2 = 0.64$ , 95% CI [0.28,

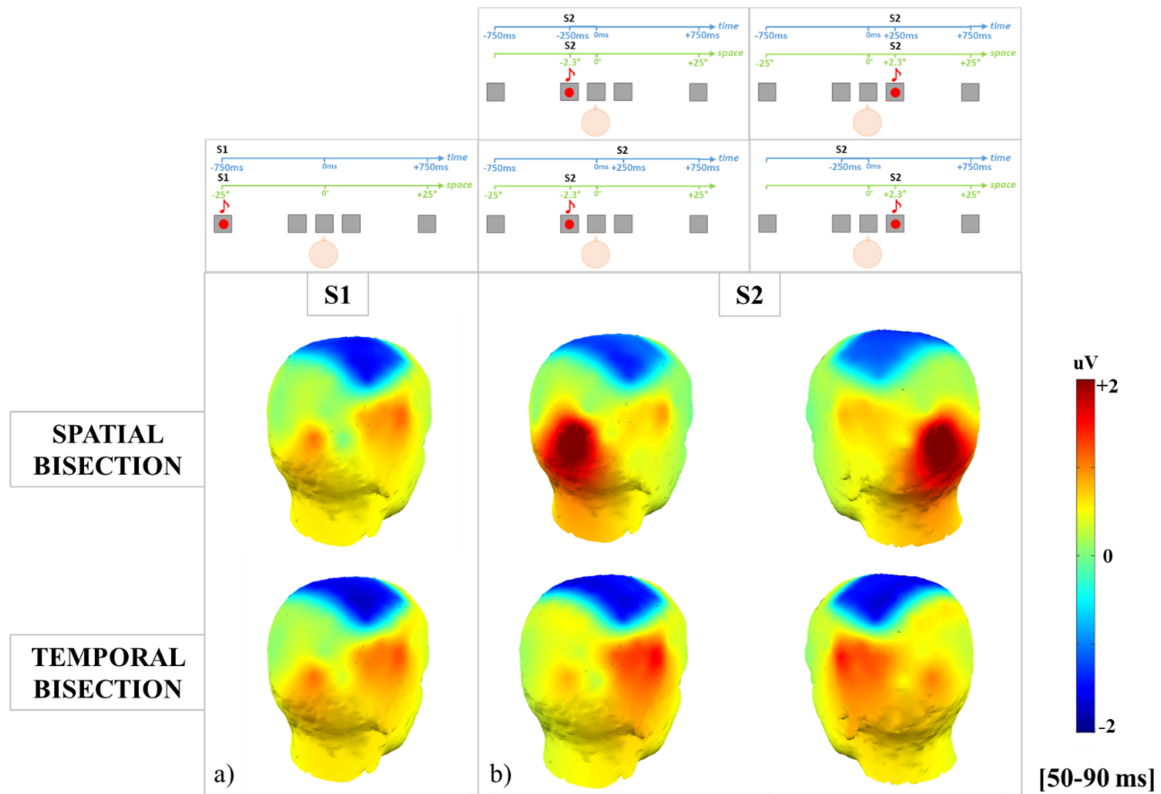


Figure 2.3 - Scalp maps of the mean ERP amplitude in the 50–90 ms time window after S1 (a) and S2 (b), for the spatial (top) and temporal (bottom) bisection task. On top, a schematic representation of each condition. *Reproduced with permission. Gori et al. (2022b). Multisensory representations of space and time in sensory cortices. Human Brain Mapping, 44(2), 656– 667.*

0.80]). However, a significant interaction between Task and AV Stimulus ( $F_{(1,15)} = 51.63$ ,  $p < 0.001$ , partial  $\eta^2 = 0.77$ , 95% CI [0.51, 0.88]) showed that the gain modulation during the temporal bisection was not similar between S1 and S2. In this regard, post hoc two-tailed t-tests (Figure 2.4) revealed that the response in the temporal regions was higher during the temporal bisection than during the spatial bisection specifically for S2 ( $t_{(15)} = -7.34$ ,  $p < 0.001$ , Cohen's  $d = -1.83$ , 95% CI = [-2.67, -0.99]), while for S1 the two tasks showed a similar temporal activation ( $t_{(15)} = -0.63$ ,  $p = 1.00$ , Cohen's  $d = -0.15$ , 95% CI = [-0.67, 0.35]). To sum up, these results on the temporal activation indicated an amplification of the contralateral response within the 50-90 ms time window that was specific for multisensory stimuli involved in the processing of a temporal metric. The follow-up ANOVA on the occipital areas showed greater lateralized ERP responses for the spatial bisection task than for the temporal bisection task ( $F_{(1,15)} = 51.73$ ,  $p < 0.001$ , partial  $\eta^2 = 0.78$ , 95% CI [0.51,

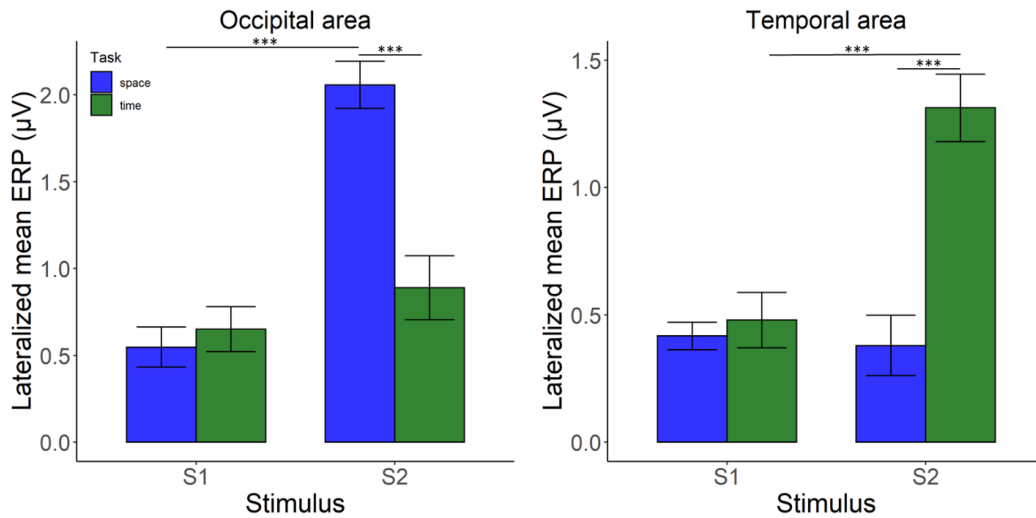


Figure 2.4 - Lateralized mean ERP amplitude (i.e., the difference between the contralateral and ipsilateral ERP responses) in the selected time window (50–90 ms) after S1 and S2 of the two bisection tasks in occipital (left panel) and temporal (right panel) areas. Error bars indicate the standard error of the mean (SEM). \*\*\*  $p < 0.001$ . *Reproduced with permission. Gori et al. (2022b). Multisensory representations of space and time in sensory cortices. Human Brain Mapping, 44(2), 656–667.*

0.88]) in the 50-90 ms time window. However, the significant interaction between Task and AV Stimulus (Figure 2.4;  $F_{(1,15)} = 44.17$ ,  $p < 0.001$ , partial  $\eta^2 = 0.75$ , 95% CI [0.45, 0.86]) revealed that this cortical gain of the spatial bisection on the temporal bisection was specific for S2 ( $t_{(15)} = 9.07$ ,  $p < 0.001$ , Cohen's  $d = 2.27$ , 95% CI = [1.30, 3.23]) and not for S1 ( $t_{(15)} = -0.91$ ,  $p = 0.373$ , Cohen's  $d = -0.22$ , 95% CI = [-0.74, 0.28]). Overall, results on the occipital cortical response suggested that the neural modulation of these regions was associated with the processing of a spatial metric more than of a temporal representation.

Over the occipital scalp (Figure 2.5), ERP waveforms recorded within the 50-90 ms time window after S2 onset showed a stronger contralateral response during the spatial bisection task than during the temporal bisection task. ERP waveforms also showed a not lateralized modulation of later neural response P140 specific to the spatial bisection task and a contralateral modulation occurring in the 250-450 ms time window, more pronounced for the spatial task. Over the temporal scalp (Figure 2.5), a contralateral ERP component after 50-90 ms from S2 onset was stronger during the temporal bisection task than during the spatial bisection task.

Finally, the source level analysis (Figure 2.6) showed that both bisection tasks elicited after S2 a cortical response contralateral to the stimulus position, in occipital and temporal regions. However, the early occipital response was stronger during the spatial bisection than

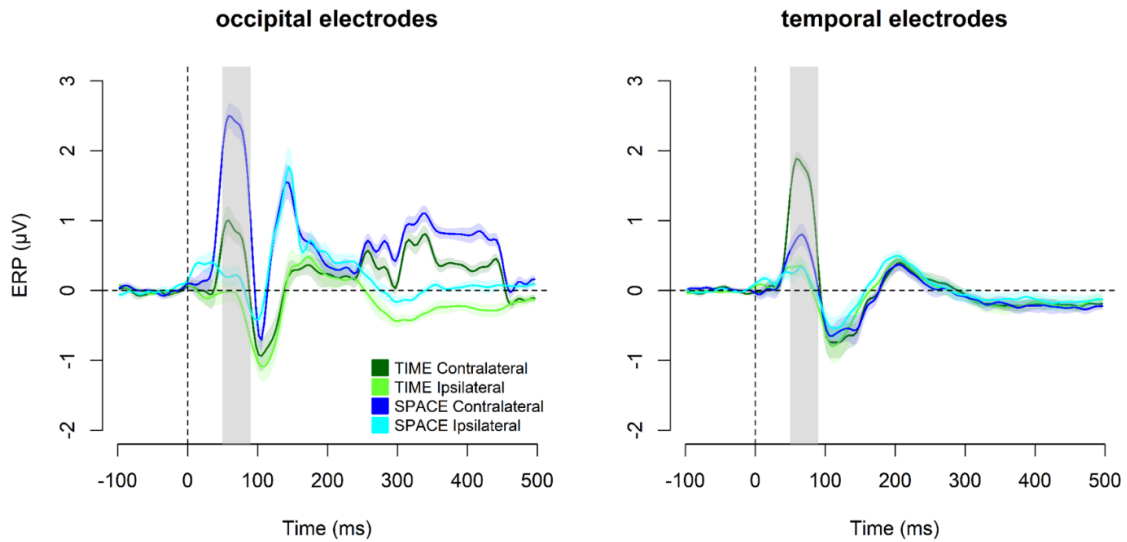


Figure 2.5 - ERPs (mean  $\pm$  SEM) elicited by S2 during the spatial bisection and the temporal bisection tasks in occipital (left panel) and temporal (right panel) areas. The gray-shaded area delimits the selected time window (50–90 ms). *Reproduced with permission. Gori et al. (2022b). Multisensory representations of space and time in sensory cortices. Human Brain Mapping, 44(2), 656–667.*

during the temporal bisection task, while the temporal activation was greater during the temporal bisection than during the spatial bisection task. Paired two-tailed t-tests confirmed the significant differences between the two tasks in the recruitment of the occipital and temporal cortices. Overall, this activation originated in wide temporal and occipital regions that likely involved the auditory and visual cortices respectively.

### 2.1.3 Discussion of Study 1

In Study 1, we tested the hypothesis that the pivotal role of occipital and temporal areas for the spatial and temporal representation, respectively (Amadeo et al., 2020a; Campus et al., 2019, 2017), acts also within a multisensory framework. In two separate tasks, participants evaluated the spatial (spatial bisection task) and the temporal (temporal bisection task) relationships between three audiovisual stimuli. Results showed an early activation (50-90 ms after S2 onset) of temporal regions that was larger during the temporal bisection task than during the spatial bisection task. Complementarily, an occipital response occurring after S2 was stronger when encoding the audiovisual stimuli in a spatial bisection task than in a temporal bisection task. This activation, originating in wide temporal and occipital regions,



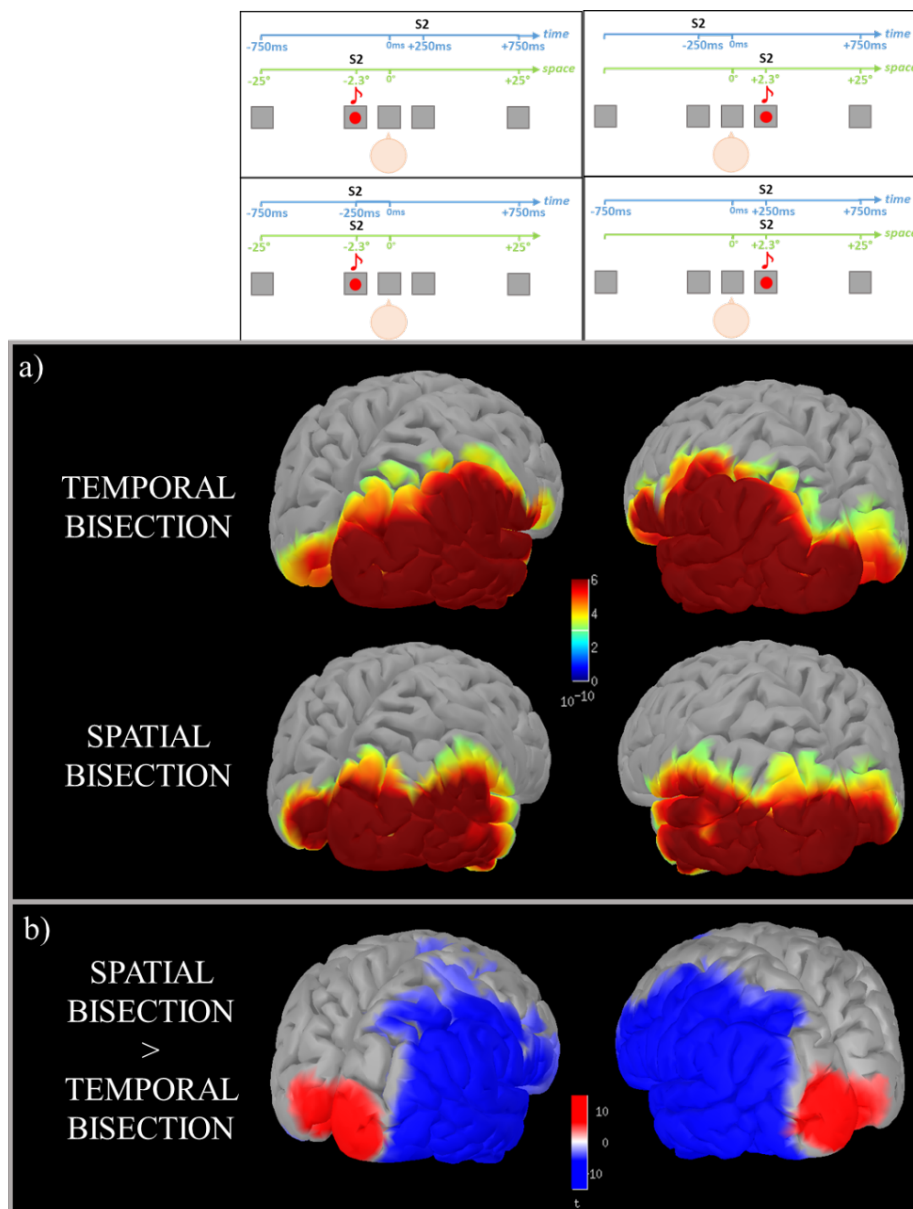


Figure 2.6 - Source level analysis. a) Average source activity after S2 in the 50–90 ms time window. b) Results of the pairwise two-tailed t-tests performed on average source activity in the 50–90 ms time window: only t values corresponding to  $p < .0001$  after FDR correction are displayed. Reddish and bluish colors indicate stronger activation in spatial and temporal bisections, respectively. Color intensity indicates the significance of the difference (i.e., the magnitude of  $t$ ). *Reproduced with permission. Gori et al. (2022b). Multisensory representations of space and time in sensory cortices. Human Brain Mapping, 44(2), 656–667.*

likely involved the auditory and visual cortices, respectively. Moreover, temporal responses

recalled the N1 component usually elicited by auditory stimuli (Näätänen and Picton, 1987), while occipital responses resembled the visual-evoked C1 (Di Russo et al., 2002).

Many past studies already revealed that visual and auditory cortices can support the encoding of multiple sensory modalities (Buetti and Macaluso, 2010; Fort et al., 2002; Foxe et al., 2000; Giard and Peronnet, 1999; Molholm et al., 2002; Murray et al., 2005; van Wassenhove and Grzeczkowski, 2015) and that this recruitment can happen in early time latency (within the first 100 ms post-stimulus onset; Bolognini et al., 2010a; De Meo et al., 2015; Fort et al., 2002; Giard and Peronnet, 1999; Murray et al., 2016b; Raji et al., 2010; Talsma et al., 2007; Teder-Sälejärvi et al., 2002). However, results of Study 1 also show that these early multisensory mechanisms can take into account some stimuli features, which in this case are the domains of representation (i.e., space and time) to which the stimuli belong. Indeed, we used the same audiovisual stimuli in the two bisection tasks and changed only the experimental question between them. Thus, this experimental paradigm allowed us to detect the cortical activation of identical sensory information but with different behavioral goals. Overall, these results support the idea that the visual and auditory circuits are crucially enrolled whenever dealing with spatial and temporal representations, respectively, also at the multisensory level. Finally, we showed that the behavioral performance was similar between the two tasks, which confirmed that the neural modulation of sensory areas referred essentially to the task request, rather than to any effect of task difficulty.

The findings of this study should be considered in light of the lack of unimodal conditions (only auditory and only visual), which did not allow us to infer that the domain-specific neural response we observed was intrinsically multisensory. In fact, although we asked participants to respond to both stimulation, people could have processed only the visual inputs (as the most relevant sensory information for spatial representation) and neglected the auditory inputs. However, a qualitative comparison between our results and those of past works using the same methodology but with unimodal conditions (visual stimuli: Amadeo et al., 2020a; auditory stimuli: Amadeo et al., 2019a; Campus et al., 2019, 2017) showed a larger neural response to our multisensory stimuli than to the unisensory stimuli of these previous studies, suggesting that the activation we revealed was likely related to multisensory processing. Future investigations could implement computational descriptions of unisensory and multisensory data at the cortical level, for example within a Bayesian model.

## **2.2 Study 2: the influence of the kind of spatial representation on the recruitment of occipital areas**

Vision is the most accurate sense for spatial representation since the visual modality enables the simultaneous acquisition of environmental properties and extracts the most detailed information about the surroundings (Thinus-Blanc and Gaunet, 1997; Tinti et al., 2006). In addition, it shapes other sensory modalities in creating an integrated spatial percept (King, 2009, 2014). The connection of the visual modality with space perception is expressed also by the recruitment of visual cortices for building fine spatial representations through different sensory modalities. Indeed, visual areas sustain the processing of multisensory stimuli involved in various spatial tasks, such as in the ventriloquist effect (where vision dominates the localization of spatially incongruent audiovisual stimuli; Busse et al., 2005; Gondan et al., 2005; Teder-Sälejärvi et al., 2005), in the tactile discrimination of stimuli orientation and motion (Hagen et al., 2002; Sathian, 2005; Zangaladze et al., 1999), and in the spatial bisection task (previously described in Section 2.1 as a layout that requires the estimation and comparison of different locations in space; Campus et al., 2019, 2017). In addition, in the posterior scalp, an auditory-evoked contralateral occipital positivity (ACOP) was found, appearing between 200-450 ms after the onset of a lateralized task-irrelevant sound (Feng et al., 2014; Matusz et al., 2016; McDonald et al., 2013; Retsa et al., 2020). Overall, these findings provided evidence of a cortical system of space perception that is recruited by different sensory modalities during various kinds of spatial representation. Further knowledge about spatial circuits derives from the study of blindness as a model to investigate the connection between vision and space perception. Auditory and somatosensory modalities were found to colonize the visual cortex of blind individuals, whose cortical reorganization was often associated with enhanced auditory and tactile skills (Amedi et al., 2007; Battal et al., 2022; Collignon et al., 2011, 2009; Fieger et al., 2006; Gougoux et al., 2005; Kujala et al., 1995; Renier et al., 2014; Ricciardi et al., 2014; Röder et al., 2022; Sadato, 2006). Parallel to these evidences, other studies showed that visual disabilities negatively influence the performance of blind individuals in more elaborate spatial processing, such as the acoustic spatial bisection task (Amadeo et al., 2019a; Campus et al., 2019; Gori et al., 2020b, 2014; Tonelli et al., 2020), the tactile spatial reasoning (Bertonati et al., 2020), and the auditory spatial memory (Setti et al., 2018, 2022a,b). Findings on the spatial abilities of visually impaired individuals seem contradictory, however, they may also simply indicate that the effect of visual deprivation on space perception is not uniform, since the brain processes the spatial domain differently according to the kind of task. The reason why the support of

the visual modality is fundamental for certain types of spatial tasks and less for others could be explained by the cross-sensory calibration theory (Gori, 2015). This theory, by assuming that during childhood vision (as the most accurate sense to process spatial information) calibrates audition and touch for some spatial skills, but not for others, suggests that visual calibration could be necessary for processing complex spatial information (such as the spatial bisection; Campus et al., 2019; Gori et al., 2020b), but less crucial for encoding other spatial configurations (such as the spatial localization; Röder et al., 1999; Rohlf et al., 2020). This characterization may also define the cortical activation associated with different spatial tasks.

In light of these considerations, in the next sections we investigate whether different configurations of spatial inputs would underlie different neural activation of occipital regions. We hypothesize finding a stronger early occipital component during a spatial metric representation than a localization task since this activation was found to be fundamental for processing an auditory spatial bisection task for which visual calibration is necessary (Campus et al., 2017; Gori et al., 2022b), but not in blind people who showed impaired performance (Amadeo et al., 2019a; Campus et al., 2019; Gori et al., 2020b). Conversely, since the spatial localization is independent of visual calibration (Rohlf et al., 2020) and associated with higher behavioral performance and cortical activity in the blind population (Fieger et al., 2006; Röder et al., 1999), we expect to find, during this task, a more robust later occipital activation which was previously associated with spatial attention processing (Di Russo et al., 2002). To test this, we proposed to a group of healthy subjects an electrophysiological study involving audiovisual spatial bisection and localization tasks, which elicit two kinds of spatial representation of multisensory stimuli to compare at the cortical level (Bertonati et al., under review).

## 2.2.1 Methods

### Sample and experimental procedure

17 adults participated in the study (10 females, mean age  $\pm$  SD:  $24 \pm 3.08$  years old). All participants reported they had no history of neurological, cognitive, or sensory deficits and they gave written informed consent before testing. We decided sample size based on previous studies testing the neural correlates of spatial abilities (Amadeo et al., 2019a; Campus et al., 2019, 2017; Gori et al., 2022b). A priori power analysis (two-tailed t-test, estimated effect size Cohen's  $d = 0.80$ ,  $\alpha = 0.05$ ) determined a minimum sample size of 15 participants. The ethics committee of the local health service (Comitato etico, ASL 3 Genova) approved the study, which was conducted in line with the Declaration of Helsinki.

Participants sat at a distance of 180 cm from the center of a horizontal array of 23 speakers/LEDs spanning  $\pm 25^\circ$  of visual angle ( $0^\circ$  represented the central speaker/LED, aligned with the subject body midline, with negative values on the left and positive values on the right). Each speaker was spatially aligned to one LED. The setup presented audiovisual (AV) stimuli, each consisting of a single sound (60 dB SPL at ears' level, 500 Hz) simultaneously reproduced for 75 ms with a single red flash ( $2.3^\circ$  diameter,  $20 \text{ cd/m}^2$  luminance). Participants performed a spatial bisection task (same as the spatial bisection presented in Section 2.1) and a spatial localization task (Figure 2.7). In the spatial bisection task, each trial consisted of three AV stimuli (namely S1, S2, and S3) played at three different spatial positions and time lags. There was a fixed time interval of 1.5 seconds between S1 and S3. These two stimuli were always played at  $-25^\circ$  and  $+25^\circ$ , respectively, while S2 could be presented randomly from either  $-2.3^\circ$  or  $+2.3^\circ$  in space, and at either  $-250 \text{ ms}$  or  $+250 \text{ ms}$  in time (with  $0 \text{ ms}$  representing the middle of the  $1.5 \text{ s}$  temporal sequence). The time separation between S1, S2, and S3 was sufficiently large to ensure a complete decay of the ERP response between the stimuli.

In each trial, participants estimated whether S2 was spatially farther from S1 or S3. In the spatial localization task, each trial consisted of one AV stimulus (S) reproduced from either  $-2.3^\circ$  or  $+2.3^\circ$ . Participants' task was to localize S by identifying whether its position was more on the left ( $-2.3^\circ$ ) or on the right ( $+2.3^\circ$ ) than the center of the array ( $0^\circ$ ). For both tasks, subjects provided their answers after the stimuli presentation by pressing the appropriate button. S2 stimulus of the spatial bisection was identical to S stimulus of the spatial localization, thus the two tasks likely differed only in the kind of spatial representation elicited by the AV stimuli. Specifically, the spatial bisection elicited the construction of a

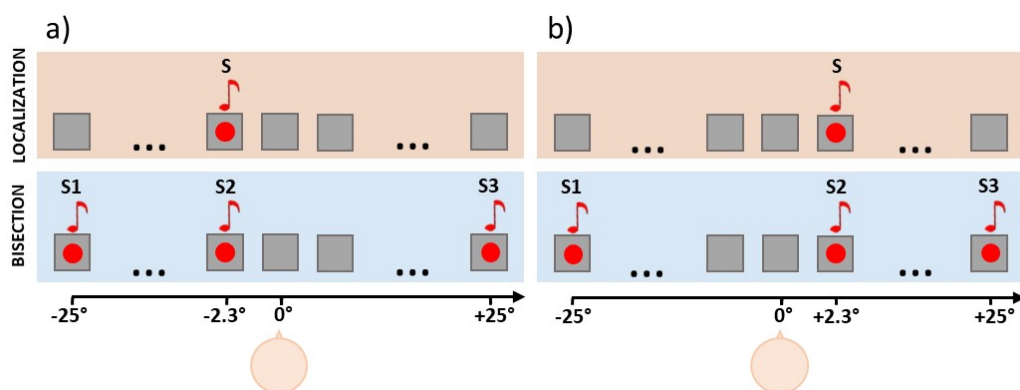


Figure 2.7 - Two experimental conditions according to S2/S positions in space, for the spatial localization task (orange rectangles above) and the spatial bisection task (blue rectangles below). a) S2 and S from  $-2.3^\circ$ ; b) S2 and S from  $+2.3^\circ$ .

spatial metric of S2 in relation to the position of S1 and S3, since participants, first, localized the AV sources in space and then connected the spatial locations to each other. Instead, the spatial localization required the identification of only S in space. However, it is worth noting that participants, also in the spatial bisection, could, in principle, evaluate the position of only S2 and not of S2 in reference to S1 and S3. Nevertheless, we excluded this possibility since a study using the same task but with S1 and S3 positions that varied across trials showed similar results to paradigms using fixed S1 and S3 as in our experiment (Aggius-Vella et al., 2020). This observation suggests that S2 of the spatial bisection and S of the spatial localization were presumably processed differently by the participants.

The spatial bisection and localization tasks were counterbalanced across subjects, in two separate blocks, and each block consisted of 240 trials and 15 catch trials (in which S2 and S were reproduced at 0° and at 0 ms, to test for participants' stereotypical responses). Participants were asked to maintain a stable head position that was continuously monitored by the experimenter, together with the electrooculogram (EOG) signal.

### **EEG data collection, pre-processing, and data analysis**

We recorded EEG and EOG (two electrodes positioned at the left and right outer canthi to check horizontal ocular movements) data with 64 active electrodes using the Biosemi ActiveTwo EEG System. We kept electrode offsets under 35 mV, and we applied a first-order analog anti-aliasing filter with a half-power cutoff at 3.6 kHz and then down-sampled data at 512 Hz (acquisition at 2048 Hz, with a decimation factor of 1/4) with pass-band from DC to 134 Hz. EEG recording was referenced to a Common Mode Sense (CMS) active electrode and a Driven Right Leg (DRL) passive electrode, which played as the ground electrodes used in conventional systems. EEG was filtered between 0.1 and 100 Hz. We applied the Artifact Subspace Reconstruction (ASR) method, available as a plug-in for EEGLAB software (Delorme and Makeig, 2004), to remove transient stereotypical (e.g., eye blinks) and non-stereotypical (e.g., movement or muscle bursts) high-amplitude artifacts. Our study selected a 500 ms sliding window and a threshold of 3 standard deviations to identify corrupted subspaces. Additionally, we removed channels that posed an inferior correlation with other channels (more than 0.85) or those with line noise relative to its signal presenting more than 4 standard deviations based on the total channel population. Finally, whenever the fraction of contaminated channels exceeded the threshold of 0.25, we removed time windows. EEG data was further cleaned by Independent Component Analysis (ICA) with two EEGLAB toolboxes, namely SASICA (Chaumon et al., 2015) and IC\_MARC (Frølich et al., 2015), keeping all parameters as their default. Finally, we rejected components based

on abnormal topographies and/or spectra and we referenced data to the average of the left and right mastoids (TP7, TP8 electrodes).

The ERPs analysis focused separately on the neural responses to S2 for the spatial bisection task and to S for the spatial localization tasks, as the two stimuli were physically identical but elicited two different kinds of spatial representation. In this way, we could draw conclusions about the neural modulation of the experimental question. Thus, we obtained the ERPs by averaging EEG data synchronously with the S2 and S onsets. We considered as baseline a time window of 200 ms before S1 onset for the spatial bisection and 200 ms before S onset for the spatial localization. The analysis focused on electrodes related to visual processing (O1, O2 in occipital areas), given that past studies revealed neural responses specific to spatial tasks at these sites (Campus et al., 2019, 2017; Gori et al., 2022b). Moreover, the mean ERP amplitudes after S2 and S onsets were computed in three time windows selected to explore some main ERP processing stages of occipital areas (Di Russo et al., 2002; Hillyard and Anllo-Vento, 1998). The three time windows are: *i*) between 50 and 90 ms (corresponding to the visual C1 component), *ii*) between 110 and 160 ms (encompassing the visual P100 (80-130 ms) and P140 (110-160 ms) components), *iii*) between 250 and 450 ms (reflecting more cognitive mechanisms). For each time window, we collapsed average ERP waveforms across conditions and hemispheres of recording to obtain ERPs recorded on the contralateral and the ipsilateral hemisphere with respect to stimulus position in space and therefore calculate lateralized ERP responses as the difference between the contralateral and ipsilateral ERP recordings. By performing three separate t-tests, we compared the cortical activity between the spatial bisection and the spatial localization tasks within the three selected time windows. We also performed a similar sensor level analysis on central (C1, C2 electrodes) and temporal areas (T7, T8 electrodes) in the three time windows. Finally, behavioral performance was computed as the percentage of correct responses for each task. Task difficulty was also calculated as the difference between spatial localization and spatial bisection behavioral performance. To check whether task difficulty played a role in the neural activation associated with the two tasks, for each time window we performed an ANCOVA on the mean ERP occipital responses, with Task as within factor and Task difficulty as covariate.

In the end, we performed distributed source analysis with Brainstorm software (Tadel et al., 2011) to provide evidence that the components we observed over the occipital scalp involved generators in the visual areas. For this analysis, we re-referenced data to the common average, we used a standard 1 mm resolution template of the Montreal Neurological Institute (non-linear average of 152 subjects, processed with FreeSurfer 5.3 ICBM152, Fonov

et al., 2009), we let the dipole orientations to be free to assume whichever (unconstrained) orientation, to avoid misleading over-interpretation in lack of individual MRI scans, and we performed the EEG forward modeling using a three-layer (head, outer and inner skull) BEM generated with OpenMEEG86. We estimated source intensities using the sLORETA approach (Gramfort et al., 2011), a technique that has been robust to noise in EEG recordings and head model approximations. We kept Brainstorm's default parameter settings for both source reconstruction and BEM creation. Source activation for each subject and condition within the three selected time windows (50-90 ms, 110-160 ms, 250-450 ms) after S2 and S onsets was separately averaged, and, subsequently, the norm of the vectorial sum of the three orientations at each vertex was estimated. Finally, paired t-tests for pairwise comparisons (corrected for multiple comparisons with the FDR method and using  $p = .00001$  as a threshold) verified the modulation of the occipital activation during the spatial bisection and the spatial localization tasks. Specifically, we compared the neural response between the two tasks, separately considering the stimulus positions in space ( $\pm 2.3^\circ$ ).

## 2.2.2 Results

We compared the occipital response between the spatial bisection and the spatial localization task within three separate time windows after S2 and S onsets. In the 50-90 ms time window (Figure 2.8a), results revealed an occipital positivity contralateral to the stimulus position in space (i.e., either  $-2.3^\circ$  or  $+2.3^\circ$ ) which was significantly stronger during the spatial bisection task than during the spatial localization task (Figure 2.8b;  $t_{(16)} = 7.39$ ,  $p < 0.001$ , Cohen's  $d = 1.79$ , 95% CI = [0.99, 2.59]). In the 110-160 ms time window (Figure 2.8a), S2 of the spatial bisection task elicited a bilateral occipital response recalling a later phase of the P1 (i.e., P140, peak latency of 146 ms), while S of the spatial localization task evoked a contralateral occipital activation resembling an earlier phase of the P1 component (peak latency of 110 ms). Paired t-test showed that the neural response in this time window was significantly greater during the spatial localization than during the spatial bisection (Figure 2.8b;  $t_{(16)} = -19.85$ ,  $p < 0.001$ , Cohen's  $d = -4.81$ , 95% CI = [-6.56, -3.06]). Finally, in the 250-450 ms time window (Figure 2.8a), no significant difference was found in the occipital response between the two spatial tasks (Figure 2.8b;  $t_{(16)} = 1.56$ ,  $p = 0.138$ , Cohen's  $d = 0.37$ , 95% CI = [-0.13, 0.88]). Statistical comparisons at temporal sites (electrodes T7, T8; Figure 2.9a) did not reveal any significant difference between the spatial bisection and spatial localization task in any of the time windows considered (50-90 ms:  $t_{(16)} = 0.04$ ,  $p = 0.964$ , Cohen's  $d = 0.01$ , 95% CI = [-0.48, 0.50]; 110-160 ms:  $t_{(16)} = -0.38$ ,  $p = 0.705$ , Cohen's  $d = -0.09$ , 95% CI =



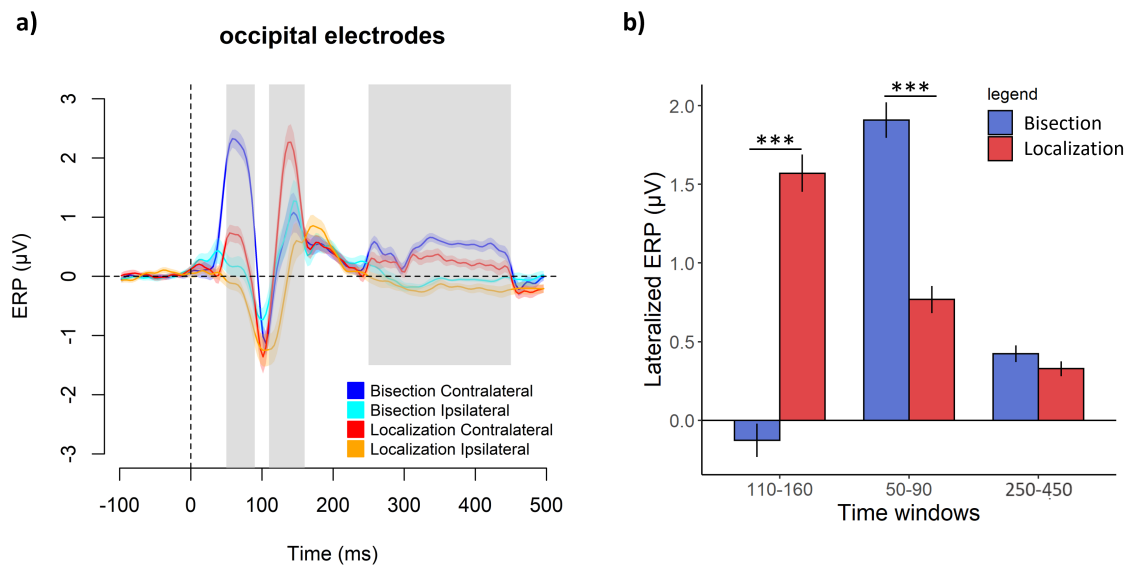


Figure 2.8 - Occipital ERPs during the spatial bisection and the spatial localization. a) Contralateral and ipsilateral occipital ERPs (mean  $\pm$  SEM) with respect to S2 during the spatial bisection task (blue and light blue curves) and with respect to S during the spatial localization task (red and orange curves). The gray-shaded areas delimit the three time windows: 50-90 ms, 110-160 ms, and 250-450 ms. b) Lateralized mean ERP amplitude, calculated as the difference between the contralateral and ipsilateral ERP occipital responses in the 50-90 ms, 110-160 ms, and 250-450 ms time windows separately, for the spatial bisection (blue bars) and spatial localization (red bars). Error bars indicate SEM. \*\*\*  $p < 0.001$ .

[-0.58, 0.40]; 250-450 ms:  $t_{(16)} = 0.70$ ,  $p = 0.489$ , Cohen's  $d = 0.17$ , 95% CI = [-0.32, 0.66]). Similar results emerged in the comparison of the two tasks at central sites (electrodes C1, C2; Figure 2.9b; 50-90 ms:  $t_{(16)} = -1.87$ ,  $p = 0.079$ , Cohen's  $d = -0.45$ , 95% CI = [-0.97, 0.06]; 110-160 ms:  $t_{(16)} = 1.66$ ,  $p = 0.115$ , Cohen's  $d = 0.40$ , 95% CI = [-0.11, 0.91]; 250-450 ms:  $t_{(16)} = -0.79$ ,  $p = 0.436$ , Cohen's  $d = -0.19$ , 95% CI = [-0.69, 0.30]). The observation that cortical activity was not influenced by the kind of task in temporal and central regions, but it was in the occipital areas, suggested that especially the latter (which play a specific role in the processing of the space domain) were subject to modulation by the kind of spatial representation. The scalp topographies of the mean ERPs in the 50-90 ms, 110-160 ms, and 250-450 ms time windows also show these results (Figure 2.10).

We provided further evidence of the cortical generators of the ERP components during the two spatial tasks, by performing statistical comparisons in a source-level analysis (Figure 2.11). In the 50-90 ms time window, a paired two-tailed t-test revealed significantly more pronounced recruitment of contralateral occipital areas during the spatial bisection than

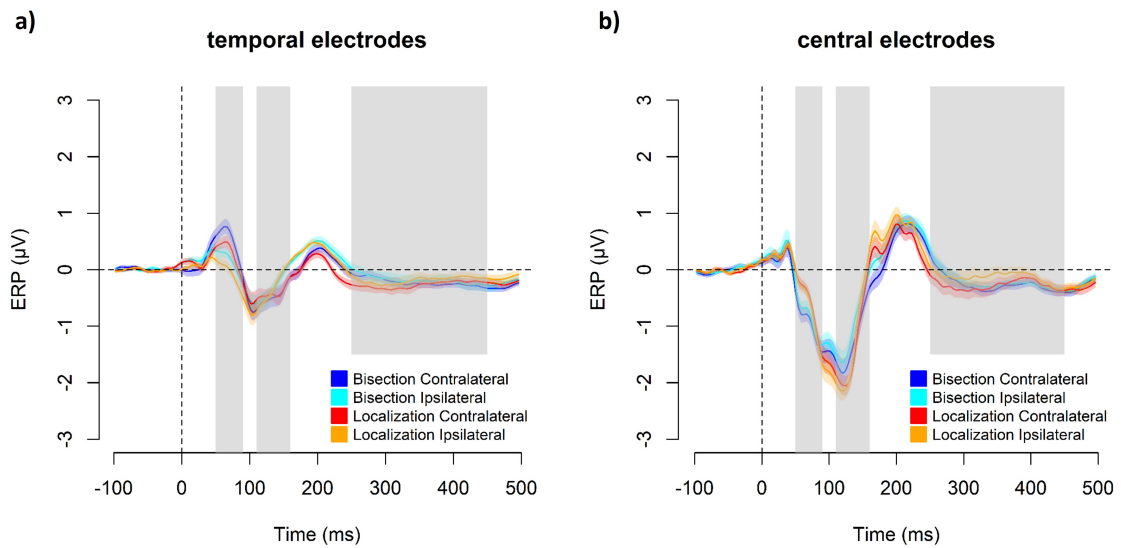


Figure 2.9 - Contralateral and ipsilateral temporal (a) and central (b) ERPs (mean  $\pm$  SEM) in respect to S2 during the spatial bisection task (blue and light blue curves) and in respect to S during the spatial localization task (red and orange curves). In both panels, the gray-shaded areas delimit the three time windows: 50-90 ms, 110-160 ms, 250-450 ms.

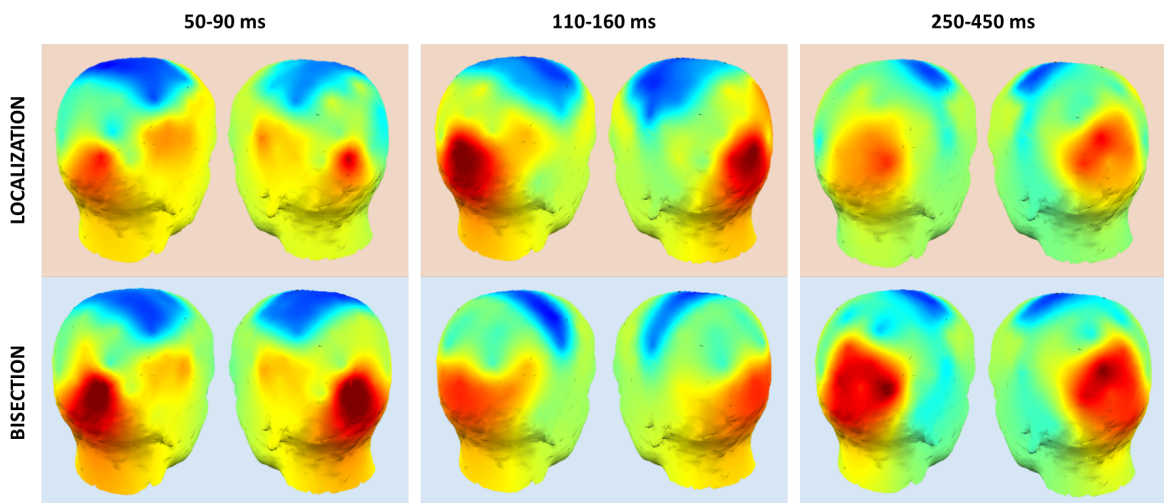


Figure 2.10 - Scalp maps of the mean ERP amplitude in the 50–90 ms (left column), 110-160 ms (central column), and 250-450 ms (right column) time windows after S2 and S onsets, for the spatial bisection (lower row) and the spatial localization (upper row), respectively.

during the spatial localization task. On the contrary, within the 110-160 ms time window, the contralateral occipital activation was greater during the spatial localization. Finally, within

### SPATIAL BISECTION > SPATIAL LOCALIZATION

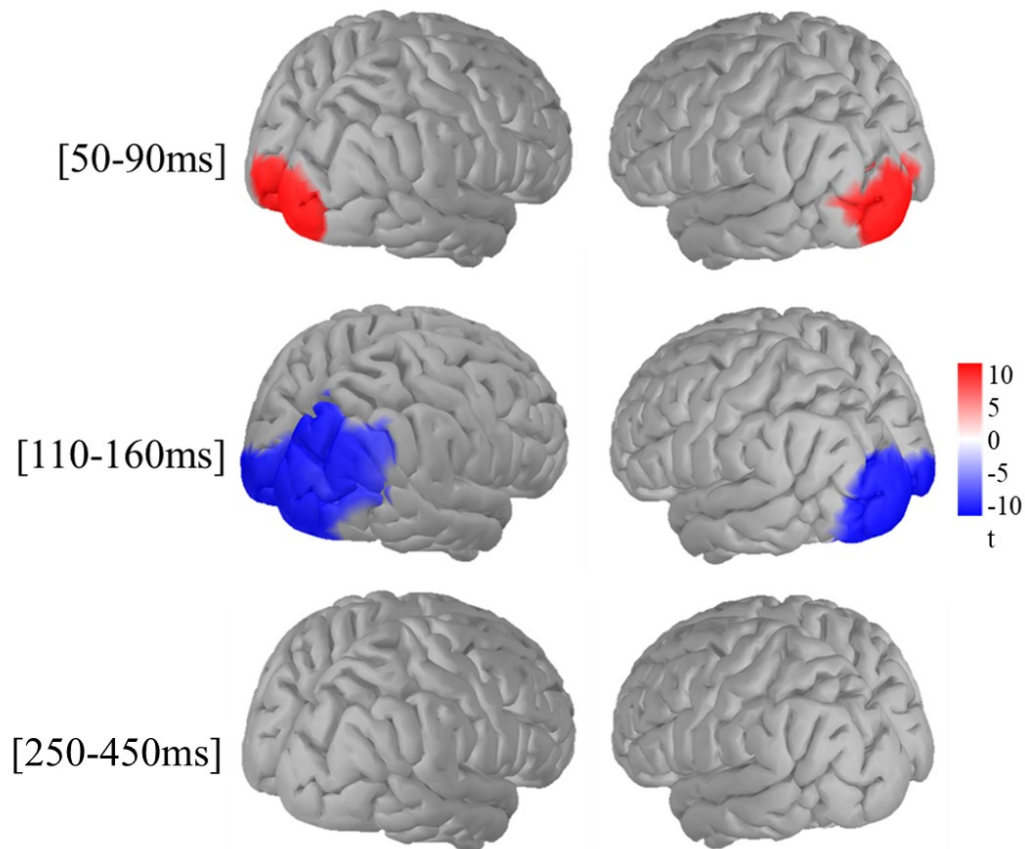


Figure 2.11 - Maps of the paired two-tailed t-tests run on the average source activity in the 50-90 ms (upper line), the 110-160 ms (middle line), and the 250-450 ms (lower line) time windows. Reddish and bluish colors indicate stronger activation during the spatial bisection and the spatial localization task, respectively. The significance of the difference (i.e., the magnitude of  $t$ ) is represented by color intensity. The left and right columns show the conditions in which S/S2 were reproduced at  $-2.3^\circ$  or  $+2.3^\circ$ , respectively.

the 250-450 ms time window, the statistical comparison between the two tasks was not significant in any cortical area.

At the behavioral level, participants correctly localized significantly more AV stimuli in the spatial localization task than in the spatial bisection task (mean  $\pm$  SD percentage of localization:  $98.18 \pm 2.17$ ; mean  $\pm$  SD percentage of bisection:  $91.76 \pm 7.93$ ;  $t_{(16)} = 4.07$ ,  $p < 0.001$ , Cohen's  $d = -0.98$ , 95% CI =  $[-1.59, -0.38]$ ). ANCOVA revealed that the main effects of Task remained under control for the task difficulty in the occipital sites for the time windows considered (50-90 ms: Task:  $F_{(1,15)} = 53.61$ ,  $p < 0.001$ ; Task difficulty:  $F_{(1,15)} = 4.67$ ,  $p = 0.047$ ; Task \* Task difficulty:  $F_{(1,15)} = 0.68$ ,  $p = 0.42$ . 110-160 ms: Task:  $F_{(1,15)} =$

377.94,  $p < 0.001$ ; Task difficulty:  $F_{(1,15)} = 14.74$ ,  $p = 0.001$ ; Task \* Task difficulty:  $F_{(1,15)} = 0.34$ ,  $p = 0.566$ . 250-450 ms: Task:  $F_{(1,15)} = 2.52$ ,  $p = 0.133$ ; Task difficulty:  $F_{(1,15)} = 2.61$ ,  $p = 0.127$ ; Task \* Task difficulty:  $F_{(1,15)} = 1.58$ ,  $p = 0.228$ ). These results allowed us to rule out the possibility of an influence of the task difficulty on the occipital modulation between the two tasks.

### 2.2.3 Discussion of Study 2

In Study 2, the recruitment of occipital areas in response to different kinds of spatial representation of multisensory stimuli was investigated. Past studies have revealed evidence of a cortical spatial representation system, where neural responses occurred at different levels of sensory processing and within several cortical areas (Gori et al., 2020a, 2022b; Molholm et al., 2002; Retsa et al., 2020). These findings must be related to the fact that external stimuli can be localized in coordinates that describe the spatial relations among inputs, or in coordinates that refer only to the position in space of the perceiving subject. Study 2 depicted this dichotomy at the cortical level, by recording EEG data during a spatial bisection and a spatial localization task. Results indicated that early occipital components were associated with a spatial representation that required a metric definition of the spatialized stimuli, while later occipital components with the localization of a single position in space. Specifically, an early occipital activation (50-90 ms) contralateral to the AV stimulus location was stronger when participants encoded the stimuli in the spatial bisection task, while in a later time window (110-160 ms), the neural modulation required a greater occipital response for the spatial localization. No cortical modulation due to the task was detected in the 250-450 ms time window suggesting that higher-level information processing does not take into account the different kinds of spatial representation.

Previous studies showed similar patterns of activation in the time ranges considered in our study, by investigating the cortical responses associated with multisensory spatial tasks (Gori et al., 2022b; Molholm et al., 2002). However, this is the first time that two spatial tasks were directly compared to each other at the cortical level. Indeed, the tasks we used relied on different spatial processing: while in the spatial bisection task participants estimated the relative position of a bimodal stimulus in relation to other spatialized stimuli, in the spatial localization task they localized the position of only one stimulus in space. Past research has already shown that subjects' goals can influence multisensory processes in some behavioral and neural responses (Talsma et al., 2007; ten Oever et al., 2016). Specifically, a cortical circuit designated to account for the complexity of the factors involved in multisensory

processing has been described (reviewed in Murray et al., 2016a). In Study 2, a similar pattern of activation may emerge for the spatial representation, with occipital recruitment that is modulated by the subject's goal, in this case by the kind of spatial representation a person builds.

For Study 2, we mention as limitations *i*) the lack of unimodal conditions (only auditory and only visual), which (similarly to Study 1) may limit the conclusion that results represented multisensory integration mechanisms, *ii*) the presence of S1 in the spatial bisection but not in the spatial localization, which should be considered in interpreting the cortical difference between the two tasks.

*Chapter 2 is partially extracted and adapted from Gori, M., Bertonati, G., Campus, C., and Amadeo, M. B. (2023). Multisensory representations of space and time in sensory cortices. Human Brain Mapping, 44(2), 656–667. <https://doi.org/10.1002/hbm.26090>*

## Chapter 3

# Object motion perception and interception

In Chapter 1, I introduced the crucial role of multisensory experience in developing a coherent percept of the world, especially of vision and audition that are essential to scaffold spatial and temporal representations, respectively. I also illustrated how space and time are interrelated dimensions that communicate to form a shared magnitude system, as theorized by the ATOM model (Buetti and Walsh, 2009; Lambrechts et al., 2013; Walsh, 2003). In Chapter 3, I further explore these domains and their interaction by focusing on motion processing intended as, on the one side, motion perception in Section 3.1, and on the other side, the manual interception of moving objects in Section 3.2.

Space and time domains are tightly coupled in object motion. For humans and animals, the ability to track moving stimuli in the surrounding environment is of fundamental importance to predict movement outcomes and guide their actions accordingly. Motion perception can be expressed through all our sensory modalities. Just consider the importance of perceiving an approaching car in the street, not only by using vision but also by hearing the engine roar if the vehicle is coming from behind a corner. Over the years, motion perception has experienced a progressive rise in investigation by the neuroscientific community. Research went from exploring the motion perception sense by sense (Burr and Thompson, 2011; Nishida et al., 2018; Pei and Bensmaia, 2014; Warren et al., 2002) to investigating how the brain perceives an object in motion when this is conveyed by multiple sensory modalities, finding in the latter a multisensory gain in terms of precision and response speed (Bentvelzen et al., 2009; Wuerger et al., 2003b, 2010). At the neural level, many structures in the visual system were early found to contain motion-responsive neurons, especially in the occipitotemporal area V5/MT, in humans and primates (Ungerleider and Desimone, 1986; Watson et al., 1993; Zihl

et al., 1991). Less is known for what concern the auditory and the somatosensory system, but neural responses to the acoustic motion were revealed in the planum temporale, the parietal cortices, and the right insula (Baumgart et al., 1999; Pavani et al., 2002), and to the tactile motion in the primary and secondary somatosensory areas (Hagen et al., 2002). The fact that motion perception undergoes a specialized treatment by the brain (probably due to its crucial role in survival) is also demonstrated by the brain imputing motion at a constant velocity to static stimuli presented in sequence (Oliveri and Koch, 2009). Specifically, we refer to the Kappa and Tau effects (Bill and Teft, 1969; Collyer, 1977; Jones and Huang, 1982; Sarrazin et al., 2004). These effects were first revealed in experiments presenting three light bulbs in sequence in a row forming two spatio-temporal intervals. Participants' time judgments increased as a function of the spatial distance between the stimuli (Kappa effect), while participants' space judgments increased as a function of temporal separation between the stimuli (Tau effect). These effects were interpreted as the result of imputed constant velocity by the participants, although there was no actual or phenomenal motion in the stimuli.

The association between the space and time domains was also described as the result of humans performing actions in the surrounding environment (Binetti et al., 2015; Walsh, 2003). Action can be defined as goal-directed behaviors that select and use relevant sensory information to achieve some desired outcomes, also on the basis of predictive mechanisms (Brenner et al., 2019; Dickinson and Balleine, 2000). This definition includes the person's manual responses, which are the specific topic of Section 3.2 and Chapter 4. Coming back to interaction with moving objects, the manual interception of items in motion is one of the ways in which we implement manual responses in everyday situations. The term "interception" is referred to any body movement directed toward a moving object with the intention of catching it (Brenner and Smeets, 2018). The moving object is considered intercepted when it is stopped or hit or kicked to prevent it from moving in a certain direction. A wide variety of tasks have been used to study the interception of moving objects, real or virtual, such as catching balls, grasping rolling objects, hitting targets moving on a screen, etc. (Bootsma and van Wieringen, 1990; Brenner and Smeets, 2015; Carnahan and McFadyen, 1996; De La Malla et al., 2017; Savelsbergh et al., 1992). Overall, these studies highlighted that interception is a challenging task that requires concurrently considering the task demands and one's own movement limitations (Brenner and Smeets, 2018). For example, when hitting a moving target, the brain is able to consider the presence of obstacles on the path by veering away from them without affecting the interceptive performance (Brenner and Smeets, 2007). Thus, although there are many ways in which one could move to intercept a target, the sensorimotor system seems to be flexible and sophisticated enough to select the best choice

for performing the task successfully. In addition, the system's flexibility is influenced by the presence or the lack of perceptual feedback of previous interceptions (Brenner et al., 2002; Smeets and Brenner, 2004).

Given the crucial tight relationship between space and time in motion perception and interception and the great flexibility of the sensorimotor system in the accomplishment of task demands, in Chapter 3 I explore the use of spatio-temporal cues in the interaction with moving objects. First, in Study 3 (Section 3.1), I investigate the role of visual experience in the relationship between the space and time domains in the acoustic motion perception of early blind and sighted individuals. Secondly, in Study 4 (Section 3.2), I explore the ability of the human sensorimotor system to adapt itself to the spatio-temporal specific demands of an interception task.

### **3.1 Study 3: the role of visual experience in the use of spatio-temporal cues in the acoustic motion processing**

Acoustic motion may be described as a change of sound location over time. Given that the auditory system was revealed to have a superior ability for treating temporal information (Burr et al., 2009), one may wonder whether temporal aspects of acoustic motion have a peculiar role in the perception of moving objects. In this regard, Freeman et al. (2014) used a discrimination contours technique (a method previously used to differentiate the perceptual components underlying color vision; Champion and Freeman, 2010; Gegenfurtner and Hawken, 1995; Noorlander et al., 1980; Poirson et al., 1990; Reisbeck and Gegenfurtner, 1999; Wardle and Alais, 2013) to determine how different object's motion cues (i.e., duration, distance, and speed) were processed and integrated during acoustic motion perception. Authors showed that in the auditory motion perception temporal cues dominated since listeners were more sensitive in processing changes in the duration of the sound than changes in the spatial displacement traveled by the moving object (distance) or in the ratio between space and time components (speed) (Freeman et al., 2014). Thus, these results suggested that speed is not a primary component in the auditory system and challenged the existence of acoustic motion-sensitive detectors like those found in vision. Specifically, the visual system holds velocity detectors at the neural and perceptual level (similar to direction-tuned sensors; Albright, 1984; Ben-Yishai et al., 1995) that lead observers to process visual moving stimuli based on the ratio between space and time, i.e., the speed of the stimulus (Gegenfurtner and Hawken, 1995; Reisbeck and Gegenfurtner, 1999; Wardle and Alais, 2013). On the



contrary, the auditory system seems to lack low-level velocity detectors and to compute auditory motion based on other acoustic cues, such as the interaural-time differences (ITDs; differences in the arrival of a sound at the two ears), the interaural-level differences (ITLs; differences in the sound pressure levels in the two ears) and the filtering properties of the pinna (Chaplin et al., 2018; Middlebrooks, 2015; Middlebrooks and Green, 1991).

Given the different contributions of single sensory modalities in motion perception, one might wonder what happens to this ability in the absence of one sense, for example in the case of blindness. Acoustic motion perception becomes particularly important for people with visual disabilities. I already introduced in Chapter 1 and Chapter 2 the effects of visual impairment on the two domains of representation underlying motion perception, i.e., space and time. Specifically, I described that in blindness, *i*) good performance is maintained in the time perception of blind people (Gori et al., 2014; Hötting and Röder, 2009; Stevens and Weaver, 2005; Van Der Lubbe et al., 2010; Vercillo et al., 2016); *ii*) controversial results emerge in the space perception of blind individuals, with enhanced auditory abilities in some spatial tasks (Battal et al., 2020; Lessard et al., 1998; Röder et al., 1999; Voss et al., 2004), and compromised skills in other and more complex spatial tasks (Gori et al., 2014; Vercillo et al., 2016; Voss et al., 2015; Zwiers et al., 2001); *iii*) blind people use the temporal properties of an acoustic event to infer its spatial information (Amadeo et al., 2020b; Gori et al., 2018, 2020b). For what concern auditory motion perception in case of blindness, literature shows conflicting results. On the one side, blind individuals were found to be able to detect the motion of a sound source (Bedny et al., 2010; Dormal et al., 2016; Jiang et al., 2016, 2014; Poirier et al., 2004, 2006; Wolbers et al., 2011) and to have enhanced abilities in judging the direction of moving sounds (Lewald, 2013). On the other side, blind people were impaired in processing a more complex motion structure (e.g., tracking a sound moving in different trajectories within a circle) (Finocchietti et al., 2015). Therefore, it is still unclear if the absence of vision promotes or negatively affects auditory motion perception and, specifically for the topic covered in this chapter, if blindness influences the use of motion cues (duration, distance, and speed) by the auditory system. Indeed, a cortical reorganization may occur with the visual speed detectors revealed in the sighted individuals being recruited by audition in the blind population (leading to a similar or better performance compared to sighted people). Alternatively, a lack of cortical plasticity may force blind individuals to use acoustic motion cues with impaired processing of spatial cues. Given the difficulties of visually impaired people with certain spatial information and the specialized support they receive from the temporal domain (Amadeo et al., 2020b; Gori et al., 2018, 2020b, 2014; Vercillo et al., 2016;

Voss et al., 2015; Zwiers et al., 2001), in Study 3 we leaned more towards the second point of view.

Specifically, Study 3 investigated how spatial and temporal information is used for sound speed discrimination in the lack of vision from birth (Bertonati et al., 2021). To this end, ten early blind and ten blindfolded sighted participants performed a speed discrimination task in which spatial, temporal, or both cues were available to infer moving sounds' velocity. We hypothesized that visual deprivation would influence the use of spatial and temporal cues underlying acoustic speed perception, by promoting the blind individuals' preference towards the processing of temporal aspects of an acoustic event.

### 3.1.1 Methods

#### Sample and experimental procedure

Ten early blind participants (EB, 6 females, mean age  $\pm$  SD: 38.22  $\pm$  11.39 years old) and ten age-matched sighted controls (SC, 5 females, mean age  $\pm$  SD: 38.99  $\pm$  10.83 years old) took part in the experiment. Visual impairment of EB individuals was attributed to peripheral deficits of the visual system and blindness was total (except for light perception in three participants and light and shape perception in two; for clinical details, see Table 3.1). None of the participants reported additional sensory, cognitive, and/or neurological deficits. All individuals gave written informed consent before starting the test, in accordance with the ethics committee of the local health service (Comitato etico, ASL3 Genovese) and the Declaration of Helsinki.

Participants were seated 2 m away from the center of an array of 24 speakers spanning 50° of visual angle. Auditory stimuli consisted of moving white noises (80 dB at the ears' level). We simulated the acoustic motion by reproducing sequential white noises along with the array of speakers and modulating the volume of the sounds. Acoustic stimuli could move, in a randomized order, either rightward or leftward, and, for each motion direction, they could randomly begin from the most extreme speaker of the array or from the two speakers immediately next to it. In each trial (Figure 3.1), the array reproduced a reference sound (R) moving at a fixed speed (26.47°/s; traveled distance: 21.17°; duration: 800 ms), and a target sound (T) moving at different speeds trial by trial (from 6.62°/s to 105.86°/s). T speeds were computed by manipulating the stimulus duration and traveled distance so that T differed from R by given proportions (Weber fraction, W) of R duration and distance. In particular,  $W_{distance} = \frac{(T_{distance} - R_{distance})}{R_{distance}}$  and  $W_{duration} = \frac{(T_{duration} - R_{duration})}{R_{duration}}$ , where distance refers to the

Table 3.1 Clinical details of early blind participants

Participants	Age	Gender	Pathology	Blindness onset	Residual vision
EB1	48	M	Atrophy optic nerve	Birth	No vision
EB2	22	F	Retinopathy of prematurity	Birth	Light and shadows
EB3	32	F	Retinitis pigmentosa	Birth	1/20 visual acuity. <1% visual field. Light and shapes.
EB4	31	F	Retinopathy of prematurity	Birth	No vision
EB5	44	F	Retinitis pigmentosa	Birth	1/50 visual acuity. Light and shapes.
EB6	33	F	Retinopathy of prematurity	Birth	No vision
EB7	28	M	Leber amaurosis	Birth	No vision
EB8	55	M	Retinopathy of prematurity	Birth	Light and shadows
EB9	54	M	Unknown	Birth	No vision
EB10	32	F	Retinitis pigmentosa	Birth	Light and shadows

spatial displacement of the stimulus along with the array (in degrees) and duration to the stimulus’s temporal length (in ms).

The resulting set of T speeds is shown in the distance-duration plane in Figure 3.2. From each orientation of the plane (i.e., the eight lines of the plane), we used a constant stimuli method and selected nine target speeds (each repeated six times for a total of 432 trials).

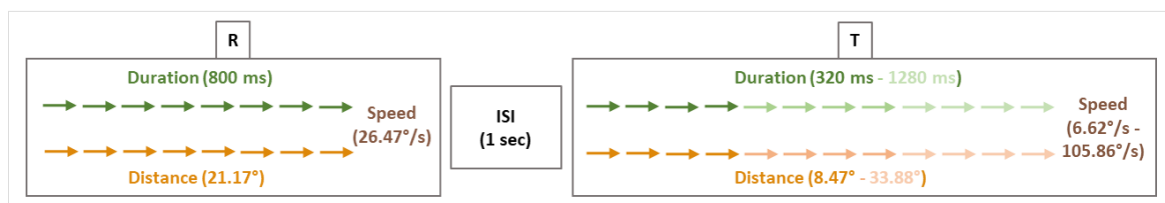


Figure 3.1 - Illustration of a trial. The reference sound (R) moved at the fixed speed of 26.47°/s, and the target sound (T) at faster or slower speeds according to the manipulation of the reference distance and/or duration. R and T were always separated by a 1-second interstimulus interval (ISI). *Reproduced with permission. Bertonati et al. (2021). Auditory speed processing in sighted and blind individuals. Plos one, 16(9), e0257676.*

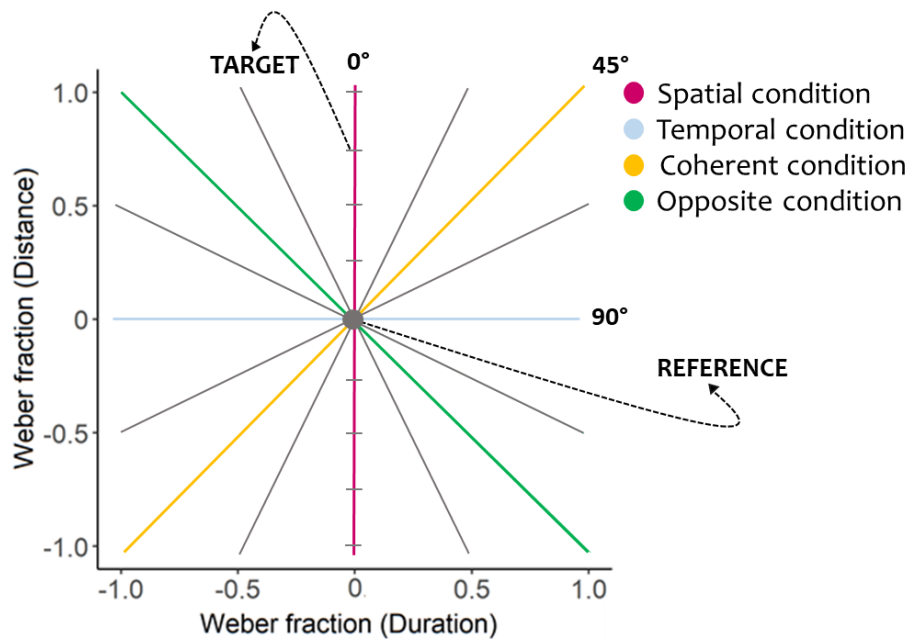


Figure 3.2 - Distance-duration plane. The eight lines composing the plane correspond to eight orientations. Along with each orientation, the T sound (here selected as an example) differed from the R sound (the grey dot) by a given proportion of spatial (distance) and temporal (duration) cues. Among the eight orientations, the colored lines correspond to the four experimental conditions (spatial, temporal, coherent, and opposite). 0°, 45° and 90° correspond to the possible orientation of participants' behavioral performance explained in more detail in Figure 3.3. *Reproduced with permission. Bertoni et al. (2021). Auditory speed processing in sighted and blind individuals. Plos one, 16(9), e0257676.*

Among the orientations of the distance-duration plane, we selected four experimental conditions (colored lines of the plane): *i*) a spatial condition where only distance cues were manipulated, so that when distance increased, speed increased; *ii*) a temporal condition where only duration cues were manipulated, so that when duration increased, speed decreased; *iii*) a coherent condition where distance and duration cues were manipulated with a directly proportional relationship, so that when distance increased, duration increased proportionally and speed was constant (i.e., T speed was equal to R speed); *iv*) an opposite condition where distance and duration cues were manipulated with an inversely proportional relationship, so that when distance increased, duration decreased and speed increased (i.e., speed was fast for short durations/long distances and slow for long durations/short distances). These conditions allowed us to evaluate how participants discriminated the acoustic speed based on spatial (traveled distance) and temporal (duration) cues of the moving sounds.

Participants performed a two-alternative forced-choice (2AFC) task in which, for each trial, they determined whether T was moving faster or slower than R, by pressing one of two corresponding keys. According to the experimental condition (spatial, temporal, coherent, or opposite), participants could rely on spatial, temporal, or both cues to identify T speed. No feedback was given to participants during the task. The experiment was divided into 6 blocks with randomly selected target speeds. We asked participants to keep their heads steady while performing the task, and we constantly monitored the subjects' head positions throughout the experiment. Sighted individuals and the two participants with residual light and shape perception (see Table 3.1 for details) were blindfolded before entering the room. This ensured that they were not influenced in their performance by having seen the experimental setup.

### Data analysis

The discrimination contours technique was used to disentangle the contribution of spatial and temporal cues in the speed discrimination task (Champion and Freeman, 2010; Freeman et al., 2014; Gegenfurtner and Hawken, 1995; Reisbeck and Gegenfurtner, 1999; Wardle and Alais, 2013). Proportion of "faster" responses was computed for each participant as a function of  $W_{distance}$  and  $W_{duration}$  in two separate analyses. For each orientation of the distance-duration plane and each subject, psychometric curves were fitted to these two proportions by cumulative Gaussian function. We extracted a total of 16 Just Noticeable Difference (JND, 8 from the  $W_{distance}$  fit and 8 from the  $W_{duration}$  fit) scores from the standard deviation of the best fitting function. Thus, we applied the set of 16 JNDs of each subject to the distance-duration plane (as a Cartesian coordinate system), and fitted ellipses to this set by using the function *fit\_ellipse* in Matlab (Gal, 2020). Interpretation of the ellipse orientations followed the discrimination contours technique. Specifically, we expected ellipses oriented *i*) along with the oblique orientation ( $45^\circ$ , speed-dominant orientation) when participants preferentially encoded speed cues; *ii*) parallel to the y-axis ( $0^\circ$ , duration-dominant orientation) when they preferentially encoded temporal cues; *iii*) parallel to the x-axis ( $90^\circ$ , distance-dominant orientation) when they preferentially encoded spatial cues. For a graphical explanation of the ellipse's orientations see Figure 3.3. Thus, for instance, a vertical ellipse ( $0^\circ$ , duration-dominant orientation) would have been the result of large JNDs when only spatial information was provided (in the spatial condition) and small JNDs when only temporal information was provided (in the temporal condition), meaning that participants were more sensitive to the temporal cues rather than to the distance or the speed cues in discriminating the target speed. To statistically evaluate if SC and EB groups were more sensitive to spatial, temporal, or speed information, we compared the orientation of

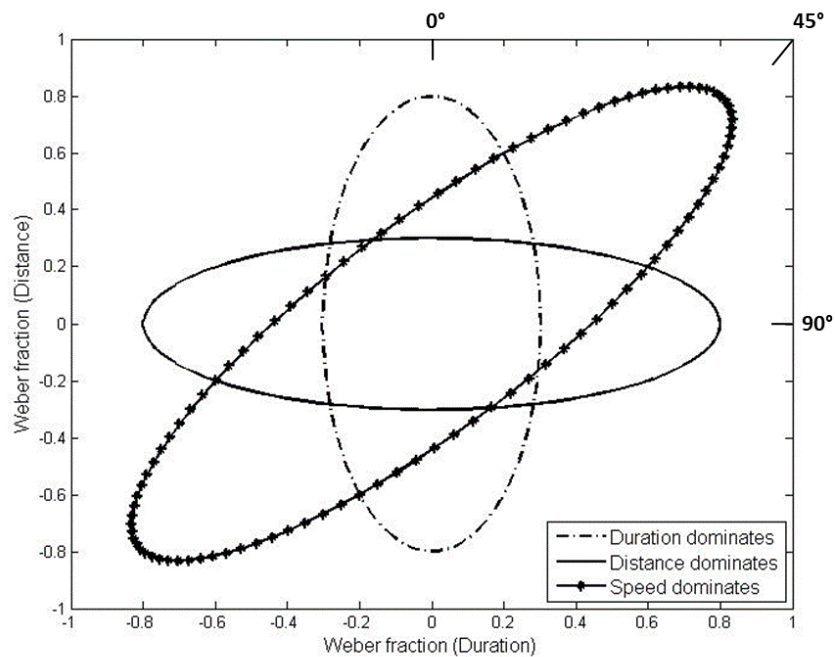


Figure 3.3 - Exemplar ellipse's orientations. Three exemplar ellipses have been applied to the distance-duration plane resembling the three dominant orientations: duration-dominant orientation ( $0^\circ$ ), speed-dominant orientation ( $45^\circ$ ), and distance-dominant orientation ( $90^\circ$ ). These ellipses' orientations would show if participants were more sensitive toward temporal, speed, or spatial cues of auditory motion. *Reproduced with permission. Bertoni et al. (2021). Auditory speed processing in sighted and blind individuals. Plos one, 16(9), e0257676.*

the subjects' ellipses (in degrees) to the three dominant orientations ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ). Since data were not normally distributed for one group (Shapiro–Wilk test) we computed three permutation paired t-tests per group through the *perm.t.test* function for R (Hervé, 2022).

In addition, we wanted to evaluate the ability to discriminate speed through audition for each condition. To do that, we calculated the proportion of "faster" responses as a function of the radial distance  $r = \sqrt{(W_{distance}^2 + W_{duration}^2)}$  (Champion and Freeman, 2010; Freeman et al., 2014). When  $r = 0$ , T and R speeds were equal, while  $r$  values associated with points on the plane with y coordinates  $< 0$ , were turned to negative values ( $r_{turned} = -r$ ).  $r$  values were fitted by cumulative Gaussian functions and JND scores were extracted (after verifying the psychometric function's goodness of fit by checking the significance of its R squared at alpha level = 0.05). When psychometric functions were not significant, we assigned the worst JND of the sample to the function (consequently interpolating four SC participants in the spatial condition). It is worth noting that we observed from this fitting, in both groups, inverted psychometric curves in the spatial and coherent conditions (an example in Figure 3.4) which

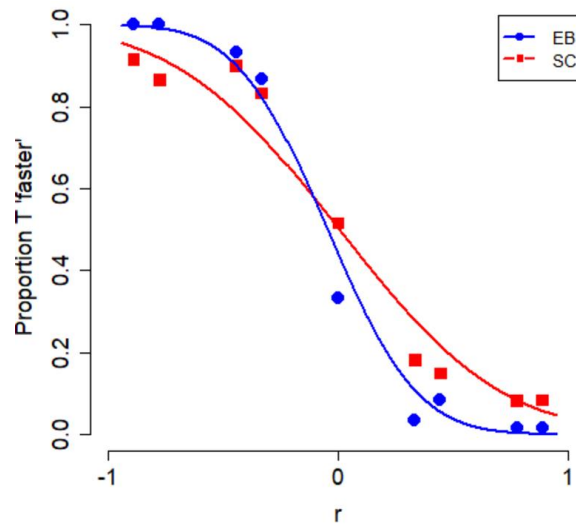


Figure 3.4 - Psychometric curves of a sighted participant (red curve) and an early blind participant (blue curve). Exemplifications of inverted psychometric curves, fitted to the proportion of “faster” responses on the radial distance ( $r$ ). Inverted psychometric curves denoted an impediment in performing the speed discrimination task. In this example, the reported psychometric curves refer to the coherent condition. *Reproduced with permission. Bertoni et al. (2021). Auditory speed processing in sighted and blind individuals. Plos one, 16(9), e0257676.*

corresponded to negative JNDs. We interpreted this result with participants not relying their answers on the stimulus’s speed, but rather only on its spatial or temporal cues (and consequentially failing the speed discrimination task). To include these conditions together with the others, we decided to apply a conversion to the negative JND ( $JND_{neg}$ ), where  $JND_{conv} = JND_{neg} - \min(JND) + \max(JND)$  for each group (Amadeo, Campus, Pavani, et al., 2019; Gori et al., 2018). By applying this conversion, low  $JND_{conv}$  indicated good precision in speed discrimination, and high  $JND_{conv}$  poor precision in speed discrimination (but good precision in either the temporal or the spatial domain, depending on the condition). We compared EB and SC groups in their ability to discriminate the speed of moving sounds by performing permutation ANOVA on  $JND_{conv}$  values (5000 iterations) with Group (EB vs SC) as a between-subjects factor, and Condition (spatial vs temporal vs coherent vs opposite) as a within-subjects factor. We used the *aovperm* function for R (Frossard and Renaud, 2019) to run the analysis. Alpha level was set at 0.05 for all statistical analyses and Bonferroni correction was used for multiple comparisons.

### 3.1.2 Results

We evaluated how EB and SC combined spatial, temporal, and speed cues in perceiving auditory motion. Four participants (3 EB and 1 SC) were excluded from the ellipse orientation analysis since their sets of JNDs were not suitable for ellipses' fitting. Graphical inspection showed that ellipses of the EB and SC groups (Figure 3.5) were oriented vertically, indicating that participants of both groups were more sensitive to changes in stimuli temporal duration than traveled distance or speed. Statistical analysis confirmed these observations, by showing that, for both groups, the orientation of the ellipses was not significantly different from duration-dominant orientation ( $0^\circ$ : SC:  $t = -1.28$ ,  $p\text{-value} = 0.234$ ,  $95\% \text{ CI} [-1.05, 0.26]$ , Hedges's  $g = -0.39$ ; EB:  $t = -2.55$ ,  $p\text{-value} = 0.078$ ,  $95\% \text{ CI} [-1.74, -0.03]$ , Hedges's  $g = -0.84$ ), while was oriented away from the distance-dominant orientation ( $90^\circ$ : SC:  $t = -27.57$ ,  $p\text{-value} = 0.005$ ,  $95\% \text{ CI} [-12.89, -4.52]$ , Hedges's  $g = -8.30$ ; EB:  $t = -20.21$ ,  $p\text{-value} = 0.015$ ,  $95\% \text{ CI} [-11.18, -3.19]$ , Hedges's  $g = -6.64$ ) and the speed-dominant orientation ( $45^\circ$ : SC:  $t = -14.43$ ,  $p\text{-value} = 0.002$ ,  $95\% \text{ CI} [-6.90, -2.31]$ , Hedges's  $g = -4.34$ ; EB:  $t = -11.38$ ,  $p\text{-value} = 0.015$ ,  $95\% \text{ CI} [-6.36, -1.17]$ , Hedges's  $g = -3.74$ ). Finally, EB and SC participants did

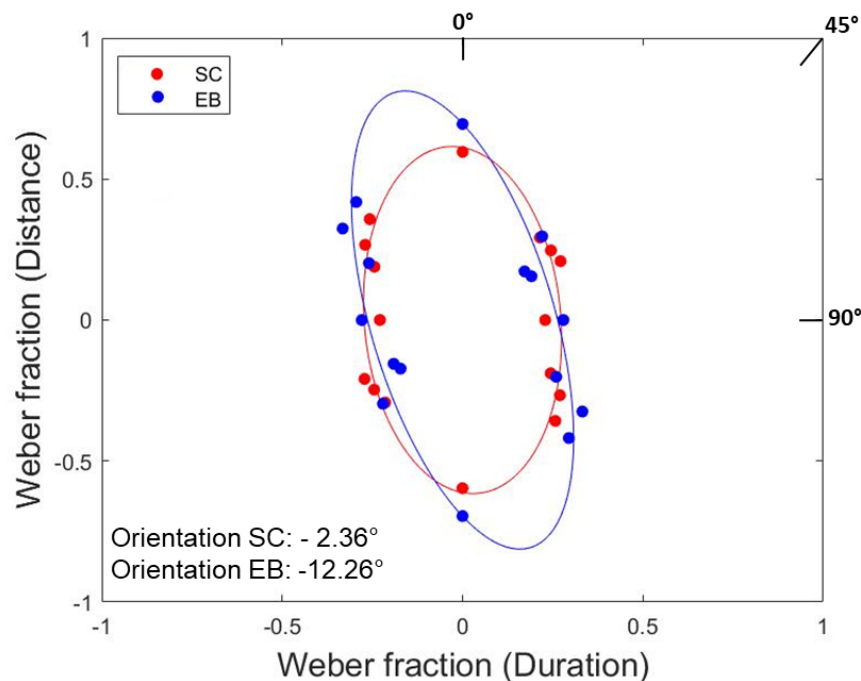


Figure 3.5 - Ellipses of EB (in blue) and SC (in red) groups. For each ellipse, data points represent the JNDs of the psychometric functions plotted along each orientation. Ellipses oriented close to the  $0^\circ$  axis suggest temporal dominance. *Reproduced with permission. Bertoni et al. (2021). Auditory speed processing in sighted and blind individuals. Plos one, 16(9), e0257676.*



not differ in terms of ellipses orientation ( $t = 1.46$ ,  $p\text{-value} = 0.178$ ,  $95\% \text{ CI} [-1.74, 0.35]$ , Hedges's  $g = -0.69$ ). Overall, these results suggest that, in the task, spatial and temporal information were separable dimensions and participants tended to base their judgments of the sound speed on the temporal components of moving sounds (temporal dominance).

Further qualitative analyses revealed that EB and SC, in addition to the temporal dominance, also followed a temporal assumption that identified as faster those sounds with a shorter duration. This aspect can be observed in the graph in (Figure 3.6) that shows subjects' proportion of "faster" responses for each point of the distance-duration plane. The graph illustrates high proportions of "faster" responses on the left side of the plane, which suggest that participants identified T sounds lasting less than the R sound as faster stimuli (temporal assumption). For both groups, the mean "faster" proportion was significantly higher in the shorter duration/longer distance point than in the longer duration/shorter distance point (EB:  $\chi^2_{(1,10)} = 87.09$ ,  $p\text{-value} < 0.001$ ,  $95\% \text{ CI} [0.32, 0.53]$ , Cramer's  $V = 0.43$ ; SC:  $\chi^2_{(1,10)} = 93.63$ ,  $p\text{-value} < 0.001$ ,  $95\% \text{ CI} [0.30, 0.52]$ , Cramer's  $V = 0.41$ ). This result means that the speed was considered faster every time T was shorter in time, independently of space, but not every time it was shorter in space. This temporal assumption led participants to be unable to perform the speed discrimination task in some conditions, specifically, in the coherent condition (in which the temporal assumption prevented participants from perceiving that T

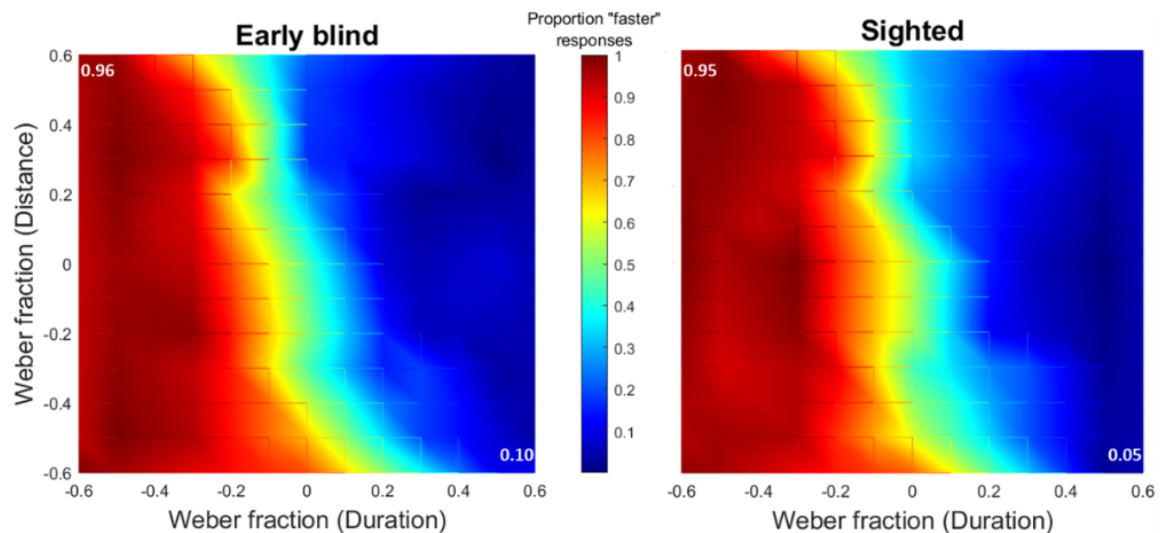


Figure 3.6 - "Faster" proportion graph. On the distance-duration plane, the proportion of "faster" responses was plotted for each point of the plane, for the two groups separately. On the left-up and right-down corners of each graph, "faster" proportion values were reported. *Reproduced with permission. Bertonati et al. (2021). Auditory speed processing in sighted and blind individuals. Plos one, 16(9), e0257676.*

speed = R speed) and in the spatial condition (in which the temporal assumption prevented participants from recognizing that shorter distances corresponded to slower speeds).

To evaluate differences in the speed discrimination performance of the two groups, we compared EB and SC in their speed discrimination precision ( $JND_{conv}$ ). Results revealed a significant main effect of Group (Iter: 5000,  $F = 97.69$ ,  $p$ -value  $< 0.001$ , 95% CI [0.19, 0.15], partial  $\eta^2 = 0.36$ ), with EB participants showing a lower precision in the speed discrimination compared to SC (Figure 3.7). The main effect of Condition (Iter: 5000,  $F = 182.34$ ,  $p$ -value  $< 0.001$ , 95% CI [0.33, 0.62], partial  $\eta^2 = 0.50$ ) revealed, for both SC and EB, higher precision in the speed discrimination when only temporal cues were available (temporal condition) compared to all other conditions (spatial:  $t = -11.48$ ,  $p$ -value  $< 0.001$ , 95% CI [-3.11, -1.43], Hedges's  $g = -2.22$ ; coherent:  $t = -12.66$ ,  $p$ -value  $< 0.001$ , 95% CI [-3.45, -1.62], Hedges's  $g = -2.48$ ; opposite:  $t = -3.23$ ,  $p$ -value = 0.008, 95% CI [-1.67, -0.56], Hedges's  $g = -1.09$ ). The condition with the lowest performance was when target and reference speeds were equal (coherent vs spatial:  $t = 4.33$ ,  $p$ -value  $< 0.001$ , 95% CI [-2.27, -0.94], Hedges's  $g = -1.57$ ;

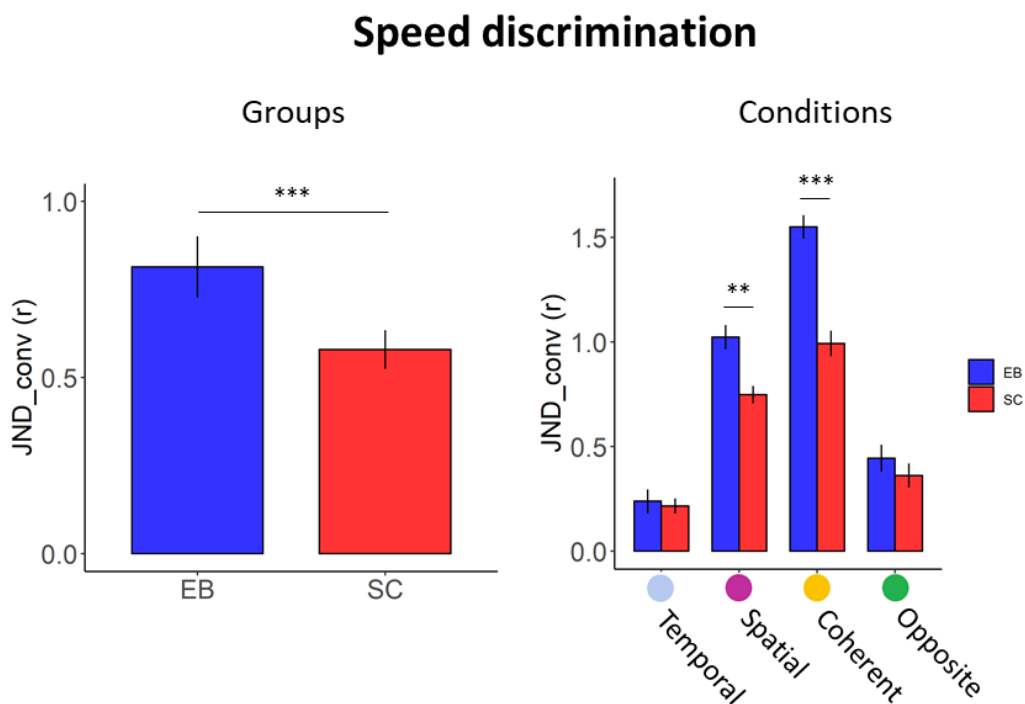


Figure 3.7 - Group performance in the speed discrimination task. The left panel reports the average  $JND_{conv}$  of each group. The right panel reports the average  $JND_{conv}$  per group for each experimental condition. Error bars show SEM. \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . *Reproduced with permission. Bertonati et al. (2021). Auditory speed processing in sighted and blind individuals. Plos one, 16(9), e0257676.*

coherent vs temporal:  $t = 12.66$ ,  $p\text{-value} < 0.001$ , 95% CI [-3.45, -1.62], Hedges's  $g = -2.48$ ; coherent vs opposite:  $t = 9.97$ ,  $p\text{-value} < 0.001$ , 95% CI [1.26, 2.82], Hedges's  $g = 2.00$ ). Lastly, participants were significantly less precise in the spatial condition than in the opposite condition ( $t = 7.57$ ,  $p\text{-value} < 0.001$ , 95% CI [1.00, 2.38], Hedges's  $g = 1.66$ ), suggesting that participants' performance dropped when temporal information was not available (as in the spatial condition in which target and reference sound had the same temporal duration), compared to when spatial but also temporal cues were available (as in the opposite condition). It is worth noting that also from these results we could reveal that the temporal properties were successfully informative only when the temporal assumption that short duration corresponds to fast speed was satisfied (i.e., in the temporal and opposite condition), and not when this temporal assumption did not subsist (i.e., in the coherent and spatial condition). Finally, results revealed a significant interaction between Group and Condition (Figure 3.7; Iter: 5000,  $F = 11.79$ ,  $p\text{-value} < 0.001$ , 95% CI [0.12, 0.45], partial  $\eta^2 = 0.31$ ) that suggested that participants' performance was influenced not only by their visual experience but also by the kind of experimental condition (i.e., the kind of cues available for speed discrimination). Post hoc comparisons (after Bonferroni correction for multiple comparisons) showed that EB individuals were significantly less precise than SC in the spatial and coherent conditions (spatial:  $t = 3.88$ ,  $p\text{-value} = 0.006$ , 95% CI [0.63, 2.50], Hedges's  $g = 1.50$ ; coherent:  $t = 6.73$ ,  $p\text{-value} = 0.001$ , 95% CI [0.72, 2.70], Hedges's  $g = 1.64$ ), while they were similar in the temporal and opposite conditions (temporal:  $t = 0.34$ ,  $p\text{-value} = 1$ , 95% CI [-0.52, 0.68], Hedges's  $g = 0.08$ ; opposite:  $t = 0.95$ ,  $p\text{-value} = 1$ , 95% CI [-0.36, 0.85], Hedges's  $g = 0.24$ ). It is worth noting that the conditions that revealed a greater difference between the EB and SC participants also inverted the psychometric curves, meaning that participants were following a misleading temporal assumption that dropped their performance. Since EB participants were less precise than SC in the conditions in which they applied this temporal assumption, results suggested that the lack of vision may play a role in the erroneous use of temporal information for motion processing.

### 3.1.3 Discussion of Study 3

The present study examined how visual experience shapes spatial and temporal representations underlying acoustic speed perception. A group of early blind individuals and one of sighted people determined how fast sounds were moving at different speeds by relying on sound's duration, traveled distance, or both components. Results showed that early blind participants were influenced in their acoustic speed discrimination by distinctive processing

of spatio-temporal motion cues. Both groups relied on a temporal dominance and a temporal assumption in their speed discrimination, but early blind participants emerged to be more influenced by these mechanisms than sighted controls.

In Study 3, we revealed a temporal attraction of both sighted and blind participants in the acoustic speed perception which was expressed in two main forms. On the one side, in both groups, the temporal components of moving sounds dominated the speed discrimination performance, with participants preferentially relying on temporal cues to determine the speed of the target sounds (temporal dominance). This result confirmed past findings on sighted adults, challenging the existence of acoustic motion-sensitive detectors similar to those in the visual system (Carlile and Best, 2002; Carlile and Leung, 2016; Freeman et al., 2014; Locke et al., 2016), but was the first, to our knowledge, that was revealed in the visually impaired population. Blind individuals have already been shown to be competent at temporal tasks (Hötting and Röder, 2009; Stevens and Weaver, 2005; Van Der Lubbe et al., 2010). Here, we also showed that temporal dominance in the acoustic motion perception occurs not only in sighted individuals but also in the lack of visual experience, suggesting that blindness does not allow the recruitment of visual motion-sensitive detectors by audition, as one could hypothesize in a cross-modal cortical reorganization point of view. On the other side, all participants tended to consider the target speeds as faster whenever the target sound was shorter than the reference, independently of space (temporal assumption). It is worth noting that, by following this temporal assumption, participants dropped in their discrimination of the sound speed in those conditions wherein shorter sound duration did not imply faster speed (i.e., in the coherent and spatial conditions). In Chapter 5, these results are further discussed in light of human perception being influenced by previous knowledge reflecting the probability distribution of the environment's speed property (Parise et al., 2014; Senna et al., 2015, 2017).

Although both early blind and sighted participants reported temporal dominance and misleading temporal assumption, early blind individuals reveal even worse performance than sighted controls in discriminating the moving sounds' speed, in spatial and coherent conditions. In the spatial condition, since participants could rely exclusively on distance cues to discriminate the sounds' speed, blind participants may have been more impaired due to the spatial deficit of this population that past studies have described (Gori et al., 2014; Vercillo et al., 2016). Moreover, since blind individuals tend to rely on the event's temporal aspects to infer spatial information (by assuming a prior of stimulus' constant velocity, Amadeo et al., 2020b; Gori et al., 2018, 2020a), in our study they may have applied, more than the sighted controls, the temporal assumption of short duration equal to fast speed to the spatial

domain. This mechanism would have made them erroneously interpret smaller traveled distances as faster speeds. In the coherent condition, in which the spatial and the temporal motion components corresponded (see Figure 3.2 for more details), early blind individuals may have selectively focused their attention on the domain they encoded better (i.e., time) and neglected the presence of spatial information, leading to a lower precision in the speed discrimination.

The findings of this study should be considered in light of a limitation regarding the small sample size of the experimental groups, which, however, resembled the sample size of other studies investigating the perceptual abilities of early blind people (Gori et al., 2014; Lessard et al., 1998; Röder et al., 1999; Vercillo et al., 2016; Voss et al., 2015).

In conclusion, our results showed that the lack of visual experience in the first years of life may influence the ability to encode the acoustic speed by increasing the preference of blind individuals towards the stimuli' temporal characteristics. Specifically, early visual perception seems to be fundamental to prevent the misleading use of compensatory behaviors in the perception and interpretation of environmental events, such as motion perception.

## **3.2 Study 4: the use of visual spatio-temporal information for the manual interception of moving objects**

In many everyday situations, we make goal-directed arm movements towards moving objects, such as when intercepting a rolling ball. Intercepting moving targets is a challenging task that requires control of where the interception will take place as well as when (i.e., spatial and temporal precision, respectively; Brenner and Smeets, 2018). Various components of a person's movements are adjusted to meet spatio-temporal precision demands of interception. For instance, movement times proportionally vary with temporal and spatial precision, so that they are shorter when the temporal precision demand of the task is high (Tresilian and Houseman, 2005; Tresilian and Plooy, 2006; Tresilian et al., 2009), and longer when the spatial precision demand (in movement endpoint or direction) is greater (Tresilian et al., 2009). However, since most interception tasks require a degree of both spatial and temporal precision, a compromise between the various movement components to adjust should be reached. In this regard, in a task in which both spatial and temporal precision was required, participants flexibly modified their movement times in relation to the spatio-temporal accuracy demands of the task (Tresilian et al., 2009), suggesting that, when interacting with a moving object, our motor behavior can become adaptable to the context.

Many daily situations that include the interaction with moving objects also require people to correct future movements in relation to previous interceptive errors. Past studies showed that the human sensorimotor system is able to use internal representations of interceptive errors in order to make corrections in the next trials and that these representations are based on the gap between the planned and the actual action outcomes (López-Moliner et al., 2019). People correct interceptive errors in response to situations that may or may not include external perturbations and/or feedback of their performance (Burge et al., 2008; de la Malla et al., 2014; Herzfeld et al., 2014; van Beers, 2009). In this regard, people take into great consideration the presence of feedback of their interception even when these are delayed (i.e., the perceptual consequences of their action are delayed from the action itself); however, sensorimotor performance is greater when feedback is accurate rather than delayed (de la Malla et al., 2014). Finally, interceptive movements are adjusted on the basis of sensory information until is no longer possible due to sensorimotor delays (Brenner and Smeets, 2011, 2015; Veerman et al., 2008) and the magnitude of such adjustment depends on the urgency of the response (Liu and Todorov, 2007; Oostwoud Wijdenes et al., 2011).

When reaching for static objects, the brain detects and distinguishes between the spatio-temporal features of the item even just by looking at it shortly before reaching (Land, 2006, 2009). Conversely, with moving objects, sensorimotor adjustment to spatio-temporal components becomes more difficult, since one has to determine, on the one side, where the target will be at the moment of impact (interception position) and, on the other side, when the target will reach the planned interception point (interception timing) (Brenner and Smeets, 2015). Judging the position and timing of reaching is more challenging with moving than with static objects because, in the first case, the two components are tightly coupled (and consequentially difficult to disentangle) by the motion performed by the object itself. For example, if we assume a target moving at a certain velocity, in the correction of interceptive errors to this target, the brain has to compute a combined compensation of the spatial (in position) and temporal (in timing) errors which has to change depending on the target velocity. In addition to this issue is the fact that adjusting the interceptive movements in a certain way can be advantageous for some aspects but disadvantageous for others (Brenner and Smeets, 2015, 2018). Thus, the brain has the difficult task to consider all constraints imposed by the object's motion and to select the sensory information that in each specific circumstance is more relevant to the task. Target velocity is one of the aspects considered by the sensorimotor system to optimize interception performance (Brouwer et al., 2000). Specifically, Smeets and Brenner (1995) showed that people moved their hand faster when they have to intercept targets perceived as moving quickly (because of background moving

in the opposite direction) than targets perceived as slower. Moving quickly to intercept fast targets decreases spatial precision but increases temporal precision (Schmidt et al., 1979), and this mechanism was explained as correctly judging the timing of interception of a fast target being more important than correctly judging its interception position (Brenner and Smeets, 2015).

In summary, when intercepting moving objects, both spatial and temporal precision are required (Brenner and Smeets, 2018) and the way the arm moves is adjusted based on spatio-temporal errors in previous attempts (López-Moliner et al., 2019). Given these considerations, in Study 4 we interrogated whether and how participants corrected the spatio-temporal components of future interception (i.e., position and timing of interception) when each of these is more relevant for the task. Specifically, we investigated whether what one adjusted in each next interception depended on the target's speed. To correct for a given spatial error, the faster the target moves, the shorter needs to be the adjustment in time to solve the spatial error. Therefore, we hypothesized that for participants would be more convenient to adjust the timing of interception for fast targets and the position of interception for slow targets. To see whether the interceptive movements adjustment of participants was flexible (i.e., considering the previous errors together with target speed), or fixed (i.e., considering previous errors only), we proposed an interception task to 64 healthy people in which performance was measured through a motion tracking system and conveyed to participants through visual error feedback.

### 3.2.1 Methods

#### Sample

64 healthy people took part in the study. Of these, 24 individuals (17 females, mean age  $\pm$  SD:  $28 \pm 8.48$  years old) participated in Experiment 1, 24 individuals (16 females, mean age  $\pm$  SD:  $22 \pm 3.03$  years old) in Experiment 2 and 16 individuals (10 females, mean age  $\pm$  SD:  $29 \pm 10.62$  years old) in Experiment 3. All participants reported no history of neurological, cognitive, and/or sensorimotor deficits and they gave written informed consent prior to testing. The study was approved by the ethics committee of the local health service (Ethics Committee of Vrije Universiteit Amsterdam) in accordance with the Declaration of Helsinki.

### **Setup and calibration**

The study consisted of three experiments. In all experiments, participants stood in front of a large screen (Techplex 150; 1.25 m x 1.00 m, tilted backward by 30°) in a normally illuminated room. Images were back-projected at 120 Hz with a resolution of 800 × 600 pixels (Figure 3.8a). An infrared camera system (Optotrak 3020) was placed at about shoulder height above the ground to the left of the screen and, recording at 500 Hz, measured the position of a marker (an infrared LED) attached to the nail of the participant's right index finger. At the beginning of each block, calibration consisted in the participants placing the fingertip with the marker at five positions indicated on the screen (four at the corners and one in the center of an imaginary 60 cm x 50 cm rectangle on the screen), and these positions measured by the Optotrak. In this way, we could determine the position of the fingertip with respect to the projected images on the screen every 2 ms throughout the experiment.

### **Stimuli and procedure**

In this section, the aspects of stimuli and procedure that were similar across the three experiments of Study 4 are described. Further details on each specific experiment are provided in the dedicated sections. In Study 4, we investigated how visual feedback was used to guide the correction of hand movements toward a moving object, based on the object's speed. To test this, participants performed a manual interception task. Each trial (Figure 3.8b) started with participants placing their right index finger at the starting point, a 1.5 cm green disk placed 20 cm to the right of and 20 cm below the screen center. Between 0.6 and 1.2 s after participants placed their finger at the starting point, a grey target disk appeared on the screen 20 cm above the screen center and started moving rightwards at a constant velocity. In different blocks counterbalanced across subjects, the target moved either slowly (0.25 m/s; start of motion on the screen at 5 cm to the right of the screen center) or fast (1.1 m/s; start of motion on the screen at 46 cm to the left of the screen center). The size of the target disk was fixed (2 cm of radius) in Experiment 2 and Experiment 3 and varied trial by trial by factor 1.1 in Experiment 1 (more details on the target size are reported in the experiments' dedicated sections). Between each trial, participants could rest by not placing their finger at the starting point.

Participants' task was to intercept the moving target disk right after its appearance by lifting their finger from the starting point and tapping with it on the moving target. While the target was moving, a systematic spatial displacement made the target jump a fixed amount (details on the size of the systematic jump are described for each experiment in the dedicated



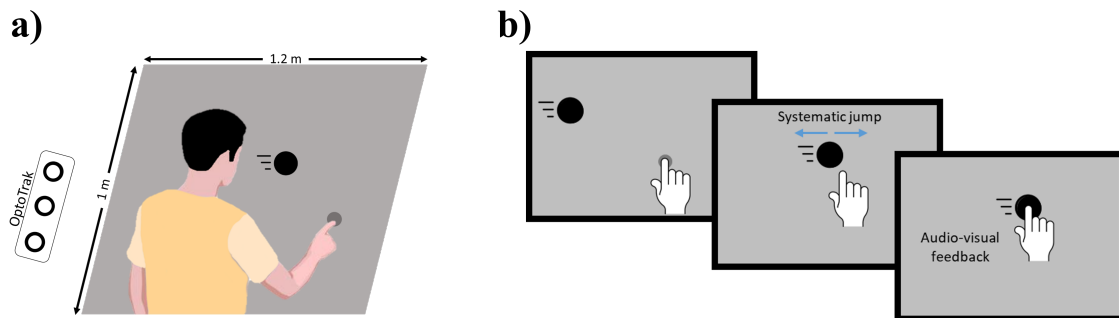


Figure 3.8 - Experimental procedure. a) Experimental setup: participants stood in front of a large screen on which they intercepted rightward moving targets with their right index finger. Optotrak system recorded participants' movements throughout the entire experiment. b) Graphical explanation of a trial: after participants placed their right index on the starting point, a target disk started moving to the right of the screen at one of two speeds (0.25 m/s or 1.1 m/s) in two different blocks. While moving, the target motion was perturbed by a systematic jump (on the left or on the right) that induced participants to make interceptive errors. After each interception (i.e., the finger tapping on the screen) audiovisual feedback on performance was provided to participants.

sections) rightward or leftward (Figure 3.8). This jump occurred when it was too late for the participants to adjust their movement to hit the target (i.e., when the participant's hand was 1.5 cm from the target) so that participants interpreted it as their own interception error and tried to correct themselves in each next trial (an example for a rightward target jump in Figure 3.9). Rightward and leftward jumps were repeated 100 times for each target jump direction (left or right) and presented in a randomized order within each block, for a total of 400 trials per participant (200 trials per block). A tap on the screen was detected when the finger's deceleration (in the direction of the screen) was larger than  $250 \text{ m/s}^2$ , and the finger was less than 0.5 cm above the screen.

**Error feedback:** after 500 ms from tap detection, audiovisual feedback was provided with respect to participants' performance. We determined whether the target was hit by comparing the position of the finger at the moment of the tap and the position of the target at the exact same moment (but after the target position had been shifted due to the systematic jump). We considered a target as hit if the position of the right index finger was within the outline of the target at the moment of the tap. In this case, the target stopped and disappeared from the screen after 500 ms, and simultaneously a bell rang (positive feedback). On the contrary, we considered a target as missed if the finger position was outside the target contours at the moment of the tap. In this case, the target deflected away from the finger at 1 m/s (e.g. if

the finger tapped below and to the right of the target, the disk moved up and to the left) and no acoustic feedback was delivered (negative feedback). We chose to provide explicit feedback on the performance for two main reasons: *i*) providing perceptual feedback encourages people to correct their sensorimotor behaviors in order to improve the interception of moving objects, more than no-feedback conditions (Gray et al., 2006); *ii*) the kind of feedback we delivered gave explicit information on the properties of the interceptive error committed by the participant. Specifically, when participants missed a target, they could easily determine whether their tap was too late or too early (error in the timing of interception), or too on the right or on the left than the target (error in the position of interception) and correct themselves in each next trial based on the feedback. It is also important to keep in mind that feedback provided to participants was related to the interceptive errors induced by the systematic spatial displacement included in each trial as experimental manipulation (Figure 3.8). Thus, this systematic jump (and the related feedback) allowed us to identify the effects of responding specifically to error feedback. The adjustment of participants' interception in each next trial would ideally correspond, in direction and magnitude, to the systematic jump made by the target in each previous trial: as shown as an example in Figure 3.9, with a target jumping on the right in the previous trial (trial  $n-1$ ), one can interpret the error *i*) as a position error (in trial  $n-1$  the participant's tap was too on the left than the target) and correct it by shifting on the right the position of interception on the next trial (trial  $n$ ), or *ii*) as a timing error (in trial  $n-1$  the participant's tap was too late) and correct it by tapping earlier on the next trial (trial  $n$ ). Thus, target jump direction (leftward or rightward) allowed us to disentangle between adjustment in position (i.e., the position on the screen where they tried to tap the target, in mm) and timing (i.e., how fast they were to hit the target, in ms).

### Experiment 1

Here, the target size (disk radius) changed trial by trial according to the participant's performance. Specifically, in each block, the target size started from 1 cm and increased by factor 1.1 (i.e., 10% of the previous target size) when participants could not hit the target in the previous trial or decreased by factor 1.1 when they could hit the target in the previous trial. We introduced changes in the target size to have participants hit about 50% of the targets, irrespective of the target speed (1.1 m/s or 0.25 m/s). In this way, the average target size for fast targets was bigger than for slow targets, preventing participants to struggle more in intercepting targets moving at a fast speed and making the performance between the two speeds comparable in terms of difficulty. The systematic target jump (to the left or to the right) was of 1 cm.

### **Experiment 2**

Experiment 2 was designed to investigate whether participants corrected their hand movement in each next trial based on the magnitude of the error (i.e., the size of the target jump). To test this, within each block target jump size randomly varied between 0.5 cm and 1 cm. Each target jump size (0.5 cm or 1 cm) was repeated 50 times for each target jump direction (left or right) and target speed (fast or slow), for a total of 400 trials per participant. In addition, to check that hand movement adjustment was related only to the magnitude of the error and not to changing of target size from trial to trial (as in Experiment 1), we kept the target size fixed at 2 cm radius.

### **Experiment 3**

Experiment 3 was designed to check whether participants corrected the interception timing more if the correction of the position was limited. To limit position correction, we instructed participants to intercept the moving targets within an indicated interception region (a grey square of 10 cm x 10 cm), whose center was placed at 20 cm to the right of and 20 cm above the screen center. A target was considered hit only if its position at the moment of the tap (together with the position of the participant's right index finger at the exact same moment) was within the interception region (providing the corresponding positive feedback). The interception region was the same in slow and fast blocks. Target size (2 cm radius) and target jump (1 cm) were kept fixed.

### **Data analysis**

For each experiment, to investigate how, when intercepting moving targets, participants corrected their hand movement in response to error feedback, and whether this correction was based on the target speed, we examined the participants' interception adjustment for the two speeds separately (i.e., 1.1 m/s and 0.25 m/s). People could respond to errors by either changing where (position; Figure 3.9 middle panel) or when (timing; Figure 3.9 lower panel) they tapped. Our question was whether they would change where they tapped less (because they would try to correct more by when they tapped) when the target moved faster. Since the time and place of the tap may gradually shift during the experiment, irrespective of errors, we concentrated on changes between consecutive trials. We averaged across trials to remove changes in response to any other errors than those imposed by the target jumps. We used the median rather than the mean to not have to worry about outliers. Specifically, for each trial, we measured the amount of adjustment for the timing of interception, that was

how fast were participants to hit the target (in ms), and for the position of interception, that was the position on the screen where they tried to tap the target (in mm), in relation to the systematic jump introduced in the previous trial (in Figure 3.9 a graphical explanation of expected adjustment in relation to a target jump to the right). To do this, we calculated the median of the difference in position and timing with respect to the previous trial:

$$\textit{position difference} = \textit{median}(\textit{position trial}_n - \textit{position trial}_{n-1})$$

$$\textit{timing difference} = \textit{median}(\textit{timing trial}_n - \textit{timing trial}_{n-1})$$

for each target jump (left or right). Then, we computed the position and timing adjustment by calculating the subtraction in the median position difference and in the median timing difference between the left and right jump groups. Finally, to compare the position and timing adjustment between the two target speeds (after checking for normality of data distribution with Shapiro-Wilk test), we ran two separate one-tailed t-tests on the position and timing adjustment with Speed (Fast, Slow) as within-subject factor. We chose to use one-tailed t-tests since in our research question we hypothesized specific directions of results, with higher position adjustment for slow speed than fast speed, and higher timing adjustment for fast speed than slow speed.

### 3.2.2 Results

Study 4 investigated the strategy adopted by our brain to use visual error feedback in guiding the correction of movements towards a moving object, based on target speed. To answer this question, we proposed to 64 participants an interception task with virtual moving objects in three separate experiments.

#### Experiment 1

Experiment 1 explored the influence of target speed on the adjustment of interception for fast and slow target motion. Success in interception (i.e., % of targets hit) was comparable for both speeds because target size (disk radius) changed trial by trial in relation to participants' performance. Consequently, fast targets were on average bigger (mean: 1.6 cm) than slow targets (mean: 1.1 cm). We compared the position and timing adjustment between the two target speeds. As expected, we found that participants adjusted the position of their interceptive behavior more for slow targets than for fast targets (Figure 3.10a;  $t_{(23)} = -1.83$ ,  $p = 0.03$ , Cohen's  $d = -0.37$ , 95% CI [-0.80, 0.05]) and adjusted the timing more

for fast targets than for slow targets although the latter difference was not statistically significant (Figure 3.10a;  $t_{(23)} = 0.87$ ,  $p = 0.19$ , Cohen's  $d = 0.17$ , 95% CI [-0.23, 0.59]). Considering the bars with standard errors in Figure 3.10a (which suggested high variability among participants' performance), in Experiment 2 and Experiment 3 we manipulated some aspects of the interception task to replicate these findings and examine whether participants would respond more consistently under slightly different circumstances (fixed target size in Experiment 2 and less variability in where they were allowed to hit the target in Experiment 3).

### Experiment 2

In Experiment 2 we kept the target size constant and varied the magnitude of the target jump (either 0.5 cm or 1 cm). For both target speeds, participants adjusted the position of interception more for the 1 cm target jump than for the 0.5 cm target jump (Figure 3.10b). They did not adjust the timing more for the larger target jump than for the small one. For neither target jump size, there was a significant difference between the two target speeds in the timing (0.5 cm jump size:  $t_{(23)} = 1.07$ ,  $p = 0.14$ , Cohen's  $d = 0.21$ , 95% CI [-0.19, 0.63]; 1 cm jump size:  $t_{(23)} = 0.83$ ,  $p = 0.20$ , Cohen's  $d = 0.17$ , 95% CI [-0.24, 0.58]) or position (0.5 cm jump size:  $t_{(23)} = -0.06$ ,  $p = 0.47$ , Cohen's  $d = -0.01$ , 95% CI [-0.42, 0.39]; 1 cm jump size:  $t_{(23)} = 0.13$ ,  $p = 0.55$ , Cohen's  $d = 0.02$ , 95% CI [-0.38, 0.43]) adjustment. However, the trend of adjusting the timing more for fast targets than for slow targets was observed for both jump sizes. Overall, results of Experiment 2 showed that, by changing the target jump size, participants were able to correct their position of interception proportionally to the magnitude of the jump and consequentially of the error experienced.

### Experiment 3

Experiment 3 checked whether participants adjusted the interception timing more if the correction of the position was limited by imposing an interception region and whether this was influenced by target speed. For both target speeds, participants no longer adjusted the position of their interception and more clearly corrected the timing of interception (Figure 3.10c). However, there was no significant difference between the two target speeds in the timing ( $t_{(15)} = -0.07$ ,  $p = 0.47$ , Cohen's  $d = -0.01$ , 95% CI [-0.52, 0.49]) or position ( $t_{(15)} = 0.56$ ,  $p = 0.71$ , Cohen's  $d = 0.14$ , 95% CI [-0.37, 0.65]) adjustment. In addition, the trend of correcting the timing more with fast speed than slow speed revealed in the other two experiments disappeared in Experiment 3. However, it is worth noting that the same absolute

adjustment to the timing corresponds with a much larger fraction of the error being adjusted for the fast target. These results confirmed that people were able to adjust their interceptive behavior to the presence of an interception region.

### 3.2.3 Discussion of Study 4

In Study 4, we investigated whether, when intercepting virtual moving targets, one corrects hand movements in response to error feedback by adjusting the interceptive component (i.e., the position or the timing of interception) that is more relevant according to some target's characteristics. In particular, we wondered whether the interception adjustment depended on the target speed. To test this, groups of healthy people were asked to join three experiments in which they had to intercept a moving disk on a screen by tapping on it with their right index finger. Feedback on performance was provided after each trial, to induce participants to correct themselves accordingly in each next trial. Overall, Study 4 showed that *i*) one can correct the interceptive movements in relation to the visual feedback presented in the previous trial, also based on the target speed, with a tendency to rely on and adjust more the timing of interception for fast targets, and the position for slow targets (Experiment 1); *ii*) this adjustment was performed by taking into account the magnitude of the error to correct (Experiment 2) and the region for intercepting (Experiment 3).

Results of Experiment 1 are in line with past studies revealing the ability of the sensorimotor system of adjusting the interceptive movements on the basis of the latest sensory information (Brenner and Smeets, 2011, 2015; Land and McLeod, 2000; López-Moliner et al., 2019). The present study added evidence that movement corrections were related to sensory information in a clever manner, that is based on the spatio-temporal component of motion interception that was more convenient to correct in each specific circumstance (i.e., spatial component for slow targets and temporal component for fast targets). This observation was particularly evident for the position adjustment, instead of only a trend for the timing adjustment. However, it is worth noting the high variability among participants. Many factors may have contributed to this, including target size changing trial by trial in Experiment 1. Indeed, participants may have been affected in the correction not only by the error feedback of trial  $n-1$  (that was associated with a certain target size), but also by the target size of trial  $n$  (which changed by increasing or decreasing by a factor 1.1 compared to the target size of trial  $n-1$ ). In this regard, Experiment 2 was designed to rule out the possibility that the high variability among participants was due to participants responding not only to target systematic jump but also to target size. Although the inter-subject variability did not

decrease (suggesting that the variability found in Experiment 1 was not due to changes in target size), Experiment 2 demonstrated that the participants' behavior was commensurate with the magnitude of the error. We also revealed that, differently from Experiment 1, here participants equally adjusted the position of interception between the two speeds. This result may be explained by the fact that, by removing changes in target size, fast objects were overall more difficult to intercept than slow objects, thus participants missed more targets in the fast block (and experienced more negative feedback). Since feedback of missed targets was revealed to improve movement adjustment and reduce systematic errors (Brenner et al., 2013), in Experiment 2 the great adjustment of position also with fast objects may be due to the higher number of missed targets in this condition. Finally, Experiment 3 confirmed the rationale that forcing participants to intercept a target within a specific region would increase their tendency to adjust the timing of interception, by reducing their possibility to correct the interception position (Brenner and Smeets, 2015). Although in Experiment 3 no significant difference was revealed in the timing adjustment between the fast and slow target speeds, it is worth noting that an equal adjustment in timing between the two target speeds actually shifts the contribution towards more timing adjustment for fast targets. The fact that the difference in timing adjustment between the two speeds was bigger in Experiment 1 and Experiment 2 may be related to the fact that in these two experiments participants were free to adjust any way they liked, while in Experiment 3 they were constrained by the interception region.

Across all experiments, other two points of discussion emerged. The first one is related to the magnitude of error correction. For both position and timing adjustment, participants tended to correct only a percentage of the error in previous trials. Although this observation may suggest that subjects were actually not able to correct their errors (despite in Experiment 2 we showed that they corrected consistently with the magnitude of the error), this result could also be explained by previous studies showing that our brain corrects about 38% of interception error in the planning of the next movements (Baddeley et al., 2003; Cheng and Sabes, 2007; Scheidt et al., 2001; Scheidt and Stoeckmann, 2007; van Beers, 2009). In addition, differently from these studies, participants may have tended not to fully correct the error because the direction of the jump changed trial by trial and, if two consecutive trials had targets jumping in two different directions, full error correction in one direction could be counterproductive for them. In this regard, it was proposed that when the environment constantly changes (i.e., the error feedback is not consistent across trials) the brain suppresses learning from errors because any learning could be disadvantageous to performance in the next (not identical) trials (Herzfeld et al., 2014). The second matter of debate is the large variability among participants already mentioned above. This variability may suggest that

the average performance was not the result of automatic sensorimotor mechanisms, but of chosen action strategies that varied among participants. In this sense, other aspects such as the use of rewards for successful interceptions (which were not considered in the present study) and the presence of external perturbations to induce errors may have led to a variety of different strategies implemented by the participants. For example, it was suggested that people, in the presence of perturbation, not only have to adjust their movement planning but also infer perturbations and that this estimation may be solved by a wider variety of sensorimotor strategies (van Beers, 2009).

Given these considerations, a future direction could be the design of a study that does not involve external perturbation to induce errors but only takes advantage of the natural movement incorrectness of people. Similarly, an alternative paradigm in which no feedback on performance is used could be implemented to show what would happen in terms of correction when people are not explicitly made aware of their errors. Finally, it would be interesting to test whether the performance would be more consistent by using real objects instead of virtual ones since the former has been proven to elicit more precise interceptive behaviors (Brenner et al., 2012; McLeod et al., 1985).

*Chapter 3 is partially extracted and adapted from Bertonati, G., Amadeo, M. B., Campus, C., and Gori, M. (2021). Auditory speed processing in sighted and blind individuals. Plos one, 16(9), e0257676.*



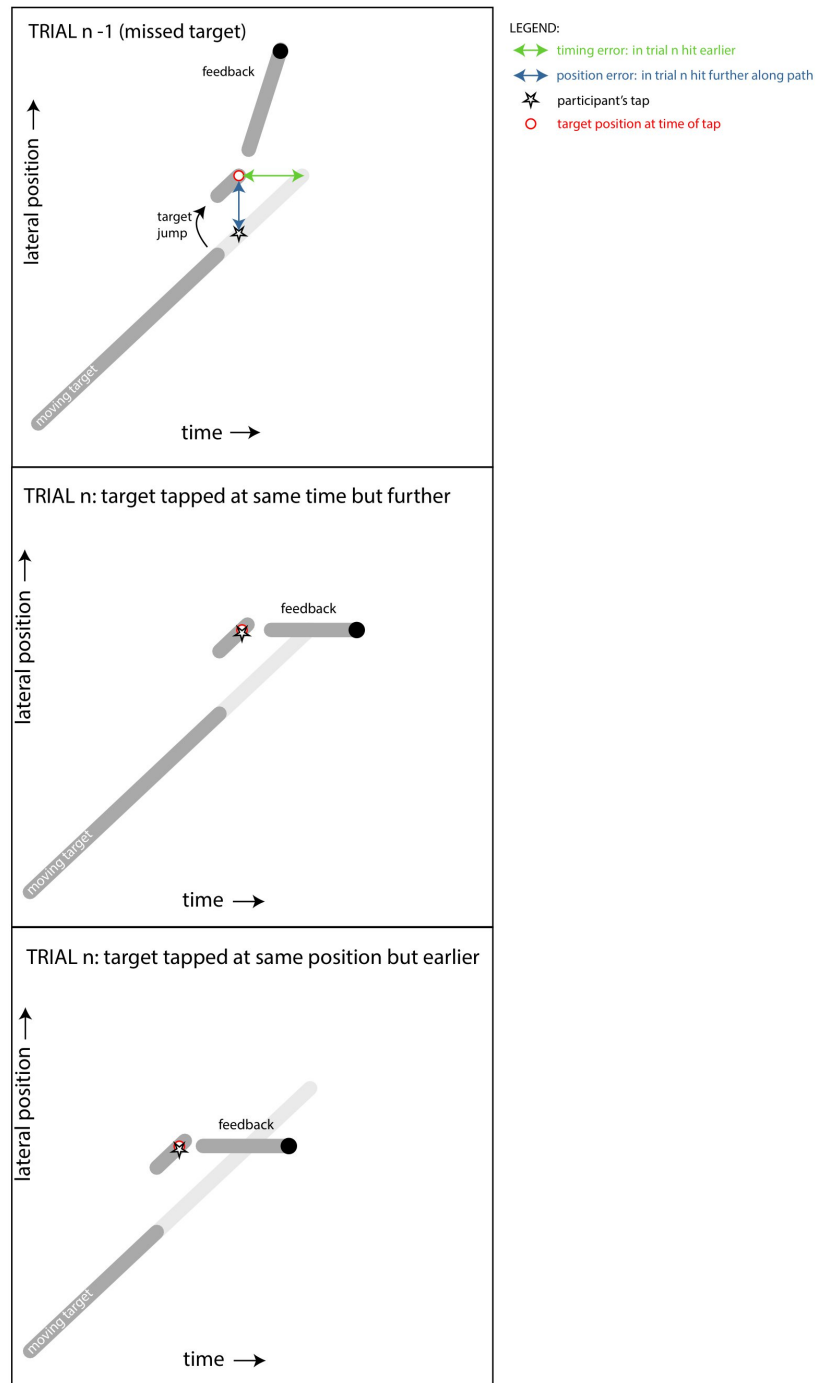


Figure 3.9 - Graphical explanation of possible responses to the target jump, in each next trial (shown here, as an example, for a target jump to the right). The upper panel shows the participant's response in trial n-1, middle and lower panels show the participant's possible responses in trial n. The star indicates the position on the screen of the participant's tap, while the red circle indicates the target's position at the time of the participant's tap. Since in trial n-1 (upper panel) the target jumps too late for the participant to respond on that trial, aiming for the target results in a systematic error that depends on the direction of the jump. Thus, one can correct this systematic error on the next trial (trial n) by shifting where (middle panel) or when (lower panel) one tap, or some combination of both. Correction of the systematic error is shown here for complete compensation.

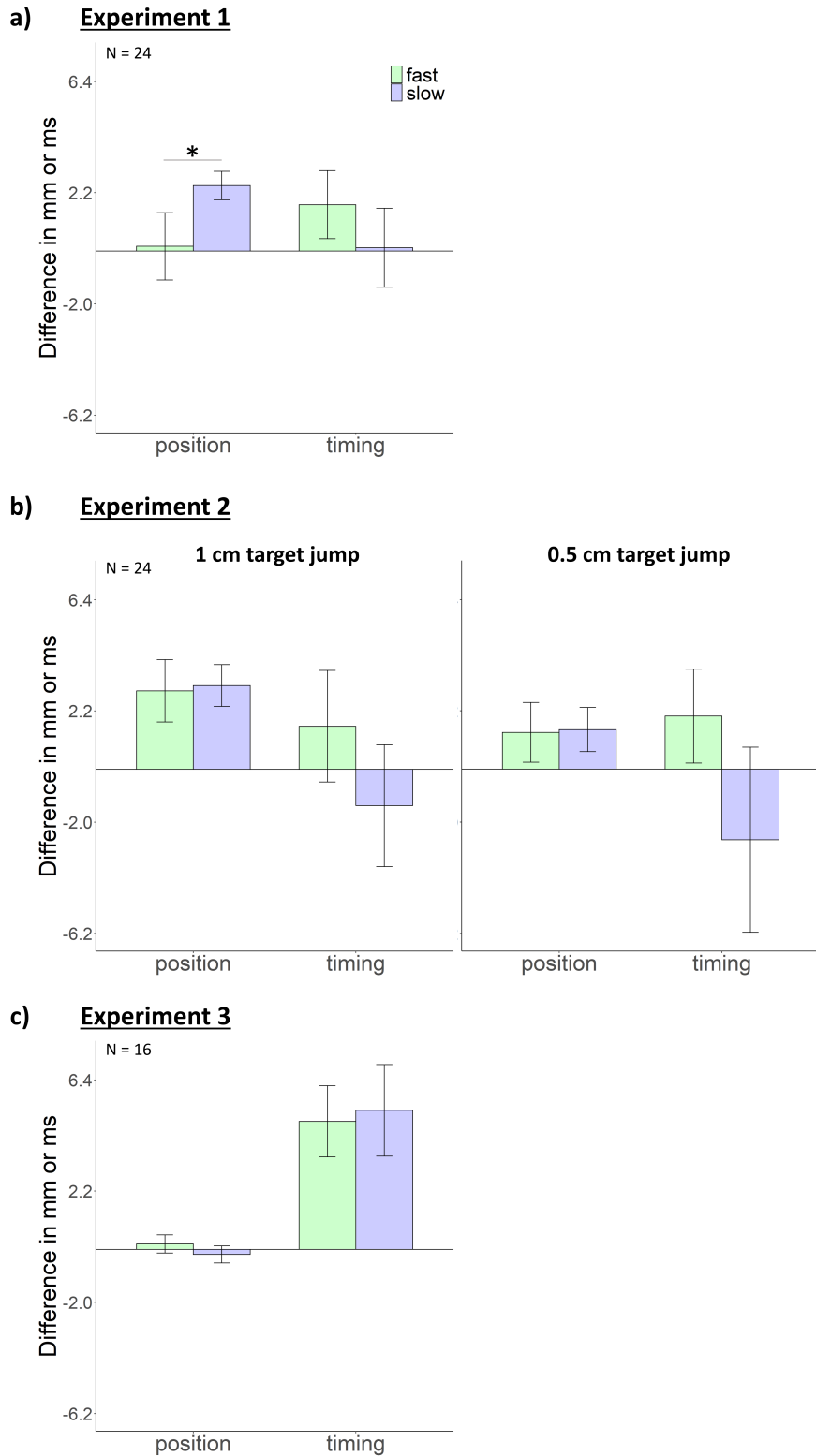


Figure 3.10 - Position and timing adjustment for the two target speeds (green bars: fast speed; violet bars: slow speed) for a) Experiment 1, b) Experiment 2, and c) Experiment 3. Positive values mean coherent adjustment of timing and position in relation to the target jump in each previous trial. Error bars show SEM. \*:  $p < 0.05$ .

## Chapter 4

# MultiTab: a novel technological device to investigate multisensory skills

In Chapter 1, I introduced space and time domains as interrelated dimensions on which we base our perception to represent the external world and directly act on it. Findings of Chapter 2 and Chapter 3 allowed me to demonstrate the various ways in which space, time, and motion communicate with each other. First, I showed domain-specific processing of multisensory inputs that were present also at the cortical level. Secondly, I brought evidence of how the brain balances the processing of spatio-temporal information in motion perception and in the manual interaction with moving objects. Within this background, in Chapter 4 I introduce the design and validation of a new portable device, named MultiTab, for an accessible and robust investigation of the multifaceted interaction between space, time, and motor responses.

As already introduced in Chapter 1, in daily life we are exposed to multisensory cues that our brain integrates to create a coherent representation of the world. Two main principles have been suggested to guide multisensory integration: the spatial and temporal congruency of multisensory stimuli. Specifically, multisensory responses at the behavioral and neural level emerged to be stronger when the sensory components of the multisensory stimulus (e.g., a sound and a light) are presented from the same spatial location (spatial congruency) and close in time (temporal congruency) (Chen and Vroomen, 2013; Meredith et al., 1987; Spence, 2013; Stein and Stanford, 2008; Stevenson et al., 2012; van der Stoep et al., 2017). Multisensory stimulation has been widely proven to boost perception and behavioral performance, compared to when only unisensory stimulations are provided (Burnett et al., 2004; Gingras et al., 2009; Jiang et al., 2002; Stein et al., 1988, 1989). For example, when a person has to detect as fast as possible unisensory and multisensory stimuli, response times are

---

generally faster with the multisensory stimulation compared to the unisensory one (Diederich and Colonius, 2004; Stevenson et al., 2012). This phenomenon of Multisensory Response Enhancement (MRE) is stronger when the spatial and temporal alignment between the bimodal inputs is maintained (Van der Stoep et al., 2015).

A task in which multisensory cues can be particularly useful involves inferring spatio-temporal features of an external event (van der Stoep et al., 2017). For instance, previous studies revealed that in localizing the position of a sound source in space humans and animals benefited from a visual stimulus presented at the same time and in the same spatial location (Bolognini et al., 2007, 2005; Colonius and Diederich, 2004; Hairston et al., 2003; Jiang et al., 2002). This process of multisensory integration in spatial localization has been described as statistically optimal and follows Bayesian rules (Rowland et al., 2007). Indeed, localization decisions are probabilistic since based on sensory information which can be more or less accurate, depending on the neural and environmental noise and on prior knowledge. Crossmodal phenomena were also found to improve spatial performance, with exogenous crossmodal spatial attention that facilitated sensory information processing and subsequent behavioral performance. In this regard, past literature revealed faster responses to auditory sources placed at locations previously cued by visual inputs compared to responses placed at uncued locations, and vice-versa for the localization of visual stimuli previously cued by auditory inputs (Ahveninen et al., 2019; Johnen et al., 2022).

To localize stimuli around us we often use manual responses. By pointing towards a sensory source (either a sound or a flash) in the space, a person translates into motor coordinates the spatial representation of the perceived stimulus, with the target and hand positions that have to be combined into a common reference frame to facilitate precision in the pointing action (Boyer et al., 2013; Redon and Hay, 2005). Similarly, hand movements can also be used to convey temporal information, such as the duration of a perceived stimulus and the rhythm of a temporal sequence (Hildebrandt et al., 2022; Roach et al., 2017). However, motor actions are not simply a measure of perception, as both processes are tightly coupled and influenced by each other (Camponogara and Volcic, 2020). In these regards, it was revealed that performance in a visual discrimination task was enhanced when relevant motor actions were directed to a visual target accompanied by a sound (Elshout et al., 2020). This mechanism seems to subsist also in spatial localization performance since people showed to significantly improve their acoustic localization after multisensory training in which both motor and acoustic cues were provided (Valzolgher et al., 2020).

Given the facilitating role of motor actions in multisensory space perception and vice versa, several experimental paradigms investigating spatial skills have involved the use of

manual responses in the tasks. However, many of these paradigms required complex setups and techniques for the reproduction and recording of these responses. To fill this gap, in Chapter 4 I introduce MultiTab, an easily portable system that provides spatialized auditory and visual stimulations (of which spatial and temporal congruency are guaranteed) and promotes and measures user's motor responses (Bertonati et al., accepted for publication). This new technological device is here presented in its technical and behavioral validation in a spatial localization task, in which we measured the response times and localization errors of participants in unisensory and multisensory conditions. We expected MultiTab to be able to provide audiovisual stimulation within reliable space and time binding windows and yield MRE in participants' performance (i.e., shorter response times when both audio and visual cues were available compared to unisensory conditions), together with smaller localization errors.

## 4.1 MultiTab system

MultiTab is a novel portable device that allows the serial reproduction of spatialized sounds and lights on a flat circuit equipped with tactile sensors for registering the user's touch (Figure 4.1).

### 4.1.1 Hardware

MultiTab consists of two identical modules (tab A and B; 10.5 cm X 10.5 cm). Each tab is equipped with 64 touch sensors (capacitive sensors) covering the entire tab's surface, eight Red Green Blue (RGB) LEDs, and four speakers (mod ASE02008MR-LW150-R). The modules are embedded into a plastic compact structure (for a total size of 12 cm X 23 cm) which portability makes it easy to use in experimental and clinical settings. A USB cable and an audio jack cable complete the system, for the device-remote PC communication and sound production, respectively. Figure 4.1 shows the final configuration of the device. The mechanical design of the device (Figure 4.2) was carried out using PTC Creo Parametric 7.0 (a parametric 3D CAD platform) and all its components were fabricated using Additive Manufacturing (AM), specifically Selective Laser Sintering (SLS) on a 3D System ProX SLS 6100 machine. The custom components were manufactured using Duraform Polyamide 12 (Nylon), a thermoplastic with desirable mechanical properties and fine-feature surface resolution.

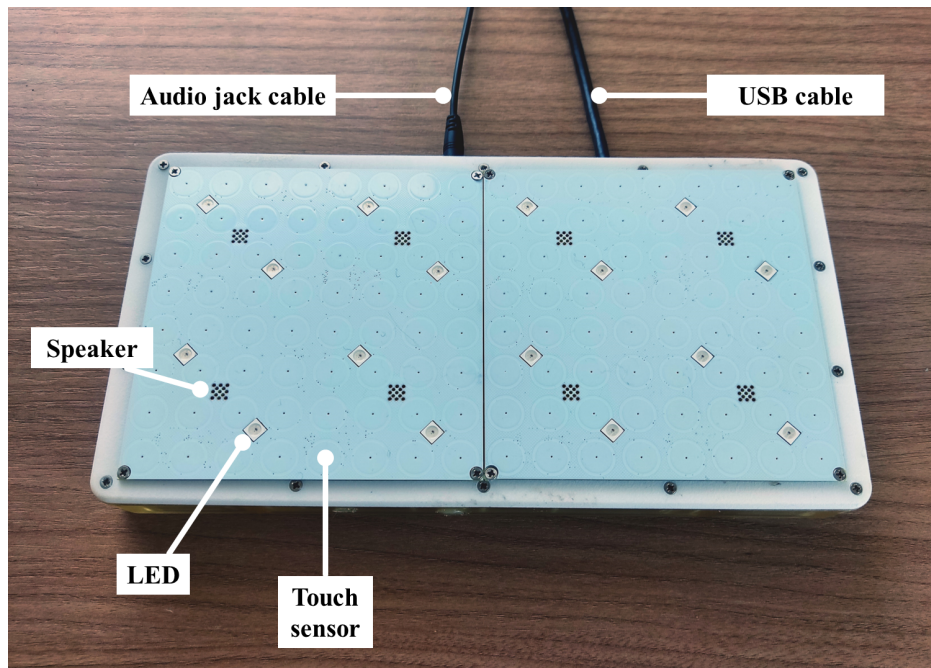


Figure 4.1 - MultiTab system. Between the two modules (tab A on the left and tab B on the right), the device comprises 128 touch sensors, 16 LEDs, 8 speakers, a USB cable, and an audio jack cable.

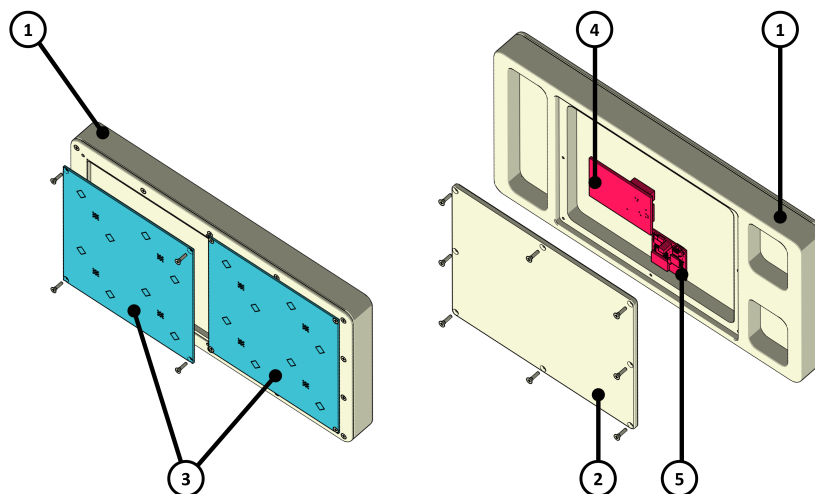


Figure 4.2 - CAD view of the device. 1) outer shell; 2) cover plate; 3) electronic tabs; 4) Seeee XIAO microcontroller; 5) service board.

The flow of information between the remote PC and MultiTab takes place through a USB port with a serial line RS232. The PC USB port also provides the necessary power

supply for the device (5V 500mA). MultiTab is controlled by three microcontrollers: two microcontrollers ATMEGA328P embedded in tab A and B and responsible for interpreting the commands received from the PC; one microcontroller ATSAM21G18P embedded in the Seeeduino XIAO, managing communications with the remote PC and converting the signal into a serial line RS485. Figure 4.3 describes the block scheme of the device.

### 4.1.2 Software

Reproduction of the auditory and visual stimuli and recording of the user's touch by the touch sensors (with a 1.5-by-1.5 cm precision) are carried out in Matlab (Inc., 2019). The flow of information between Matlab and MultiTab takes place by sending commands suitably formatted on the RS232 serial line. Each command consists of:

- 1 starting byte
- 1 peripheral address byte
- 1 unit address byte
- 1 command byte
- 3 bytes identifying any parameters of each command
- 1 ending byte

Operatively, the two tabs of MultiTab work independently and, when activated together, they receive commands from the remote PC simultaneously. Temporal limits of sound and light reproduction, as well as LED colors, are customizable through Matlab. Therefore,

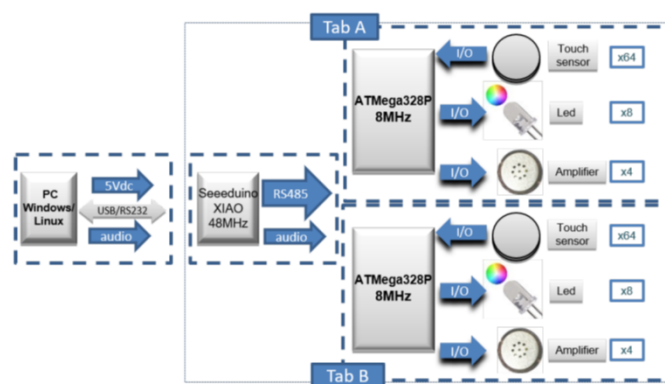


Figure 4.3 - MultiTab block scheme.

MultiTab represents a lightweight and portable multisensory tablet, with highly customizable and strictly controlled stimulation based on experimental and/or rehabilitation needs.

## 4.2 Technical validation

### 4.2.1 Methods

To test the accuracy and precision of MultiTab in playing sounds and lights, we used the Biosemi ActiveTwo System to record auditory and visual stimulation. From each unit (i.e., speaker or LED) of MultiTab, auditory (70 dB, pink noise) and visual (2.8 cd, blue light) stimuli were recorded for 100 ms through a wired PGA58 vocal microphone and a photodiode, respectively. The two sensors were connected to two different channels of the Biosemi system. We recorded 100 auditory and 100 visual stimuli from each unit during unimodal (auditory or visual) and bimodal (audiovisual) reproductions, separately. To evaluate the stimuli production's accuracy and precision, the recorded traces were compared with the supposed stimuli's onset, offset, and duration (100 ms). Specifically, differences between the actual and the supposed beginning or end of stimulation were calculated to determine any delays or advances in playback. The mean duration and standard deviation of the recorded traces were calculated as a measure of the stimuli's production accuracy and precision, respectively. Finally, for the audiovisual stimuli, we graphically inspected the onset and offset of the auditory and visual traces to ensure the simultaneity of the two sensory inputs.

### 4.2.2 Results

Technical validation of MultiTab checked the stimuli production's accuracy and precision, specifically the onset, offset, and duration of the auditory and visual stimulation (unimodal and bimodal). The visual stimuli were precise in the onset and offset of the reproduction (mean onset: 0.24 ms; SD onset: 0.56 ms; mean offset: -1.23 ms; SD offset: 1.80 ms), while the acoustic stimuli resulted to be less precise in the onset of the reproduction (mean onset: 4.15 ms; SD onset: 1.87 ms; mean offset: -1.38 ms; SD offset: 1.68 ms), possibly due to a delay in the execution of the commands. For all the sensory stimulations (i.e., the visual and auditory stimuli from all the units of MultiTab), the mean recorded duration corresponded to the supposed stimulus's duration (100 ms), with negligible variation between multiple recordings from the same unit (Audio: mean: 96.25 ms; SD: 1.85 ms; Visual: mean: 99.69 ms; SD: 1.78 ms; Audiovisual: mean visual: 98.06 ms; SD visual: 0.48 ms; mean audio:



98.21 ms; SD audio: 2.56 ms). For the audiovisual condition, we checked the simultaneity of the stimulation between the auditory and the visual inputs through a graphical inspection of the bimodal traces (an example in Figure 4.4a). In addition, differences in the onset and offset of reproduction between the two stimuli were calculated and did not seem to be substantial (Figure 4.4b).

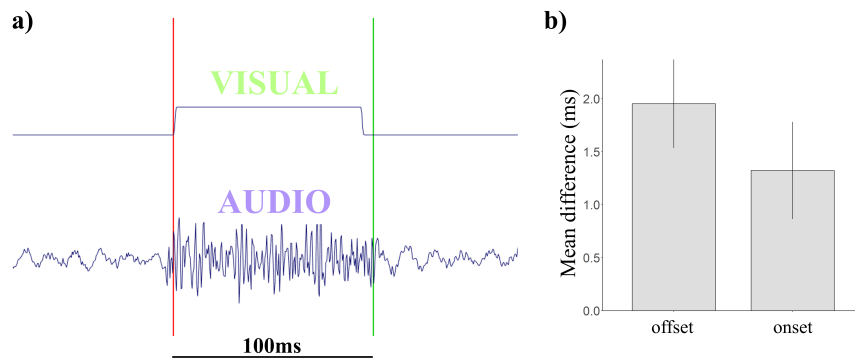


Figure 4.4 - Simultaneity of the auditory and visual stimulation during the audiovisual condition. a) Example of visual (upper) and auditory (lower) recordings' traces of 100 ms temporal interval. b) Mean difference in the onset and offset of the stimulation between the auditory and the visual inputs during the audiovisual condition.

## 4.3 Behavioral validation

### 4.3.1 Methods

#### Participants

13 participants (9 female; mean  $\pm$  SD age:  $26.07 \pm 3.86$  years old) took part in the experimental validation of MultiTab. All participants reported no history of neurological, cognitive, and/or sensory deficits. They gave their written consent in accordance with the local Ethics committee (Comitato etico, ASL3 Genovese) and the Declaration of Helsinki prior to testing.

#### Experimental procedure

After assessing the technical accuracy and precision of MultiTab, we evaluated whether the device can be used to assess multisensory skills. To test this, we asked participants to perform a spatial localization task in auditory, visual, audiovisual congruent (i.e., with spatial and

temporal congruency of audiovisual stimulation), and audiovisual incongruent (i.e., with spatial but without temporal congruency of audiovisual stimulation) conditions.

Participants rested their heads on a chinrest placed at 36 cm from MultiTab and wore specially designed sunglasses with increased opacity due to the addition of transparent plastic material over the lenses. These sunglasses were designed to blur participants' vision (visual stimuli were perceived as degraded) and to likely equalize the localization of the degraded visual stimuli to that of auditory stimuli (which are known to be more difficult to localize in space than visual inputs, Alais and Burr, 2004). The blurred sunglasses were worn for the entire duration of the experiment so that participants were not influenced by seeing the device with their normal vision. The auditory stimuli consisted in a 100 ms pink noise burst (70 dB) and the visual stimuli in a 100 ms blue light (2.8 cd). Each trial started with participants placing their dominant-hand index finger on a starting point marked on the bottom of the device and stimuli delivered after a random duration of between 950 and 1450 ms to prevent stereotyped responses by subjects. Participants were instructed to manually localize the stimuli by tapping the top surface of MultiTab at the position where they perceived the stimuli and then return to the starting point. We asked participants to respond as rapidly and as accurately as possible. Touch sensors of MultiTab recorded the subjects' touch on the surface. Manual responses provided after 1.5 s from the stimulus onset were considered incorrect and deleted from the analysis. The paradigm included four conditions: *i*) in the auditory condition, participants localized a total of 80 auditory stimuli (8 positions repeated 10 times each); *ii*) in the visual condition, a total of 160 visual stimuli (16 positions repeated 10 times each); *iii*) in the audiovisual congruent condition, a total of 160 audiovisual stimuli (16 positions repeated 10 times each), constituted by a pink noise and a blue light reproduced simultaneously for 100 ms (to guarantee temporal congruency), with a 1.5 cm distance between speaker and LED (to guarantee spatial congruency); *iv*) in the audiovisual incongruent condition, a total of 160 audiovisual stimuli (16 positions repeated 10 times each), for which spatial congruency was maintained, while temporal congruency was disrupted by auditory and visual inputs temporally separated by a stimulus onset asynchrony (SOA) of 200 ms. In this last condition, auditory stimulation was always reproduced before the visual stimulation and this was implemented to test whether multisensory integration would occur in lack of temporal congruency (Chen and Vroomen, 2013; Meredith et al., 1987; Spence, 2013; Stein and Stanford, 2008; Stevenson et al., 2012; van der Stoep et al., 2017). In the audiovisual congruent and incongruent conditions, each speaker was possibly accompanied by two LEDs in proximity to it (both at 1.5 cm from it), therefore, in these conditions, one audiovisual position referred to one speaker together with one of the two LEDs, for a total of 16

audiovisual positions covering the entire device. Auditory, visual, and audiovisual (congruent and incongruent) stimuli were reproduced half on the left side and half on the right side of the device (i.e., tab A or B of MultiTab). Conditions (auditory, visual, audiovisual congruent, and audiovisual incongruent) and sides (left and right) were counterbalanced across participants in different blocks, and subjects were informed accordingly at the beginning of each block.

### Data analysis

For each participant, Response Time (RT), i.e., the time between the start of the stimulus and the first touch detected by the device, and Position error, i.e., the distance in cm between the position touched by the participant on the device and the position of the reproduced stimulus, per trial was measured. To investigate participants' multisensory gain in the spatial localization task, the medians RTs and Position errors per condition were calculated and, after assessing that data followed a normal distribution (Shapiro-Wilk test), compared by running two separate one-way ANOVAs with Condition as a within-subject factor with four levels (auditory, visual, audiovisual congruent and audiovisual incongruent). RTs and Position errors of left and right side blocks were merged across conditions as no difference was found between the RTs and Position errors of the two device's sides.

To explore the magnitude of the multisensory gain in RTs, for each participant we calculated the absolute amount of MRE for the audiovisual congruent and audiovisual incongruent conditions separately, as follows (Van der Stoep et al., 2015):

$$MRE = \min[\text{median } RT_{\text{audio}}, \text{median } RT_{\text{visual}}] - \text{median } RT_{\text{audiovisual}}$$

Thus, we statistically compared the average absolute MREs against zero to detect significant multisensory gain (positive MREs indicate multisensory integration, negative MREs multisensory inhibition). In addition, to investigate whether MRE was due to statistical facilitation or multisensory integration, we calculated the audiovisual cumulative distributive function (CDF) and we compared it with the race model given by the sum of the unimodal (auditory + visual) CDFs, for the audiovisual congruent and audiovisual incongruent conditions separately (Miller, 1982; Raab, 1961; Stevenson et al., 2014). The race model assumes that if the visual and auditory channels are independent, responses to redundant signals (i.e., in the audiovisual conditions) are especially fast since they are produced by the fastest (the winner of the race) of the response times across the unisensory conditions. We checked for violations of the race model inequality to establish the occurrence of multisensory integration mechanisms instead of statistical facilitation within our data. To test this, we run one-tailed

t-tests against zero for nine deciles (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9; p values were corrected for nine comparisons using the Bonferroni method) of the difference between the race model and the audiovisual CDFs.

### 4.3.2 Results

MultiTab was used to test participants' multisensory spatial skills. We compared participants' median RTs during unisensory (visual or auditory) and multisensory (audiovisual congruent and incongruent) conditions. Results showed a significant difference among the four conditions in the median RTs of participants localizing stimuli on MultiTab (Figure 4.5a;  $F_{(3,36)} = 35.05$ ,  $p < 0.001$ , partial  $\eta^2 = 0.74$ , 95% CI [0.61, 1]). Post-hoc two-tailed t-tests (after Bonferroni correction for multiple comparisons) revealed that participants were significantly faster in localizing audiovisual congruent (i.e., SOA = 0 ms) stimuli than visual-only ( $t_{(12)} = -4.64$ ,  $p < 0.01$ , Cohen's  $d = -1.28$ , 95% CI = [-2.06, -0.51]), auditory-only ( $t_{(12)} = -7.82$ ,  $p < 0.001$ , Cohen's  $d = -2.17$ , 95% CI = [-3.21, -1.12]) and audiovisual incongruent (i.e., SOA = 200 ms;  $t_{(12)} = -10.199$ ,  $p < 0.001$ , Cohen's  $d = -2.82$ , 95% CI = [-4.10, -1.64]) stimuli. In addition, statistical analysis showed a significant difference between the performance in the auditory and visual conditions, with shorter median RTs in the localization of visual inputs ( $t_{(12)} = 4.47$ ,  $p = 0.004$ , Cohen's  $d = 1.24$ , 95% CI = [0.47, 2.00]), which can be explained by the greater precision of the visual system in processing spatial information compared to the auditory modality (Alais and Burr, 2004). Finally, while median RTs were similar between the auditory and audiovisual incongruent condition ( $t_{(12)} = 1.26$ ,  $p = 1$ , Cohen's  $d = 0.34$ , 95% CI = [-0.23, 0.93]), a significant difference was revealed between the latter and the visual condition, with longer RTs when both visual and auditory stimuli were presented but the sound preceded the light by 200 ms ( $t_{(12)} = 12.49$ ,  $p < 0.001$ , Cohen's  $d = 3.46$ , 95% CI = [1.94, 4.97]).

For what concern Position errors, statistical analysis revealed a significant difference between conditions (Figure 4.5b;  $F_{(3,36)} = 264.2$ ,  $p < 0.001$ , partial  $\eta^2 = 0.96$ , 95% CI [0.93, 1]), which was explained by post hoc t-tests (after Bonferroni correction for multiple comparisons) as auditory stimuli being localized with less accuracy by participants compared to visual ( $t_{(12)} = 17.45$ ,  $p < 0.001$ , Cohen's  $d = 4.83$ , 95% CI = [2.79, 6.88]), audiovisual congruent ( $t_{(12)} = 19.85$ ,  $p < 0.001$ , Cohen's  $d = 5.50$ , 95% CI = [3.20, 7.80]) and audiovisual incongruent ( $t_{(12)} = 18.72$ ,  $p < 0.001$ , Cohen's  $d = 5.19$ , 95% CI = [3.01, 7.37]) stimuli. Overall, these results suggested that auditory and visual inputs, when presented simultaneously in space and time (i.e., SOA = 0 ms), elicited shorter RTs in the localization compared to

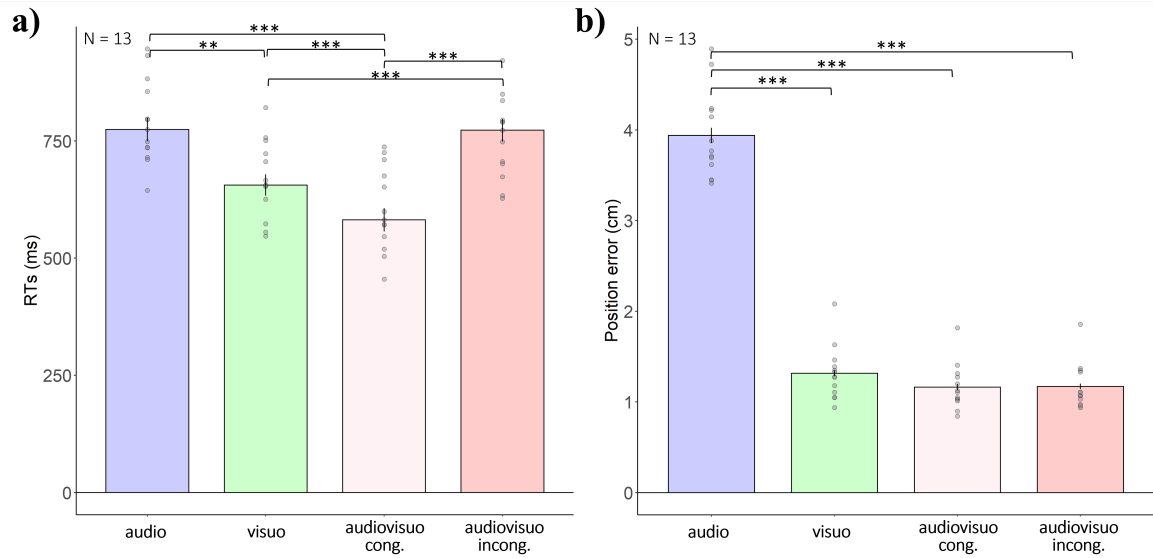


Figure 4.5 - Behavioral performance in the spatial localization task. a) Median RTs (ms) in the auditory (blue), visual (green), and audiovisual (pink and red) conditions. b) Median Position errors (in cm, calculated as the difference between the actual and the touched source position) in the auditory (blue), visual (green), and audiovisual (pink and red) conditions. Error bars show standard errors. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

when unisensory (auditory or visual) stimuli were presented. However, for what concern accuracy in the localization performance, audiovisual stimuli elicited benefits compared to inputs presented in the auditory modality only. Finally, our results revealed that, when audiovisual stimulations were not congruent in time (i.e.,  $SOA = 200$  ms), the multisensory gain in RTs did not occur, with performance similar to that in the auditory condition and worse than that in the visual condition. Since participants were instructed to respond as fast and accurately as possible, in this condition participants may have responded to the first reproduced stimulation (i.e., the sound), without waiting for the occurrence of the visual stimulation, and therefore they may not have benefited from the speeding up of the latter.

To further explore the multisensory gain in RTs, first, we calculated the absolute amount of MRE, to rule out the possibility that, given the great performance of vision in spatial processing (Alais and Burr, 2004), participants preferentially responded to the visual inputs even when localizing audiovisual stimuli (positive MREs indicate a multisensory enhancement with respect to the faster unisensory condition, negative MREs multisensory response inhibition). Then, we analyzed the amount of race model violation (i.e.,  $CDF_{A+V} - CDF_{AV} > 0$ ), to determine whether the observed multisensory gain could be explained in terms of statistical facilitation or multisensory integration. In the audiovisual congruent condition, participants

showed a positive multisensory gain that significantly differed from zero (Figure 4.6;  $t_{(12)} = 4.64$ ,  $p < 0.001$ , Cohen's  $d = 1.28$ , 95% CI = [-0.04, 2.61]), suggesting that the observed MRE was the result of multisensory processing and not of participants preferentially responding to the fastest unisensory condition. On the contrary, in the audiovisual incongruent condition, results revealed a significant negative MRE (Figure 4.6;  $t_{(12)} = -12.49$ ,  $p < 0.001$ , Cohen's  $d = -3.46$ , 95% CI = [-5.37, -1.55]) which indicates that, when auditory and visual stimuli were separated in time, participants did not combine the two percepts and showed a multisensory response inhibition. These results were confirmed by statistical analysis on the race model violation (Figure 4.7). Comparing the CDFs of the audiovisual congruent condition to those of the race model established that race model inequality was significantly violated at the 10th, 20th, 30th, and 40th percentile of the distribution (Figure 4.7b), thus, the amount of MRE observed in this condition was the result of multisensory integration. Significant differences between the race model and the observed data in the audiovisual incongruent condition were also revealed for all percentiles, but in this case multisensory integration mechanisms did not occur since  $CDF_{A+V}$  was higher than  $CDF_{AVinc}$ . Taken together, these results showed multisensory integration effects in the localization of audiovisual spatialized stimuli with MultiTab, but only when the spatial and temporal congruency assumptions were warranted. Overall, these findings were in line with the multisensory response enhancement revealed in

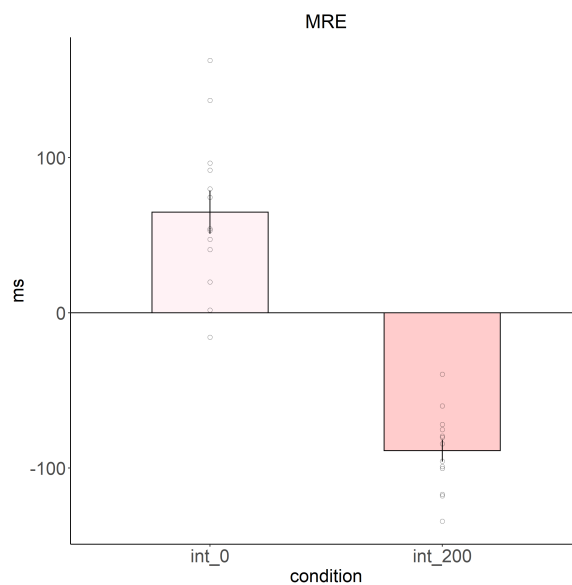


Figure 4.6 - MREs of participants in the audiovisual congruent (pink) and incongruent (red) conditions. Positive MREs indicate a multisensory gain in RTs with respect to the faster unisensory condition. Negative MREs indicate multisensory response inhibition.

previous studies investigating multisensory skills (Diederich and Colonius, 2004; Stevenson et al., 2012; Van der Stoep et al., 2015).

## 4.4 General discussion

In Chapter 4, I presented and validated a novel technology, named MultiTab, able to reproduce auditory and visual inputs, as well as to measure the user's manual responses on the device's top surface. Results indicated that MultiTab reliably reproduced auditory and visual stimulation in unisensory and multisensory conditions and induced multisensory integration in a spatial localization task.

From technical validation of MultiTab, we revealed that the device allowed the accurate reproduction of auditory (pink noise) and visual (blue light) stimuli. Although we measured some small imprecision of the device in reproducing sounds, we also showed that these inaccuracies were stable (i.e., SDs of multiple sound recordings from the same unit were small), thus they may be easily corrected by acting at the device programming level. Behavioral validation of MultiTab, implemented in the test of spatial localization abilities of a group of

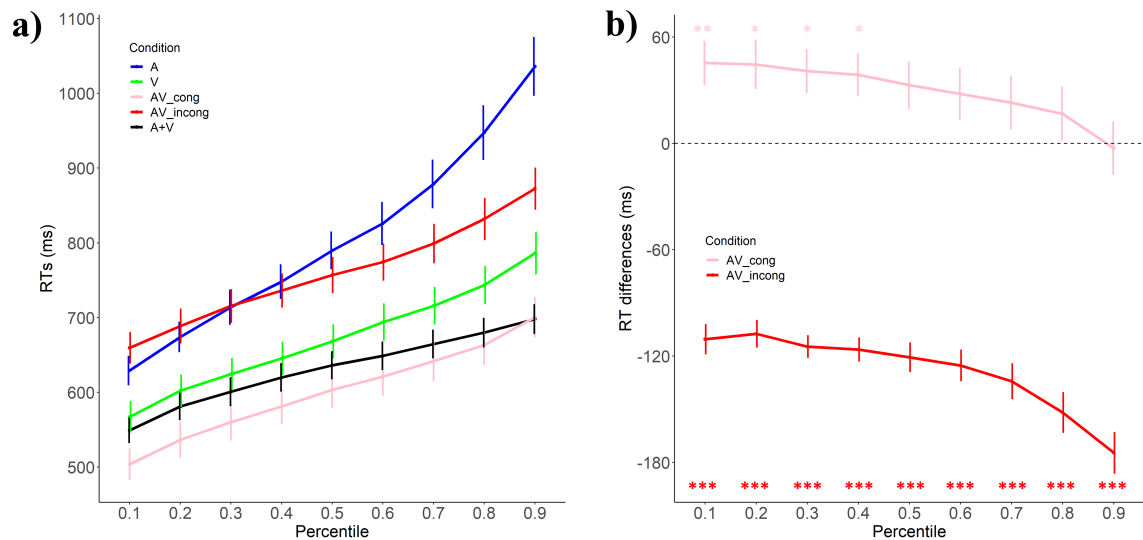


Figure 4.7 - Race model analysis. a) Mean CDFs of response times to auditory-only (blue), visual-only (green), audiovisual congruent (pink), audiovisual incongruent (red) stimuli, and A+V model (black). b) Tests of race model inequality violation (i.e.,  $CDF_{A+V} - CDF_{AV} > 0$ ), are presented for each multisensory condition (pink: audiovisual congruent, red: audiovisual incongruent) and percentile (0.1-0.9). Asterisks indicate significant differences between the observed data and the race model (after Bonferroni correction for nine comparisons). Error bars indicate SEM. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

13 healthy people, revealed that participants exhibited smaller errors in touching the stimuli positions when these were bimodal or visual, compared to auditory-only stimuli. This finding is in line with past results showing benefits in the localization of a sound source in space when also a visual stimulus was presented (Bolognini et al., 2007, 2005). The fact that we could not reveal the same result with visual stimuli (i.e., significant improvements in the audiovisual conditions compared to the visual-only condition) is probably due to the fact that the visual system alone was already sufficiently accurate in the localization of stimuli on MultiTab (Alais and Burr, 2004), so as not to benefit from the addition of an acoustic stimulus, at least as regards position errors. In this sense, a limit of MultiTab could be reproducing stimuli that are not equally reliable between the visual and acoustic modalities, at least for tasks of a spatial nature. We tried to address this inequality by blurring participants' vision and reducing the reliability of their visual perception. Future implementations of MultiTab may also include more salient acoustic stimuli which would increase their relevance for spatial unisensory and multisensory localization. In addition, it is worth noting that we asked participants to be as fast and accurate as possible in responding to the stimuli, thus they may have prioritized speed of response over accuracy. Confirmation of this hypothesis comes from the results on response times. In this analysis, we found significantly reduced manual response times in the localization of audiovisual stimuli, compared to auditory-only and visual-only conditions. This multisensory advantage was the result of multisensory integration mechanisms rather than statistical facilitation, as proved by the violation of race model inequality. Overall, these results were in line with past studies revealing MRE and multisensory integration in various experimental tasks (Diederich and Colonius, 2004; Stevenson et al., 2012; Van der Stoep et al., 2015). It is worth noting the lack of multisensory integration and MRE exhibited in the audiovisual incongruent condition (i.e., with the onsets of auditory and visual inputs separated by 200 ms), which, on the one side, confirmed the temporal congruency principle of multisensory integration (Chen and Vroomen, 2013; Meredith et al., 1987; Spence, 2013; Stein and Stanford, 2008; Stevenson et al., 2012; van der Stoep et al., 2017), and, on the other side, suggested that MultiTab is able to provide reliable time binding windows thanks to the customization of user-defined temporal limits of stimuli reproduction. For what concern spatial congruency, the spatial proximity between auditory and visual sources seemed to be sufficient to produce multisensory integration. Future works may take advantage of the design of the device to propose conflicting spatial information from different senses (sounds and lights reproduced in two different spatial locations on MultiTab) and investigate the degree to which each sense dominates perception (Alais and Burr, 2004; Ernst and Banks, 2002).



Overall, the technical and behavioral validation demonstrated the device's reliability in the investigation of multisensory spatial skills, suggesting the use of MultiTab as a more portable alternative than the canonical experimental methods, which otherwise may require complex systems for promoting and recording manual responses. This perspective stimulates to implementation of the device in the evaluation of other multisensory skills in which an action from the participant is required. In addition, MultiTab, thanks to its versatility and portability, is a promising technology that may have relevant contributions in the clinical field, facilitating assessment and rehabilitation protocols. This last aspect is further discussed in Chapter 5.

*Chapter 4 is partially extracted and adapted from Bertonati, G., Casado-Palacios, M., Crepaldi, M., Parmiggiani, A., Maviglia, A., Torazza, D., Campus, C. and Gori, M. (2023). MultiTab: A Novel Portable Device to Evaluate Multisensory Skills. Accepted for publication in IEEE EMBC.*

# Chapter 5

## General conclusions

The overall purpose of the current doctoral thesis is to investigate how visual and auditory sensory modalities shape the representations of space, time, and motion. To this end, in Chapter 1 I summarized the main scientific findings related to the encoding of spatio-temporal information within a multisensory framework. In Chapter 2 I presented two electrophysiological studies exploring the domain-specific organization of multisensory spatial and temporal processing at the cortical level. In Chapter 3 I investigated the interplay between space and time domains in the perception and interaction with moving objects. Finally, in Chapter 4 I presented the development of a new science-driven technological device named MultiTab. In the present chapter, I will summarize the main results of Chapter 2, Chapter 3, and Chapter 4, I will reconcile these outcomes with the existing research literature, and I will draw a conclusion on the implications of the present dissertation for future research questions.

### 5.1 The contextual nature of multisensory processing

This thesis presented different studies employing a variety of methodological approaches (EEG technique, psychophysical methods, motion capture technique, and technological development) intending to increase knowledge about how humans, with and without sensory disabilities, process space, time, and motion at the behavioral and cortical level.

Following the theoretical background illustrated in Chapter 1, which highlighted the crucial role of vision and audition in shaping the spatial and temporal representation, respectively, in Chapter 2 I deepened this sense-domain relationship at the cortical level, in a double-sided investigation. On the one hand, results showed a domain-specific cortical organization of multisensory processing, with occipital areas specialized in the processing of

the space domain and temporal areas of the time domain (Section 2.1). On the other hand, findings revealed that the neural modulation of occipital areas was also influenced by the nature of the multisensory spatial information to represent, with an early occipital component (50-90 ms) sustaining more the spatial metric representation of three audiovisual stimuli in space (spatial bisection), while a later occipital response (110-160 ms) more strongly characterizing the processing of a single audiovisual stimulus in space (spatial localization) (Section 2.2). In Chapter 3 I explored the use of spatio-temporal cues in motion processing. First, given the crucial role of vision in the processing of the space domain, I investigated the acoustic motion perception of early blind and sighted individuals, and I revealed that the lack of visual experience influences the ability to process the speed of moving sounds by altering how blind individuals make use of the sounds' temporal features (Section 3.1). Secondly, on the basis of findings that outlined the great flexibility of the sensorimotor system in the accomplishment of task demands, I studied how this aspect emerges in the motor interception of moving objects. Results showed that typical people manually intercepting a moving target took into consideration the spatio-temporal cues of the item when they adjusted their interceptive movements (Section 3.2). Finally, in Chapter 4 I presented MultiTab, a novel portable device that induced multisensory integration mechanisms in the manual localization of unisensory and multisensory stimuli on its surface. Overall, the studies conducted and included in the present thesis have outlined some of the ways in which space, time, and motion are conveyed and regulated by the different senses, to shape our interaction with a multisensory world.

### **5.1.1 Domain-specificity of multisensory representations at the cortical level**

Empirical evidence suggests a supramodal nature of brain organization that reflects an abstract representation of multisensory stimuli at the cortical level, which does not depend uniquely on the sensory modality delivering the inputs (Cecchetti et al., 2016), but also on other aspects such as the domain of representation (i.e., space and time) (Amadeo et al., 2020a; Campus et al., 2017; Gori et al., 2022b). Despite the increasing knowledge of the mechanisms underlying multisensory perception, several aspects of how multisensory processing fits into this theoretical framework were still unknown. Studies presented in Chapter 2 aimed to further investigate this topic. Results of Study 1 and Study 2 fit into a framework delineated by Murray et al. (2016a), who proposed that the multisensory perception does not always involve a single and fixed schema of neural activation, but encompasses different cortical

circuits. Specifically, they proposed the existence of a neural circuit that involves high-order association cortices, such as the prefrontal cortex and the superior temporal sulcus, and of a second neural circuit where interactions occur directly between low-level cortices (somatosensory, visual, and auditory cortices). The authors suggested that multisensory processes can involve both kinds of schema based on the nature of the multisensory stimuli to be processed and on the kind of stimuli features guiding multisensory mechanisms across the lifespan. Multisensory abilities emerge gradually during development, starting from favoring the processing of low-level stimulus characteristics, up to the elaboration of higher-level multisensory representations, as a result of one's experience with the world. Flexibility of multisensory processing also persists in adulthood, within the learning of new multisensory relations in a more short-term time scale (e.g., during training). Results of Chapter 2 fit into this dynamic and flexible view of multisensory processes at the behavioral and cortical levels. Specifically, Study 1 showed that the neural activation associated with multisensory stimuli was modulated by the domain of representation (space or time) of the audiovisual inputs, with stimuli' spatial characteristics especially elaborated by occipital regions and stimuli' temporal features by temporal areas. In Study 2, we provided further evidence of this view by revealing that audiovisual inputs that were part of a type of spatial representation (e.g., a single point space layout) rather than another (e.g., a metric space representation of three stimuli) modulated cortical activity of occipital areas. Overall, domain-specificity (see Study 1) and representation-specificity (see Study 2) of multisensory processing may be the result of sensory experience-dependent developmental mechanisms that select vision (and the associated occipital cortices' activity) as the main support for the spatial abilities, and hearing (and the associated temporal cortices' activity) for the temporal abilities. However, these findings should be considered in light of the lack of unimodal (auditory or visual) conditions in both Study 1 and Study 2 which limited the possibility of confidently assuming that the activation of sensory cortices we revealed resulted entirely from multisensory processing. Thus, further investigations in this direction are still needed.

Another important point of discussion derived from findings exposed in Chapter 2 are cortical multisensory processes occurring over a wide range of temporal latencies. Past studies already showed multisensory mechanisms arising at different steps of the perceptual encoding, also including latencies before 100 ms from the stimulus onset (Bolognini et al., 2010b; Bueti and Macaluso, 2010; Fort et al., 2002; Foxe et al., 2000; Ghazanfar and Schroeder, 2006; Giard and Peronnet, 1999; Molholm et al., 2002; Murray et al., 2005; van Wassenhove and Grzeczowski, 2015). So far, early multisensory integration (eMSI; De Meo et al., 2015) was generally shown to be elicited by simple tasks such as discrimination

or detection tasks. In the two studies of Chapter 2, we could investigate the occurrence of multisensory mechanisms at early latencies also with more complex requests, such as the bisection tasks, which explored the human ability to build a metric representation of the environment by estimating and comparing different inputs in space and time. The fact that we revealed a domain-specificity of multisensory processing within early time latency (50-90 ms from stimulus onset) with this kind of task can be regarded as a controversial point since early multisensory integration is typically considered an automatic process, i.e., a hallmark of bottom-up mechanisms (De Meo et al., 2015). Nevertheless, past studies revealed that also top-down factors, such as attention, influenced multisensory integration within very premature stages of stimulus processing (Talsma et al., 2010; Talsma and Woldorff, 2005), and that high-level cognitive processes can directly involve the recruitment of auditory and visual areas (reviews on the visual areas: Ricciardi et al., 2020; Roelfsema and de Lange, 2016; review on the auditory areas: Zatorre, 2007). However, the reason why in our studies the neural modulation of sensory cortices by domains of representation occurred across both earlier (before 100 ms from the stimulus onset) and later (110-160 ms) neural components, remains a matter of interpretation. We speculate that the cross-sensory calibration theory (Gori, 2015) may explain these results. This theory hypothesizes that for a kind of task such as the spatial and the temporal bisection, the visual and the auditory systems may calibrate the other senses for the spatial and the temporal representations, respectively, and this may require the support of early responses (i.e., 50-90 ms after stimulus onset) of sensory areas. In this regard, it has been suggested that “the functional implications of such “early” multisensory influences are likely to be substantial in that they suggest a previously unrealized level of interaction that could confer tremendous flexibility on sensory function (Van Atteveldt et al., 2014)” (Murray et al., 2016a). In other words, an early neural computation of multisensory stimuli can provide the flexibility needed for the system to respond to complex requests such as those of the bisection task. On the contrary, since it was revealed that the ability to localize stimuli in the space develops before multisensory integration (Rohlf et al., 2020) and even in the absence of visual experience (Gori et al., 2021), for a spatial localization task visual calibration may not be necessary. Thus, the recruitment of early occipital components may be less substantial for a spatial localization task, which conversely involves later cortical processes (after 100 ms from stimulus onset). That developmental processes play a role in the modulation of multisensory mechanisms is also suggested by the fact that children, who can well perceive low-level stimulus characteristics, gradually learn through experience to make use of these features to scaffold the representation of higher-level stimulus characteristics which progressively become part of more complex and

sophisticated mental processing (Murray et al., 2016a). In adulthood, this balance between stimulus characteristics and more complex learned associations would end with context (i.e., task, domain, goal) completely shaping multisensory mechanisms, also independently of the physical stimulus characteristics (De Meo et al., 2015; Sarmiento et al., 2015; Talsma et al., 2010; ten Oever et al., 2016; Van Atteveldt et al., 2014).

In this section, I discussed the results of Chapter 2 in light of the dynamic interaction between the sensory modalities conveying environmental inputs and the more complex associations acquired through the experience. Both these aspects can fall within the definition of context, which is the aspect that determines the final product of humans' multisensory abilities. In the next section (5.1.2), I will discuss how context can be used to cope with sensory deprivation and to correct motor behaviors.

### **5.1.2 Adaptive and maladaptive use of spatio-temporal sensory cues**

In the previous section (5.1.1), I discussed that multisensory mechanisms are not fixed or uniform but depend on contextual factors. For the sake of clarity, I mean by "context" those (short-term and long-term) external situations in which the brain operates to process inputs coming from multiple sensory modalities. The external factors defining the context can be the physical properties of a sensory event, the behavioral goal, the individual's past experience, etc. (Van Atteveldt et al., 2014). In this regard, in the present thesis, I already showed that the sensory experience of individuals shapes their multisensory abilities, as suggested by the modality appropriateness hypothesis (Welch and Warren, 1980). This hypothesis proposes that the sensory experience determines the sense that provides the most accurate information for multisensory processing, with visual stimuli that dominate the processing of the spatial characteristics of an audiovisual event, and auditory stimuli the encoding of temporal characteristics. Then, I discussed how this contextual influence transforms over time into a goal/task-dependent influence on multisensory mechanisms (Talsma et al., 2010; ten Oever et al., 2016), for which aspects such as the domain of representation (i.e., space and time) influence the behavioral and cortical processing of multisensory inputs. Together with this top-down control, also bottom-up mechanisms are part of the contextual influence on multisensory perception. This aspect was shown in Chapter 4, in the investigation of multisensory spatial localization abilities through the novel device MultiTab, where the spatial and temporal congruency assumptions of audiovisual inputs for multisensory integration were demonstrated. So far, the discussion on the influence of context on multisensory perception has focused on the healthy population. However, studies of Chapter 3, which altogether

investigated the interaction between space and time domains in motion processing, gave further insights into the context-multisensory processing relationship, by revealing how the human brain *i*) adapts to sensory deprivation (in the case of blindness, as shown in Study 3, and *ii*) optimizes motor behaviors to meet task demands (for interception tasks, as described in Study 4). Overall, these results suggested that responding to contextual demands can lead to both adaptive and maladaptive behaviors.

An example of adaptive performance is shown in Study 4, in which I investigated how a person, when intercepting moving objects, corrects hand movements in relation to previous interceptive errors and if this mechanism is regulated by the target speed. Results revealed that interceptive movements were corrected in relation to the feedback presented in previous attempts, also based on the target speed, with a tendency to adjust the timing of interception more for fast targets, and the position of interception more for slow targets. Thus, sensorimotor adjustment was adaptive to task demands, that in this case was the interceptive component (either spatial or temporal) that was more relevant according to the target speed. Previous studies already revealed our ability to adapt motor performance to contextual factors. For instance, the timing of interception was revealed to be adapted to some impact requirements (e.g., the place of the interception), by balancing the sensory information-movement coupling of the sensorimotor event (Caljouw et al., 2004; Rushton and Wann, 1999). Always with regards to interception, Brenner and Smeets (2015) delineated a model of how we modulate different aspects of hand movements to meet high precision in the interception of moving objects. This model considers all relevant sources of interceptive errors, together with the sensorimotor delays with which the position and timing of interception can be adjusted. Specifically, the authors sustained that the brain, to achieve the amazing precision previously revealed in different interception tasks, first judges the time needed to reach a target and then tunes the precise spatial position for the interception. This is done by considering the higher delay in updating the temporal properties of interception, compared to the delay in updating the spatial features. Overall, these examples showed the human ability to adapt interceptive movements in a clever manner, by considering task and sensorimotor system constraints. With the results of Study 4, we added further evidence of this ability by showing that it occurs also when the brain is asked to correct interceptive movements on the basis of previous errors, thus also during motor learning. In addition, these results told something about how motor and visual processes communicate with each other. This link was found, for example, in the beneficial effects of real-time visual feedback of one's own action (and, conversely, in the negative influences of presenting delayed visual feedback with respect to motor performance on aspects such as perceived sense of agency, sense of ownership, and the

motor performance itself; Imaizumi and Asai, 2015; Jörg et al., 2012; Longo and Haggard, 2009; Waltemate et al., 2016). Results of Study 4 deepened the link between vision and action by showing that visual feedback can influence not only the immediate motor performance but also the programming of future manual responses. However, this aspect needs to be further discussed. For instance, it would be interesting to investigate whether the influence of previous interceptive movements on the future estimation of target visual properties and manual responses towards this target, follows serial dependencies rules (Cicchini et al., 2014; Fischer and Whitney, 2014). In this regard, recent studies revealed that in the localization of visual items, motor actions produced the necessary error signal on which to base the recalibration of visual localization, suggesting a strong communication between motor and visual processes (Cont and Zimmermann, 2021; Zimmermann and Lappe, 2016). Another way in which context (in this case the sensory experience) has been shown to lead to at least partially adaptive behaviors is the motion acoustic perception investigated in Study 3. There I described that sighted and early blind individuals reported a temporal dominance in the discrimination of sounds' speed through the use of spatio-temporal cues (graphically shown by the vertically oriented ellipses in Figure 3.5). This result confirmed past findings on sighted adults (Carlile and Best, 2002; Carlile and Leung, 2016; Freeman et al., 2014; Locke et al., 2016), but was the first revealed in the visually impaired population. The fact that participants based their speed discrimination on the temporal components of moving sounds can be considered an adaptive behavior to the contextual sensory experience since it would rely on the well-known highest temporal acuity of the acoustic system (Barakat et al., 2015; Bresciani and Ernst, 2007; Burr et al., 2009; Guttman et al., 2005).

However, apparently adaptive mechanisms (such as the temporal dominance described above) may also turn into misleading behaviors in some specific circumstances. In this regard, a qualitative analysis of Study 3 revealed that both sighted and early blind individuals followed a temporal assumption that identified as faster those sounds with a shorter duration, and that this assumption led to a drop in the speed discrimination performance under some conditions. Thus, related to this finding, in the following lines maladaptive behaviors in response to contextual factors will be described. At first sight, this temporal assumption recalls a model described by Gori and colleagues (Gori et al., 2020a) explaining a possible strategy implemented by blind individuals to overcome their spatial metric issues. The authors hypothesized that when visual information is missing (due to blindness) spatial abilities seem to be compromised so blind people rely on temporal features of external events to infer the spatial information of the same event (when spatio-temporal information is conflicting; see Amadeo et al., 2020b; Gori et al., 2018). This mechanism was explained by



visually impaired individuals assuming a constant velocity of environmental stimuli. Despite this compensatory behavior being shown to be efficient with static stimuli (implemented in a bisection task similar to that presented in Chapter 2), it may not be useful with moving stimuli like those used in Study 3. Indeed, in this case, the stimuli presented to sighted and early blind participants already contained velocity information (which in turn determined the spatial and temporal information of the stimuli), which were, in some circumstances, in conflict with the temporal assumption of “shorter sound duration is equal to faster speed” (for more details on this temporal assumption see Section 3.1). In other words, by ignoring the velocity information already inherent in the moving stimuli, blind individuals were less precise than sighted controls in speed discrimination. Where this temporal assumption comes from is ground for speculation. It may be attributed to previous knowledge of individuals that reflects the probability distribution of the environment’s speed property (Parise et al., 2014; Senna et al., 2015, 2017). Since environmental statistics help humans to predict the events of the world, we speculate that relating fast speeds to short duration may be a preventive way to react rapidly to stimuli that are approaching quickly. Further investigations in this direction are desirable, also considering that statistical learning (intended as the ability to extract information about the distribution of a certain event in the environment and to learn from this in order to predict the same or similar events in the future) has been described as a form of the context-dependent control of multisensory processing (ten Oever et al., 2016). In this section, I discussed the results of Chapter 3 in light of how the influence of context on multisensory processing can lead to adaptive or maladaptive behaviors. Nevertheless, it is important to underline that these two outcomes should not be taken as uniquely distinct consequences, as they are interrelated parts of the same phenomenon. Just consider the concept of bias which, depending on the situation, can be interpreted only as a deviation from the norm, or as a mechanism functionally relevant in daily life. For instance, in the study of the vestibular system, overestimation of self-motion perception (Israël et al., 1995; Mackrous and Simoneau, 2011) has been interpreted as an adaptive way to maintaining the straight-ahead direction, thus as an attempt of the sensory system to optimize behavioral performance (Crane, 2012; Cuturi, 2022; Cuturi and MacNeilage, 2013; Zanchi et al., 2022). Nonetheless, considering both possible scenarios of the influence of context on our multisensory perception is crucial to better define our abilities and prevent deficits of adaptation to the environment. This perspective will be discussed in the next concluding section (5.2).

## 5.2 Concluding remarks

The current thesis points out that the nature of multisensory processing is multidimensional, with multisensory mechanisms being elicited by various contextual factors (sensory experience, domains of representation, task goals, etc.). It also describes the interaction between the dimensions of space, time, and motion in its most florid function, which is keeping the complex representation of the surrounding world together. Finally, the present dissertation illustrates how a disequilibrium between these dimensions may also lead to maladaptive phenomena in the interaction with the environment. Assessing and dealing with the presence of maladaptive behaviors is extremely important to consider, especially to address possible compensatory mechanisms implemented by populations with specific deficits (sensory or others). In fact, changes due to impairments to any human system can be considered adaptive when there is at least partial compensation for the damage, or maladaptive when the rearrangement of multisensory networks leads to atypical multisensory experiences (Bolognini et al., 2013). Chapter 4 of the present thesis proposed MultiTab as a technological solution to this perspective. Having confirmed its technical and experimental reliability, MultiTab can be used for the evaluation and training of multisensory abilities. The hope is to exploit the neuroscientific results discussed in the present thesis in the application of MultiTab also to clinical contexts. For instance, the lack of assessment and rehabilitation tools providing multisensory contingencies in the treatment of visual disabilities was pointed out in some reviews, which identified the main issues of the existing systems in their invasiveness, low acceptance, and long training (Cuturi et al., 2016; Gori et al., 2016). There are also good examples of assessment (Bertonati et al., 2020) and rehabilitation (Cappagli et al., 2017; Morelli et al., 2020; Purpura et al., 2017) tools being effective if used within a multisensory-based approach. However, the ultimate goal of this rehabilitation process, i.e., the best possible quality of life for the person with visual disabilities, is still not often completely achieved (Elsman et al., 2019). MultiTab could be used in this context to assess and train acoustic spatio-temporal abilities of visually impaired children and adults, for example, to mitigate temporal attraction during perception when non-functional. In addition, MultiTab, as a highly portable system, may elicit a greater degree of user acceptability due to its compact design and the possibility for the user to interact with it through a natural behavioral response such as the manual response (e.g., pointing to localize a source). Similarly, but in the case of brain damage, multisensory technologies providing auditory, visual, tactile, kinaesthetic, and sensorimotor feedback to the person were found to be at least comparable, if not better, to standard rehabilitation treatment of adults and children with brain lesions (Bolognini et al.,

---

2013; Bolognini and Vallar, 2019; Fortis et al., 2010; Frassinetti et al., 2005; Johansson, 2012; Parisi et al., 2022; Tinelli et al., 2017, 2022; Zigiotta et al., 2021). Given that these patients are often affected by sensory, motor, and cognitive dysfunctions, MultiTab could contribute to providing multisensory stimulation and promoting manual actions from the patient, within multisensory motor training.

To conclude, the results of the current thesis shed light on the relationship between space, time, and motion through the auditory and visual modalities, together with filling the lack of technological solutions for the multifaceted evaluation of these dimensions. This thesis opened to exploring new methods for the assessment and training of multisensory abilities in typical and atypical populations.

# Scientific production

## Papers

- **Published**

Gori, M.\*, Bertonati, G.\*, Campus, C. and Amadeo, M. B. (2023). Multisensory representations of space and time in sensory cortices, *Human Brain Mapping*, 44(2), 656-667.

Gori M., Bertonati G., Mazzoni E., Freddi E. and Amadeo M.B. (2022). The impact of COVID-19 on the everyday life of blind and sighted individuals, *Frontiers in Psychology*, 13.

Bertonati G., Amadeo M.B., Campus C. and Gori M. (2021). Auditory speed processing in sighted and blind individuals, *PLoS ONE*, 16.

Bertonati G., Tonelli A., Cuturi F. C., Setti W. and Gori M. (2020). Assessment of spatial reasoning in blind individuals using a haptic version of the Kohs Block Design Test, *Current Research in Behavioral Sciences*, 1.

- **Accepted for publication**

Bertonati, G., Casado-Palacios, M., Crepaldi, M., Parmiggiani, A., Maviglia, A., Torazza, D., Campus, C. and Gori, M. MultiTab: A Novel Portable Device to Evaluate Multisensory Skills, *IEEE EMBC*.

- **Under review**

Bertonati, G., Amadeo, M.B., Campus C. and Gori, M. Task-specific cortical organization of multisensory spatial representation. Under review in *Human Brain Mapping*.

- **In preparation**

Bertonati, G., Gori, M. and Brenner E. How we adjust our movements in response to error feedback when intercepting targets moving at different speeds.

---

*\*Gori M. and Bertonati G. contributed equally to this study.*

### **Talks**

Bertonati, G., Casado-Palacios, M., Crepaldi, M., Parmiggiani, A., Maviglia, A., Torazza, D., Campus, C., Gori, M. (2023). MultiTab: A Novel Portable Device to Evaluate Multisensory Skills. 45th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Sydney, oral presentation.

Bertonati G., Amadeo M.B., Campus C., Gori M. (2022). Organizzazione dominio-specifica dei processi multisensoriali a livello corticale. XXX Congresso Nazionale Associazione Italiana di Psicologia, Padova, oral presentation.

Gori M., Bertonati G., Amadeo M.B, Pavani F., Campus C. (2022). Hearing the visual rhythm and seeing the audio space: early cortical sensory responses in typical but not in blind and deaf individuals. International Multisensory Research Forum, Ulm.

Bertonati G., Amadeo M.B., Campus C., Gori M. (2021). Multisensory cortical representations of space and time. European Conference on Visual Perception ECVF 2021, online oral presentation.

Bertonati G., Tonelli A., Cuturi L., Setti W., Gori M. (2021). Valutazione del ragionamento spaziale in persone con e senza disabilità visiva: una versione tattile del Test dei cubi di Kohs. XXVII Congresso Nazionale Associazione Italiana di Psicologia della Sezione di Psicologia Sperimentale, Lecce, oral presentation.

### **Posters**

Gori M., Bertonati G., Amadeo M.B., Campus C. (2023). Domain-specific cortical organization of multisensory processing. International Multisensory Research Forum, Bruxelles.

Bertonati G., Amadeo M.B., Campus C., Gori M. (2022). ERP components of multisensory spatial representation in the visual cortex. FENS Forum 2022, Paris.

Bertonati G., Gori M., Brenner E. (2022). How we adjust our movements in response to error feedback when intercepting moving targets. International Multisensory Research Forum, Ulm.

Bertonati G., Amadeo M.B., Campus C., Gori M. (2021). Task-specific cortical organization of multisensory spatial representation. 50th annual meeting Society for Neuroscience, online.

Bertonati G., Amadeo M.B., Campus C., Gori M. (2021). Auditory speed processing in sighted and blind individuals. Vision Sciences Society, online.

# References

- Aggius-Vella, E., Kolarik, A. J., Gori, M., Cirstea, S., Campus, C., Moore, B. C., and Pardhan, S. (2020). Comparison of auditory spatial bisection and minimum audible angle in front, lateral, and back space. *Scientific Reports*, 10(1):1–9.
- Ahveninen, J., Ingalls, G., Yildirim, F., Calabro, F. J., and Vaina, L. M. (2019). Peripheral visual localization is degraded by globally incongruent auditory - spatial attention cues. *Experimental Brain Research*, 237(9):2137–2143.
- Alais, D. and Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, 14(3):257–262.
- Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of neurophysiology*, 52(6):1106–1130.
- Albright, T. D. and Stoner, G. R. (1995). Visual motion perception. *Proceedings of the National Academy of Sciences of the United States of America*, 92(7):2433–2440.
- Amadeo, M. B., Campus, C., and Gori, M. (2019a). Impact of years of blindness on neural circuits underlying auditory spatial representation. *NeuroImage*, 191(January):140–149.
- Amadeo, M. B., Campus, C., and Gori, M. (2020a). Visual representations of time elicit early responses in human temporal cortex. *NeuroImage*, 217(January):116912.
- Amadeo, M. B., Campus, C., and Gori, M. (2020b). Years of Blindness Lead to “Visualize” Space Through Time. *Frontiers in Neuroscience*, 14(August):1–14.
- Amadeo, M. B., Campus, C., Pavani, F., and Gori, M. (2019b). Spatial Cues Influence Time Estimations in Deaf Individuals. *iScience*, 19:369–377.
- Amadeo, M. B., Tonelli, A., Campus, C., and Gori, M. (2022). Reduced flash lag illusion in early deaf individuals. *Brain Research*, 1776:147744.
- Amedi, A., Hofstetter, S., Maidenbaum, S., and Heimler, B. (2017). Task Selectivity as a Comprehensive Principle for Brain Organization. *Trends in Cognitive Sciences*, 21(5):307–310.
- Amedi, A., Stern, W. M., Camprodon, J. A., Bermpohl, F., Merabet, L., Rotman, S., Hemond, C., Meijer, P., and Pascual-Leone, A. (2007). Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nature Neuroscience*, 10(6):687–689.

- Andersen, R. A., Snyder, L. H., Bradley, D. C., and Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience*, 20:303–330.
- Anderson, P. W. and Zahorik, P. (2011). Auditory and visual distance estimation. *Proceedings of Meetings on Acoustics*, 12(2011).
- Auer, E. T. J., Bernstein, L. E., Sungkarat, W., and Singh, M. (2007). Vibrotactile activation of the auditory cortices in deaf versus hearing adults. *NeuroReport*, 18(7).
- Baddeley, R., Ingram, H., and Miall, R. (2003). System identification applied to a visuomotor task: near-optimal human performance in a noisy changing task. *Journal of Neuroscience*, 23(7):3066–3075.
- Baillargeon, R. and DeVos, J. (1991). Object permanence in young infants: Further evidence. *Child development*, 62(6):1227–1246.
- Baillargeon, R., Spelke, E. S., and Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, 20(3):191–208.
- Barakat, B., Seitz, A. R., and Shams, L. (2015). Visual rhythm perception improves through auditory but not visual training. *Current Biology*, 25(2):R60–R61.
- Battal, C., Gurtubay-Antolin, A., Rezk, M., Mattioni, S., Bertonati, G., Occelli, V., Bottini, R., Targher, S., Maffei, C., Jovicich, J., and Collignon, O. (2022). Structural and Functional Network-Level Reorganization in the Coding of Auditory Motion Directions and Sound Source Locations in the Absence of Vision. *Journal of Neuroscience*, 42(23):4652–4668.
- Battal, C., Occelli, V., Bertonati, G., Falagiarda, F., and Collignon, O. (2020). General Enhancement of Spatial Hearing in Congenitally Blind People.
- Baumgart, F., Gaschler-Markefski, B., Woldorff, M. G., Heinze, H.-J., and Scheich, H. (1999). A movement-sensitive area in auditory cortex. *Nature*, 400(6746):724–726.
- Bavelier, D., Dye, M. W. G., and Hauser, P. C. (2006). Do deaf individuals see better? *Trends in Cognitive Sciences*, 10(11):512–518.
- Beauchamp, M. S. (2005). See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. *Current opinion in neurobiology*, 15(2):145–153.
- Bedny, M., Konkle, T., Pelphrey, K., Saxe, R., and Pascual-Leone, A. (2010). Sensitive period for a multimodal response in human visual motion area MT/MST. *Current Biology*, 20(21):1900–1906.
- Ben-Yishai, R., Bar-Or, R. L., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences*, 92(9):3844–3848.
- Bentvelzen, A., Leung, J., and Alais, D. (2009). Discriminating audiovisual speed: Optimal integration of speed defaults to probability summation when component reliabilities diverge. *Perception*, 38(7):966–987.



- Bertelson, P. and Aschersleben, G. (2003). Temporal ventriloquism: Crossmodal interaction on the time dimension: 1. Evidence from auditory-visual temporal order judgment. *International Journal of Psychophysiology*, 50(1-2):147–155.
- Bertonati, G., Amadeo, M. B., Campus, C., and Gori, M. (2021). Auditory speed processing in sighted and blind individuals. *PLoS ONE*, 16(9 September):1–14.
- Bertonati, G., Tonelli, A., Cuturi, L. F., Setti, W., and Gori, M. (2020). Assessment of spatial reasoning in blind individuals using a haptic version of the Kohs Block Design Test. *Current Research in Behavioral Sciences*, 1(October):100004.
- Best, V., Van Schaik, A., Jin, C., and Carlile, S. (2005). Auditory spatial perception with sources overlapping in frequency and time. *Acta Acustica united with Acustica*, 91(3):421–428.
- Bill, J. C. and Teft, L. W. (1969). Space-time relations: effects of time on perceived visual extent. *Journal of Experimental Psychology*, 81(1):196.
- Binetti, N., Hagura, N., Fadipe, C., Tomassini, A., Walsh, V., and Bestmann, S. (2015). Binding space and time through action. *Proceedings of the Royal Society B: Biological Sciences*, 282(1805):22–27.
- Bjoertomt, O., Cowey, A., and Walsh, V. (2002). Spatial neglect in near and far space investigated by repetitive transcranial magnetic stimulation. *Brain : a journal of neurology*, 125(Pt 9):2012–2022.
- Bolognini, N., Convento, S., Rossetti, A., and Merabet, L. B. (2013). Multisensory processing after a brain damage: Clues on post-injury crossmodal plasticity from neuropsychology. *Neuroscience and Biobehavioral Reviews*, 37(3):269–278.
- Bolognini, N., Leo, F., Passamonti, C., Stein, B. E., and Làdavas, E. (2007). Multisensory-mediated auditory localization. *Perception*, 36(10):1477–1485.
- Bolognini, N., Papagno, C., Moroni, D., and Maravita, A. (2010a). Tactile temporal processing in the auditory cortex. *Journal of Cognitive Neuroscience*, 22(6):1201–1211.
- Bolognini, N., Rasi, F., and Làdavas, E. (2005). Visual localization of sounds. *Neuropsychologia*, 43(11):1655–1661.
- Bolognini, N., Senna, I., Maravita, A., Pascual-Leone, A., and Merabet, L. B. (2010b). Auditory enhancement of visual phosphene perception: The effect of temporal and spatial factors and of stimulus intensity. *Neuroscience Letters*, 477(3):109–114.
- Bolognini, N. and Vallar, G. (2019). *Hemianopia, spatial neglect, and their multisensory rehabilitation*. Elsevier Inc.
- Bonino, D., Ricciardi, E., Bernardi, G., Sani, L., Gentili, C., Vecchi, T., and Pietrini, P. (2015). Spatial imagery relies on a sensory independent, though sensory sensitive, functional organization within the parietal cortex: A fMRI study of angle discrimination in sighted and congenitally blind individuals. *Neuropsychologia*, 68:59–70.

- Bootsma, R. J. and van Wieringen, P. C. (1990). Timing an attacking forehand drive in table tennis. *Journal of experimental psychology: Human perception and performance*, 16(1):21.
- Bottari, D., Heimler, B., Caclin, A., Dalmolin, A., Giard, M. H., and Pavani, F. (2014). Visual change detection recruits auditory cortices in early deafness. *NeuroImage*, 94:172–184.
- Boyer, E. O., Babayan, B. M., Bevilacqua, F., Noisternig, M., Warusfel, O., Roby-Brami, A., Hannequin, S., and Viaud-Delmon, I. (2013). From ear to hand: The role of the auditory-motor loop in pointing to an auditory source. *Frontiers in Computational Neuroscience*, 7(MAR):1–9.
- Brackbill, Y. and Fitzgerald, H. E. (1972). Stereotype temporal conditioning in infants. *Psychophysiology*, 9(6):569–577.
- Braddick, O., Atkinson, J., and Wattam-Bell, J. (2003). Normal and anomalous development of visual motion processing: Motion coherence and 'dorsal-stream vulnerability'. *Neuropsychologia*, 41(13):1769–1784.
- Brandwein, A. B., Foxe, J. J., Russo, N. N., Altschuler, T. S., Gomes, H., and Molholm, S. (2011). The development of audiovisual multisensory integration across childhood and early adolescence: A high-density electrical mapping study. *Cerebral Cortex*, 21(5):1042–1055.
- Brannon, E. M., Roussel, L. W., Meck, W. H., and Woldorff, M. (2004). Timing in the baby brain. *Cognitive Brain Research*, 21(2):227–233.
- Brannon, E. M., Suanda, S., and Libertus, K. (2007). Temporal discrimination increases in precision over development and parallels the development of numerosity discrimination. *Developmental science*, 10(6):770–777.
- Brenner, E., Cañal-Bruland, R., and Van Beers, R. J. (2013). How the required precision influences the way we intercept a moving object. *Experimental Brain Research*, 230(2):207–218.
- Brenner, E., Smeets, J., and Remijnse-Tamerius, H. (2002). Curvature in hand movements as a result of visual misjudgements of direction. *Spatial Vision*, 15(4):393–414.
- Brenner, E. and Smeets, J. B. (2007). Flexibility in intercepting moving objects. *Journal of Vision*, 7(5):1–17.
- Brenner, E. and Smeets, J. B. (2011). Continuous visual control of interception. *Human Movement Science*, 30(3):475–494.
- Brenner, E. and Smeets, J. B. (2015). How people achieve their amazing temporal precision in interception. *Journal of Vision*, 15(3):1–21.
- Brenner, E. and Smeets, J. B. (2018). Continuously updating one's predictions underlies successful interception. *Journal of Neurophysiology*, 120(6):3257–3274.
- Brenner, E., Spering, M., and Fiehler, K. (2019). Prediction in goal-directed action. *Journal of Vision*, 19:1–21.

- Brenner, E., van Dam, M., Berkhout, S., and Smeets, J. B. (2012). Timing the moment of impact in fast human movements. *Acta Psychologica*, 141(1):104–111.
- Bresciani, J.-P. and Ernst, M. O. (2007). Signal reliability modulates auditory–tactile integration for event counting. *Neuroreport*, 18(11):1157–1161.
- Brouwer, A. M., Brenner, E., and Smeets, J. B. (2000). Hitting moving objects. The dependency of hand velocity on the speed of the target. *Experimental Brain Research*, 133(2):242–248.
- Buetti, D. and Macaluso, E. (2010). Auditory temporal expectations modulate activity in visual cortex. *NeuroImage*, 51(3):1168–1183.
- Buetti, D. and Walsh, V. (2009). The parietal cortex and the representation of time , space , number and other magnitudes. pages 1831–1840.
- Burbaud, P., Degreze, P., Lafon, P., Franconi, J.-M., Bouligand, B., Bioulac, B., Caille, J., and Allard, M. (1995). Lateralization of prefrontal activation during internal mental calculation: a functional magnetic resonance imaging study. *Journal of Neurophysiology*, 74(5):2194–2200.
- Burge, J., Ernst, M. O., and Banks, M. S. (2008). The statistical determinants of adaptation rate in human reaching. *Journal of Vision*, 8(4):1–2.
- Burnett, L. R., Stein, B. E., Chaponis, D., and Wallace, M. T. (2004). Superior colliculus lesions preferentially disrupt multisensory orientation. *Neuroscience*, 124(3):535–547.
- Burr, D., Banks, M. S., and Morrone, M. C. (2009). Auditory dominance over vision in the perception of interval duration. *Experimental Brain Research*, 198(1):49–57.
- Burr, D. and Thompson, P. (2011). Motion psychophysics : 1985 – 2010. *Vision Research*, 51(13):1431–1456.
- Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H., and Woldorff, M. G. (2005). The spread of attention across modalities and space in a multisensory object. *Proceedings of the National Academy of Sciences of the United States of America*, 102(51):18751–18756.
- Caljouw, S. R., Van Der Kamp, J., and Savelsbergh, G. J. (2004). Timing of goal-directed hitting: Impact requirements change the information-movement coupling. *Experimental Brain Research*, 155(2):135–144.
- Camponogara, I. and Volcic, R. (2020). ScienceDirect Integration of haptics and vision in human multisensory grasping. *CORTEX*, 135:173–185.
- Campus, C., Sandini, G., Amadeo, M. B., and Gori, M. (2019). Stronger responses in the visual cortex of sighted compared to blind individuals during auditory space representation. *Scientific Reports*, 9(1):1–12.
- Campus, C., Sandini, G., Concetta Morrone, M., and Gori, M. (2017). Spatial localization of sound elicits early responses from occipital visual cortex in humans. *Scientific Reports*, 7(1):1–12.

- Cappagli, G., Finocchietti, S., Baud-Bovy, G., Cocchi, E., and Gori, M. (2017). Multisensory rehabilitation training improves spatial perception in totally but not partially visually deprived children. *Frontiers in Integrative Neuroscience*, 11(October):1–11.
- Cappe, C. and Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *European Journal of Neuroscience*, 22(11):2886–2902.
- Cappe, C., Thut, G., Romei, V., and Murray, M. M. (2010). Auditory-visual multisensory interactions in humans: Timing, topography, directionality, and sources. *Journal of Neuroscience*, 30(38):12572–12580.
- Carlile, S. and Best, V. (2002). Discrimination of sound source velocity in human listeners. *The Journal of the Acoustical Society of America*, 111(2):1026–1035.
- Carlile, S. and Leung, J. (2016). The Perception of Auditory Motion. *Trends in Hearing*, 20:1–19.
- Carnahan, H. and McFadyen, B. J. (1996). Visuomotor control when reaching toward and grasping moving targets. *Acta Psychologica*, 92(1):17–32.
- Cecchetti, L., Kupers, R., Ptito, M., Pietrini, P., and Ricciardi, E. (2016). Are Supramodality and Cross-Modal Plasticity the Yin and Yang of Brain Development ? From Blindness to Rehabilitation. *Frontiers in Systems Neuroscience*, 10(November):1–8.
- Champion, R. A. and Freeman, T. C. A. (2010). Discrimination contours for the perception of head-centred velocity. *Journal of vision*, 10(6).
- Chaplin, T. A., Rosa, M. G. P., Lui, L. L., and Stuart, G. (2018). Auditory and Visual Motion Processing and Integration in the Primate Cerebral Cortex. *Frontiers in neural circuits*, 12(October):1–9.
- Chaumon, M., Bishop, D. V., and Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of Neuroscience Methods*, 250(March):47–63.
- Chen, L. and Vroomen, J. (2013). Intersensory binding across space and time: A tutorial review. *Attention, Perception, and Psychophysics*, 75(5):790–811.
- Cheng, S. and Sabes, P. N. (2007). Calibration of visually guided reaching is driven by error-corrective learning and internal dynamics. *Journal of neurophysiology*, 97(4):3057–3069.
- Cicchini, G. M., Anobile, G., and Burr, D. C. (2014). Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proceedings of the National Academy of Sciences of the United States of America*, 111(21):7867–7872.
- Collignon, O., Vandewalle, G., Voss, P., Albouy, G., Charbonneau, G., Lassonde, M., and Lepore, F. (2011). Functional specialization for auditory-spatial processing in the occipital cortex of congenitally blind humans. *Proceedings of the National Academy of Sciences of the United States of America*, 108(11):4435–4440.

- Collignon, O., Voss, P., Lassonde, M., and Lepore, F. (2009). Cross-modal plasticity for the spatial processing of sounds in visually deprived subjects. *Experimental Brain Research*, 192(3):343–358.
- Collyer, C. E. (1977). Discrimination of spatial and temporal intervals defined by three light flashes: Effects of spacing on temporal judgments and of timing on spatial judgments. *Perception & Psychophysics*, 21:357–364.
- Colonus, H. and Diederich, A. (2004). Multisensory interaction in saccadic reaction time: a time-window-of-integration model. *Journal of cognitive neuroscience*, 16(6):1000–1009.
- Cont, C. and Zimmermann, E. (2021). The Motor Representation of Sensory Experience. *Current Biology*, 31(5):1029–1036.e2.
- Coull, J. T., Vidal, F., Nazarian, B., and Macar, F. (2004). Functional anatomy of the attentional modulation of time estimation. *Science*, 303(5663):1506–1508.
- Crane, B. T. (2012). Direction specific biases in human visual and vestibular heading perception. *PloS one*, 7(12):e51383.
- Culham, J., He, S., Dukelow, S., and Verstraten, F. A. (2001). Visual motion and the human brain: What has neuroimaging told us? *Acta Psychologica*, 107(1-3):69–94.
- Cuturi, L. F. (2022). Perceptual biases as the side effect of a multisensory adaptive system: Insights from verticality and self-motion perception. *Vision*, 6(3):53.
- Cuturi, L. F., Aggius-Vella, E., Campus, C., Parmiggiani, A., and Gori, M. (2016). From science to technology: Orientation and mobility in blind children and adults. *Neuroscience and Biobehavioral Reviews*, 71:240–251.
- Cuturi, L. F. and MacNeilage, P. R. (2013). Systematic biases in human heading estimation. *PloS one*, 8(2):e56862.
- Danckert, J., Ferber, S., Pun, C., Broderick, C., Striemer, C., Rock, S., and Stewart, D. (2007). Neglected time: impaired temporal perception of multisecond intervals in unilateral neglect. *Journal of cognitive neuroscience*, 19(10):1706–1720.
- de la Malla, C., López-Moliner, J., and Brenner, E. (2014). Dealing with delays does not transfer across sensorimotor tasks. *Journal of Vision*, 14(12):8.
- De La Malla, C., Smeets, J. B., and Brenner, E. (2017). Potential Systematic Interception Errors are Avoided When Tracking the Target with One's Eyes. *Scientific Reports*, 7(1):1–12.
- De Lussanet, M. H., Smeets, J. B., and Brenner, E. (2001). The effect of expectations on hitting moving targets: Influence of the preceding target's speed. *Experimental Brain Research*, 137(2):246–248.
- De Meo, R., Murray, M. M., Clarke, S., and Matusz, P. J. (2015). Top-down control and early multisensory processes: Chicken vs. egg. *Frontiers in Integrative Neuroscience*, 9(MAR):1–6.

- Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., and Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, 284(5416):970–974.
- Delorme, A. and Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21.
- DeLucia, P. R. (2004). Multiple sources of information influence time-to-contact judgments: Do heuristics accommodate limits in sensory and cognitive processes? In *Advances in psychology*, volume 135, pages 243–285. Elsevier.
- Di Russo, F., Martínez, A., Sereno, M. I., Pitzalis, S., and Hillyard, S. A. (2002). Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping*, 15(2):95–111.
- Dickinson, A. and Balleine, B. W. (2000). Causal cognition and goal-directed action. In *The evolution of cognition.*, Vienna series in theoretical biology., pages 185–204. The MIT Press, Cambridge, MA, US.
- Diederich, A. and Colonius, H. (2004). Bimodal and trimodal multisensory enhancement : Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics*, 66(8):1388–1404.
- Dormal, G., Rezk, M., Yakobov, E., Lepore, F., and Collignon, O. (2016). Auditory motion in the sighted and blind: Early visual deprivation triggers a large-scale imbalance between auditory and "visual" brain regions. *NeuroImage*, 134:630–644.
- Droit-Volet, S. (2013). Time perception in children: A neurodevelopmental approach. *Neuropsychologia*, 51(2):220–234.
- Elshout, J. A., Van der Stoep, N., Nijboer, T. C., and Van der Stigchel, S. (2020). Motor congruency and multisensory integration jointly facilitate visual information processing before movement execution. *Experimental Brain Research*, 238(3):667–673.
- Elsman, E. B., Al Baaj, M., van Rens, G. H., Sijbrandi, W., van den Broek, E. G., van der Aa, H. P., Schakel, W., Heymans, M. W., de Vries, R., Vervloed, M. P., Steenbergen, B., and van Nispen, R. M. (2019). Interventions to improve functioning, participation, and quality of life in children with visual impairment: a systematic review. *Survey of Ophthalmology*, 64(4):512–557.
- Ernst, M. O. and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429–433.
- Falchier, A., Clavagnier, S., Barone, P., and Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *Journal of Neuroscience*, 22(13):5749–5759.
- Feng, W., Störmer, V. S., Martinez, A., McDonald, J. J., and Hillyard, S. A. (2014). Sounds activate visual cortex and improve visual discrimination. *Journal of Neuroscience*, 34(29):9817–9824.

- Ferrandez, A.-M., Hugueville, L., Lehericy, S., Poline, J.-B., Marsault, C., and Pouthas, V. (2003). Basal ganglia and supplementary motor area subsecond duration perception: an fmri study. *Neuroimage*, 19(4):1532–1544.
- Fieger, A., Röder, B., Teder-Sälejärvi, W., Hillyard, S. A., and Neville, H. J. (2006). Auditory spatial tuning in late-onset blindness in humans. *Journal of Cognitive Neuroscience*, 18(2):149–157.
- Finocchietti, S., Cappagli, G., Porquis, L. B., Baud-Bovy, G., Cocchi, E., and Gori, M. (2015). Evaluation of the Audio Bracelet for Blind Interaction for improving mobility and spatial cognition in early blind children - A pilot study. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2015-Novem:7998–8001.
- Fischer, J. and Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, 17(5):738–743.
- Flanagan, J. R. and Beltzner, M. A. (2000). Independence of perceptual and sensorimotor predictions in the size–weight illusion. *Nature Neuroscience*, 3(7):737–741.
- Fonov, V. S., Evans, A. C., McKinstry, R. C., Almlí, C. R., and Collins, D. L. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, 47:S102.
- Fort, A., Delpuech, C., Pernier, J., and Giard, M. H. (2002). Early auditory–visual interactions in human cortex during nonredundant target identification. *Cognitive Brain Research*, 14(1):20–30.
- Fortis, P., Maravita, A., Gallucci, M., Ronchi, R., Grassi, E., Senna, I., Olgiati, E., Perucca, L., Banco, E., Posteraro, L., Tesio, L., and Vallar, G. (2010). Rehabilitating patients with left spatial neglect by prism exposure during a visuomotor activity. *Neuropsychology*, 24(6):681–697.
- Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C., and Schroeder, C. E. (2000). Multisensory auditory–somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Cognitive Brain Research*, 10(1-2):77–83.
- Fraisse, P. (1963). *The psychology of time*. Harper & Row, Oxford, England.
- Frassinetti, F., Bolognini, N., Bottari, D., Bonora, A., and Ladavas, E. (2005). Audiovisual integration in patients with visual deficit. *Journal of Cognitive Neuroscience*, 17(9):1442–1452.
- Freeman, T. C., Leung, J., Wufong, E., Orchard-Mills, E., Carlile, S., and Alais, D. (2014). Discrimination contours for moving sounds reveal duration and distance cues dominate auditory speed perception. *PLoS ONE*, 9(7):27–29.
- Freund, H.-J. (2001). The Parietal Lobe as a Sensorimotor Interface: A Perspective from Clinical and Neuroimaging Data. *NeuroImage*, 14(1):S142–S146.
- Frølich, L., Andersen, T. S., and Mørup, M. (2015). Classification of independent components of EEG into multiple artifact classes. *Psychophysiology*, 52(1):32–45.

- Frossard, J. and Renaud, O. (2019). *permuco: Permutation Tests for Regression, (Repeated Measures) ANOVA/ANCOVA and Comparison of Signals*. R package version 1.1.0.
- Fuster, J. M., Bodner, M., and Kroger, J. K. (2000). Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature*, 405(6784):347–351.
- Gal, O. (Retrieved in 2020). fit\_ellipse ([https://www.mathworks.com/matlabcentral/fileexchange/3215-fit\\_ellipse](https://www.mathworks.com/matlabcentral/fileexchange/3215-fit_ellipse)).
- Gavazzi, G., Bisio, A., and Pozzo, T. (2013). Time perception of visual motion is tuned by the motor representation of human actions. *Scientific Reports*, 3.
- Gebhard, J. W. and Mowbray, G. H. (1959). On discriminating the rate of visual flicker and auditory flutter. *The American journal of psychology*, 72(4):521–529.
- Gegenfurtner, K. R. and Hawken, M. J. (1995). Temporal and chromatic properties of motion mechanisms. *Vision Research*, 35(11):1547–1563.
- Ghazanfar, A. A. and Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, 10(6):278–285.
- Giard, M. H. and Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, 11(5):473–490.
- Gingras, G., Rowland, B. A., and Stein, B. E. (2009). The differing impact of multisensory and unisensory integration on behavior. *Journal of Neuroscience*, 29(15):4897–4902.
- Gondan, M., Niederhaus, B., Rösler, F., and Röder, B. (2005). Multisensory processing in the redundant-target effect: A behavioral and event-related potential study. *Perception and Psychophysics*, 67(4):713–726.
- Gondan, M. and Röder, B. (2006). A new method for detecting interactions between the senses in event-related potentials. *Brain Research*, 1073-1074(1):389–397.
- Goodale, M. A. and Milner, A. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1):20–25.
- Gori, M. (2015). Multisensory Integration and Calibration in Children and Adults with and without Sensory and Motor Disabilities. *Multisensory Research*, 28:71–99.
- Gori, M., Amadeo, M. B., and Campus, C. (2018). Temporal Cues Influence Space Estimations in Visually Impaired Individuals. *iScience*, 6:319–326.
- Gori, M., Amadeo, M. B., and Campus, C. (2020a). Spatial metric in blindness: behavioural and cortical processing. *Neuroscience and Biobehavioral Reviews*, 109:54–62.
- Gori, M., Amadeo, M. B., and Campus, C. (2020b). Temporal cues trick the visual and auditory cortices mimicking spatial cues in blind individuals. *Human Brain Mapping*, (January):2077–2091.



- Gori, M., Amadeo, M. B., Pavani, F., Valzolgher, C., and Campus, C. (2022a). Temporal visual representation elicits early auditory-like responses in hearing but not in deaf individuals. *Scientific Reports*, 12(1):1–13.
- Gori, M., Bertoni, G., Campus, C., and Amadeo, M. B. (2022b). Multisensory representations of space and time in sensory cortices. *Human Brain Mapping*, (September):1–12.
- Gori, M., Campus, C., Signorini, S., Rivara, E., and Bremner, A. J. (2021). Multisensory spatial perception in visually impaired infants. *Current Biology*, 31(22):5093–5101.e5.
- Gori, M., Cappagli, G., Tonelli, A., Baud-Bovy, G., and Finocchietti, S. (2016). Devices for visually impaired people: High technological devices with low user acceptance and no adaptability for children. *Neuroscience and Biobehavioral Reviews*, 69:79–88.
- Gori, M., Chilosi, A., Forli, F., and Burr, D. (2017). Audio-visual temporal perception in children with restored hearing. *Neuropsychologia*, 99:350–359.
- Gori, M., Del Viva, M., Sandini, G., and Burr, D. C. (2008). Young Children Do Not Integrate Visual and Haptic Form Information. *Current Biology*, 18(9):694–698.
- Gori, M., Sandini, G., and Burr, D. (2012). Development of visuo-auditory integration in space and time. *Frontiers in Integrative Neuroscience*, 6(SEPTEMBER):1–8.
- Gori, M., Sandini, G., Martinoli, C., and Burr, D. C. (2014). Impairment of auditory spatial localization in congenitally blind human subjects. *Brain*, 137(1):288–293.
- Gottlieb, G. (1971). *Development of species identification in birds: An inquiry into the prenatal determinants of perception*. U. Chicago Press, Oxford, England.
- Gougoux, F., Zatorre, R. J., Lassonde, M., Voss, P., and Lepore, F. (2005). A functional neuroimaging study of sound localization: Visual cortex activity predicts performance in early-blind individuals. *PLoS Biology*, 3(2):0324–0333.
- Gramfort, A., Strohmeier, D., Haueisen, J., Hamalainen, M., and Kowalski, M. (2011). Functional Brain Imaging with M/EEG Using Structured Sparsity in Time-Frequency Dictionaries. In Székely, G. and Hahn, H. K., editors, *Information Processing in Medical Imaging*, pages 600–611, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Gray, R., Regan, D., Castaneda, B., and Sieffert, R. (2006). Role of feedback in the accuracy of perceived direction of motion-in-depth and control of interceptive action. *Vision Research*, 46(10):1676–1694.
- Guttman, S. E., Gilroy, L. A., and Blake, R. (2005). Hearing what the eyes see: Auditory encoding of visual temporal sequences. *Psychological science*, 16(3):228–235.
- Hagen, M. C., Franzén, O., McGlone, F., Essick, G., Dancer, C., and Pardo, J. V. (2002). Tactile motion activates the human middle temporal/V5 (MT/V5) complex. *European Journal of Neuroscience*, 16(5):957–964.
- Hagura, N., Kanai, R., Orgs, G., and Haggard, P. (2012). Ready steady slow: Action preparation slows the subjective passage of time. *Proceedings of the Royal Society B: Biological Sciences*, 279(1746):4399–4406.

- Hairston, W. D., Laurienti, P. J., Mishra, G., Burdette, J. H., and Wallace, M. T. (2003). Multisensory enhancement of localization under conditions of induced myopia. *Experimental Brain Research*, 152(3):404–408.
- Hecht, H. and Savelsbergh, G. J. (2004). Theories of time-to-contact judgment. In *Advances in psychology*, volume 135, pages 1–11. Elsevier.
- Heimler, B. and Amedi, A. (2020). Are critical periods reversible in the adult brain? Insights on cortical specializations based on sensory deprivation studies. *Neuroscience and Biobehavioral Reviews*, 116(June 2019):494–507.
- Heimler, B., Striem-Amit, E., and Amedi, A. (2015). Origins of task-specific sensory-independent organization in the visual and auditory brain: Neuroscience evidence, open questions and clinical implications. *Current Opinion in Neurobiology*, 35:169–177.
- Hervé, M. (2022). *RVAideMemoire: Testing and Plotting Procedures for Biostatistics*. R package version 0.9-81-2.
- Herzfeld, D. J., Vaswani, P. A., Marko, M. K., and Shadmehr, R. (2014). A memory of errors in sensorimotor learning. *Science*, 345(6202):1349–1353.
- Hildebrandt, A., Griebßbach, E., and Cañal-Bruland, R. (2022). Auditory perception dominates in motor rhythm reproduction. *Perception*, 51(6):403–416.
- Hillyard, S. A. and Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences of the United States of America*, 95(3):781–787.
- Hötting, K. and Röder, B. (2009). Auditory and auditory-tactile processing in congenitally blind humans. *Hearing Research*, 258(1-2):165–174.
- Hyde, K. L., Peretz, I., and Zatorre, R. J. (2008). Evidence for the role of the right auditory cortex in fine pitch resolution. *Neuropsychologia*, 46(2):632–639.
- Imaizumi, S. and Asai, T. (2015). Dissociation of agency and body ownership following visuomotor temporal recalibration. *Frontiers in integrative neuroscience*, 9:35.
- Inc., T. M. (2019). Matlab version: 9.7.0.1296695 (r2019b).
- Israël, I., Sievering, D., and Koenig, E. (1995). Self-rotation estimate about the vertical axis. *Acta oto-laryngologica*, 115(1):3–8.
- Jiang, F., Stecker, G. C., Boynton, G. M., and Fine, I. (2016). Early blindness results in developmental plasticity for auditory motion processing within auditory and occipital cortex. *Frontiers in Human Neuroscience*, 10(JULY2016).
- Jiang, F., Stecker, G. C., and Fine, I. (2014). Auditory motion processing after early blindness. *Journal of Vision*, 14(13):4.
- Jiang, W., Jiang, H., and Stein, B. E. (2002). Two corticotectal areas facilitate multisensory orientation behavior. *Journal of Cognitive Neuroscience*, 14(8):1240–1255.

- Jiang, W. and Stein, B. E. (2003). Cortex Controls Multisensory Depression in Superior Colliculus. *Journal of Neurophysiology*, 90(4):2123–2135.
- Johansson, B. B. (2012). Multisensory stimulation in stroke rehabilitation. *Frontiers in Human Neuroscience*, 6(April):1–11.
- Johnen, A., Wagner, H., and Gaese, B. H. (2022). Spatial Attention Modulates Sound Localization in Barn Owls. pages 1009–1012.
- Jones, B. and Huang, Y. L. (1982). Space-time dependencies in psychophysical judgment of extent and duration: Algebraic models of the tau and kappa effects. *Psychological bulletin*, 91(1):128.
- Jörg, S., Normoyle, A., and Safonova, A. (2012). How responsiveness affects players' perception in digital games. In *Proceedings of the ACM symposium on applied perception*, pages 33–38.
- Kanai, R., Lloyd, H., Buetti, D., and Walsh, V. (2011). Modality-independent role of the primary auditory cortex in time estimation. *Experimental Brain Research*, 209:465–471.
- Kansaku, K., Carver, B., Johnson, A., Matsuda, K., Sadato, N., and Hallett, M. (2007). The role of the human ventral premotor cortex in counting successive stimuli. *Experimental brain research*, 178:339–350.
- Karns, C. M., Dow, M. W., and Neville, H. J. (2012). Altered cross-modal processing in the primary auditory cortex of congenitally deaf adults: A visual-somatosensory fMRI study with a double-flash illusion. *Journal of Neuroscience*, 32(28):9626–9638.
- Kaufmann, L., Vogel, S. E., Wood, G., Kremser, C., Schocke, M., Zimmerhackl, L.-B., and Koten, J. W. (2008). A developmental fMRI study of nonsymbolic numerical and spatial processing. *Cortex*, 44(4):376–385.
- King, A. J. (2009). Visual influences on auditory spatial learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1515):331–339.
- King, A. J. (2014). What happens to your hearing if you are born blind? *Brain*, 137(1):6–8.
- Kirkham, N. Z., Slemmer, J. A., and Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition*, 83(2):B35–B42.
- Kitagawa, N. and Ichihara, S. (2002). Hearing visual motion in depth. *Nature*, 416(6877):172–174.
- Kitajima, N. and Yamashita, Y. (1999). Dynamic capture of sound motion by light stimuli moving in three-dimensional space. *Perceptual and Motor Skills*, 89(3 PART 1):1139–1158.
- Kujala, T., Alho, K., Kekoni, J., Hämäläinen, H., Reinikainen, K., Salonen, O., Standertskjöld-Nordenstam, C. G., and Näätänen, R. (1995). Auditory and somatosensory event-related brain potentials in early blind humans. *Experimental Brain Research*, 104(3):519–526.
- Kupers, R. and Ptito, M. (2014). Compensatory plasticity and cross-modal reorganization following early visual deprivation. *Neuroscience and Biobehavioral Reviews*, 41:36–52.

- Kuśmierk, P. and Rauschecker, J. P. (2014). Selectivity for space and time in early areas of the auditory dorsal stream in the rhesus monkey. *Journal of neurophysiology*, 111(8):1671–1685.
- Lambrechts, A., Walsh, V., and Wassenhove, V. V. (2013). Evidence Accumulation in the Magnitude System. 8(12).
- Land, M., Mennie, N., and Rusted, J. (1999). The Roles of Vision and Eye Movements in the Control of Activities of Daily Living. *Perception*, 28(11):1311–1328.
- Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research*, 25(3):296–324.
- Land, M. F. (2009). Vision, eye movements, and natural behavior. *Visual Neuroscience*, 26(1):51–62.
- Land, M. F. and McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball. *Nature Neuroscience*, 3(12):1340–1345.
- Lessard, N., Paré, M., Lepore, F., and Lassonde, M. (1998). Early-blind human subjects localize sound sources better than sighted subjects. *Nature*, 395(6699):278–280.
- Levänen, S., Jousmäki, V., and Hari, R. (1998). Vibration-induced auditory-cortex activation in a congenitally deaf adult. *Current Biology*, 8(15):869–872.
- Lewald, J. (2013). Exceptional ability of blind humans to hear sound motion: Implications for the emergence of auditory space. *Neuropsychologia*, 51(1):181–186.
- Lewald, J., Meister, I. G., Weidemann, J., and Töpper, R. (2004). Involvement of the superior temporal cortex and the occipital cortex in spatial hearing: Evidence from repetitive transcranial magnetic stimulation. *Journal of Cognitive Neuroscience*, 16(5):828–838.
- Lewis, P. A. and Miall, R. C. (2003). Brain activation patterns during measurement of sub-and supra-second intervals. *Neuropsychologia*, 41(12):1583–1592.
- Lewkowicz, D. J. (1986). Developmental changes in infants' bisensory response to synchronous durations. *Infant Behavior and Development*, 9(3):335–353.
- Lewkowicz, D. J. (2000). Infants' perception of the audible, visible, and bimodal attributes of multimodal syllables. *Child Development*, 71(5):1241–1257.
- Lewkowicz, D. J. and Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. *Proceedings of the National Academy of Sciences of the United States of America*, 103(17):6771–6774.
- Lewkowicz, D. J., Leo, I., and Simion, F. (2010). Intersensory perception at birth: Newborns match nonhuman primate faces and voices. *Infancy*, 15(1):46–60.
- Lindemann, O., Abolafia, J. M., Girardi, G., and Bekkering, H. (2007). Getting a grip on numbers: numerical magnitude priming in object grasping. *Journal of Experimental Psychology: Human Perception and Performance*, 33(6):1400.

- Liu, D. and Todorov, E. (2007). Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. *Journal of Neuroscience*, 27(35):9354–9368.
- Locke, S. M., Leung, J., and Carlile, S. (2016). Sensitivity to auditory velocity contrast. *Scientific Reports*, 6:1–5.
- Longo, M. R. and Haggard, P. (2009). Sense of agency primes manual motor responses. *Perception*, 38(1):69–78.
- López-Moliner, J., Vullings, C., Madelain, L., and van Beers, R. J. (2019). Prediction and final temporal errors are used for trial-to-trial motor corrections. *Scientific Reports*, 9(1).
- Lourenco, S. F. and Huttenlocher, J. (2008). The representation of geometric cues in infancy. *Infancy*, 13(2):103–127.
- Lucan, J. N., Foxe, J. J., Gomez-Ramirez, M., Sathian, K., and Molholm, S. (2010). Tactile shape discrimination recruits human lateral occipital complex during early perceptual processing. *Human Brain Mapping*, 31(11):1813–1821.
- Mackrout, I. and Simoneau, M. (2011). Visuo-vestibular interaction: predicting the position of a visual target during passive body rotation. *Neuroscience*, 195:45–53.
- Manabe, K. and Riquimaroux, H. (2000). Sound controls velocity perception of visual apparent motion. *Acoustical Science and Technology*, 21(3):171–174.
- Martolini, C., Cappagli, G., Campus, C., and Gori, M. (2020). Shape Recognition With Sounds : Improvement in Sighted Individuals After Audio – Motor Training. 33:417–431.
- Mateeff, S., Hohnsbein, J., and Noack, T. (1985). Dynamic visual capture: apparent auditory motion induced by a moving visual target. *Perception*, 14(6):721–727.
- Matusz, P. J., Retsa, C., and Murray, M. M. (2016). The context-contingent nature of cross-modal activations of the visual cortex. *NeuroImage*, 125:996–1004.
- McDonald, J. J., Störmer, V. S., Martinez, A., Feng, W., and Hillyard, S. A. (2013). Salient sounds activate human visual cortex automatically. *Journal of Neuroscience*, 33(21):9194–9201.
- McLeod, P., McLaughlin, C., Nimmo-Smith, I., et al. (1985). Information encapsulation and automaticity: Evidence from the visual control of finely timed actions. *Attention and performance XI*, pages 391–406.
- Meredith, M. A., Nemitz, J. W., and Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience*, 7(10):3215–3229.
- Meyer, G. F. and Wuerger, S. M. (2001). Cross-modal integration of auditory and visual motion signals. *NeuroReport*, 12(11):2557–2560.
- Middlebrooks, J. C. (2015). Sound localization. *Handbook of Clinical Neurology*, 129(February):99–116.

- Middlebrooks, J. C. and Green, D. M. (1991). Sound Localization by Human Listeners. *Annual Review of Psychology*, 42:135–159.
- Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, 14(2):247–279.
- Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., and Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Cognitive Brain Research*, 14(1):115–128.
- Morelli, F., Aprile, G., Cappagli, G., Luparia, A., Decortes, F., Gori, M., Signorini, S., Heimler, B., and Dutton, G. N. (2020). A Multidimensional , Multisensory and Comprehensive Rehabilitation Intervention to Improve Spatial Functioning in the Visually Impaired Child : A Community Case Study. 14(July):1–13.
- Moretto, G. and Di Pellegrino, G. (2008). Grasping numbers. *Experimental brain research*, 188:505–515.
- Morrone, M. C., Ross, J., and Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. *Nature Neuroscience*, 8(7):950–954.
- Murray, M. M., Lewkowicz, D. J., Amedi, A., and Wallace, M. T. (2016a). Multisensory Processes: A Balancing Act across the Lifespan. *Trends in Neurosciences*, 39(8):567–579.
- Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., Schroeder, C. E., and Foxe, J. J. (2005). Grabbing your ear: Rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cerebral Cortex*, 15(7):963–974.
- Murray, M. M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., and Matusz, P. J. (2016b). The multisensory function of the human primary visual cortex. *Neuropsychologia*, 83:161–169.
- Murray, M. M. and Wallace, M. T. (2011). *The neural bases of multisensory processes*. CRC Press.
- Näätänen, R. and Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, 24(4):375–425.
- Nardini, M., Jones, P., Bedford, R., and Braddick, O. (2008). Development of Cue Integration in Human Navigation. *Current Biology*, 18(9):689–693.
- Nazzi, T. and Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, 41(1):233–243.
- Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., and Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. *Developmental Science*, 9(5):454–464.
- Newcombe, N. S., Sluzenski, J., and Huttenlocher, J. (2005). Preexisting knowledge versus on-line learning: What do young infants really know about spatial location? *Psychological Science*, 16(3):222–227.

- Nijhawan, R. and Khurana, B. (2010). *Space and time in perception and action*. Cambridge University Press.
- Nishida, S., Kawabe, T., Sawayama, M., and Fukiage, T. (2018). Motion perception: From detection to interpretation. *Annual Review of Vision Science*, 4(July):501–523.
- Noorlander, C., Heuts, M. J. G., and Koenderink, J. J. (1980). Influence of the target size on the detection threshold for luminance and chromaticity contrast. *JOSA*, 70(9):1116–1121.
- Obleser, J., Zimmermann, J., Van Meter, J., and Rauschecker, J. P. (2007). Multiple stages of auditory speech perception reflected in event-related fmri. *Cerebral Cortex*, 17(10):2251–2257.
- Ohmura, H. (1987). Intersensory influences on the perception of apparent movement. *Japanese psychological research*, 29(1):1–9.
- Oliveri, M. and Koch, G. (2009). Spatial – temporal interactions in the human brain. pages 489–497.
- Oliveri, M., Vicario, C. M., Salerno, S., Koch, G., Turriziani, P., Mangano, R., Chillemi, G., and Caltagirone, C. (2008). Perceiving numbers alters time perception. *Neuroscience letters*, 438(3):308–311.
- Oostwoud Wijdenes, L., Brenner, E., and Smeets, J. B. J. (2011). Fast and fine-tuned corrections when the target of a hand movement is displaced. *Experimental Brain Research*, 214:453–462.
- Parise, C. V., Knorre, K., and Ernst, M. O. (2014). Natural auditory scene statistics shapes human spatial hearing. *PNAS*, 111(16).
- Parisi, A., Bellinzona, F., Di Lernia, D., Repetto, C., De Gaspari, S., Brizzi, G., Riva, G., and Tuena, C. (2022). Efficacy of Multisensory Technology in Post-Stroke Cognitive Rehabilitation: A Systematic Review. *Journal of Clinical Medicine*, 11(21):6324.
- Pavani, F., Macaluso, E., Warren, J. D., Driver, J., and Griffiths, T. D. (2002). A common cortical substrate activated by horizontal and vertical sound movement in the human brain. *Current Biology*, 12(18):1584–1590.
- Pei, Y.-C. and Bensmaia, S. J. (2014). The neural basis of tactile motion perception. *Journal of Neurophysiology*, 112(12):3023–3032.
- Pick, H. L., Warren, D. H., and Hay, J. C. (1969). Sensory conflict in judgments of spatial direction. *Perception & Psychophysics*, 6(4):203–205.
- Pietrini, P., Furey, M. L., Ricciardi, E., Gobbini, M. I., Wu, W. H., Cohen, L., Guazzelli, M., and Haxby, J. V. (2004). Beyond sensory images: Object-based representation in the human ventral pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 101(15):5658–5663.
- Pisella, L., Gréa, H., Tilikete, C., Vighetto, A., Desmurget, M., Rode, G., Boisson, D., and Rossetti, Y. (2000). An 'automatic pilot' for the hand in human posterior parietal cortex: Toward reinterpreting optic ataxia. *Nature Neuroscience*, 3(7):729–736.

- Pochon, J.-B., Levy, R., Poline, J.-B., Crozier, S., Lehericy, S., Pillon, B., Deweer, B., Le Bihan, D., and Dubois, B. (2001). The role of dorsolateral prefrontal cortex in the preparation of forthcoming actions: an fmri study. *Cerebral cortex*, 11(3):260–266.
- Poirier, C., Collignon, O., Scheiber, C., and De Volder, A. (2004). Auditory motion processing in early blind subjects. *Cognitive Processing*, 5(4):254–256.
- Poirier, C., Collignon, O., Scheiber, C., Renier, L., Vanlierde, A., Tranduy, D., Veraart, C., and Volder, A. G. D. (2006). Auditory motion perception activates visual motion areas in early blind subjects. 31:279–285.
- Poirson, A. B., Wandell, B. A., Varner, D. C., and Brainard, D. H. (1990). Surface characterizations of color thresholds. *JOSA A*, 7(4):783–789.
- Purpura, G., Cioni, G., and Tinelli, F. (2017). Multisensory-Based Rehabilitation Approach : Translational Insights from Animal Models to Early Intervention. 11(July):1–6.
- Quinn, P. C., Cummins, M., Kase, J., Martin, E., and Weissman, S. (1996). Development of categorical representations for above and below spatial relations in 3- to 7-month-old infants. *Developmental Psychology*, 32:942–950.
- Raab, D. H. (1961). Statistical facilitation of simple reaction-times.
- Raij, T., Ahveninen, J., Lin, F. H., Witzel, T., Jääskeläinen, I. P., Letham, B., Israeli, E., Sahyoun, C., Vasios, C., Stufflebeam, S., Hämäläinen, M., and Belliveau, J. W. (2010). Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *European Journal of Neuroscience*, 31(10):1772–1782.
- Rao, S. M., Mayer, A. R., and Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature neuroscience*, 4(3):317–323.
- Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. *Journal of neurophysiology*, 89(2):1078–1093.
- Redon, C. and Hay, L. (2005). Role of visual context and oculomotor conditions in pointing accuracy. *NeuroReport*, 16(18).
- Regan, D. and Gray, R. (2000). Visually guided collision avoidance and collision achievement. *Trends in cognitive sciences*, 4(3):99–107.
- Reisbeck, T. E. and Gegenfurtner, K. R. (1999). Velocity tuned mechanisms in human motion processing. *Vision Research*, 39(19):3267–3286.
- Renier, L., De Volder, A., and Rauschecker, J. (2014). Cortical plasticity and preserved function in early blindness. *Neuroscience and Biobehavioral Reviews*, 41:53–63.
- Retsa, C., Matusz, P. J., Schnupp, J. W., and Murray, M. M. (2020). Selective attention to sound features mediates cross-modal activation of visual cortices. *Neuropsychologia*, 144(March).
- Ricciardi, E., Bonino, D., Pellegrini, S., and Pietrini, P. (2014). Mind the blind brain to understand the sighted one! Is there a supramodal cortical functional architecture? *Neuroscience and Biobehavioral Reviews*, 41:64–77.



- Ricciardi, E., Papale, P., Cecchetti, L., and Pietrini, P. (2020). Does (lack of) sight matter for V1? New light from the study of the blind brain. *Neuroscience and Biobehavioral Reviews*, 118(July):1–2.
- Riečanský, I. (2004). Extrastriate area V5 (MT) and its role in the processing of visual motion. *Ceskoslovenská fyziologie / Ústřední ústav biologický*, 53(1):17–22.
- Roach, N. W., McGraw, P. V., Whitaker, D. J., and Heron, J. (2017). Generalization of prior information for rapid Bayesian time estimation. *Proceedings of the National Academy of Sciences of the United States of America*, 114(2):412–417.
- Rockland, K. S. and Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *International Journal of Psychophysiology*, 50(1-2):19–26.
- Röder, B., Hamburg, U., and Kekunnaya, R. (2022). *Effects of Early Visual Deprivation Consequences of Total Blindness in Humans Behavioral and Functional Brain Changes Following Congenital Blindness in Hu*. Number July.
- Röder, B., Stock, O., Bien, S., Neville, H., and Rösler, F. (2002). Speech processing activates visual cortex in congenitally blind humans. *European Journal of Neuroscience*, 16(5):930–936.
- Röder, B., Teder-Sälejärvi, W., Sterr, A., Rösler, F., Hillyard, S. A., and Neville, H. J. (1999). Improved auditory spatial tuning in blind humans. *Nature*, 400(6740):162–166.
- Roelfsema, P. R. and de Lange, F. P. (2016). Early Visual Cortex as a Multiscale Cognitive Blackboard. *Annual review of vision science*, 2:131–151.
- Rohlf, S., Li, L., Bruns, P., and Röder, B. (2020). Multisensory Integration Develops Prior to Crossmodal Recalibration. *Current Biology*, 30(9):1726–1732.e7.
- Romei, V., Murray, M. M., Cappe, C., and Thut, G. (2009). Preperceptual and Stimulus-Selective Enhancement of Low-Level Human Visual Cortex Excitability by Sounds. *Current Biology*, 19(21):1799–1805.
- Rosenblum, L. D., Dias, J. W., and Dorsi, J. (2017). The supramodal brain: implications for auditory perception. *Journal of Cognitive Psychology*, 29(1):65–87.
- Rowland, B., Stanford, T., and Stein, B. (2007). A Bayesian model unifies multisensory spatial localization with the physiological properties of the superior colliculus. *Experimental Brain Research*, 180(1):153–161.
- Rushton, S. K. and Wann, J. P. (1999). Weighted combination of size and disparity: A computational model for timing a ball catch. *Nature Neuroscience*, 2(2):186–190.
- Sadato, N. (2006). Chapter 11 Cross-modal plasticity in the blind revealed by functional neuroimaging. *Supplements to Clinical Neurophysiology*, 59(C):75–79.
- Sarmiento, B. R., Matusz, P. J., Sanabria, D., Murray, M. M., and Murray, M. (2015). Contextual Factors Multiplex to Control Multisensory Processes.

- Sarrazin, J.-C., Giraudo, M.-D., Pailhous, J., and Bootsma, R. J. (2004). Dynamics of balancing space and time in memory: tau and kappa effects revisited. *Journal of Experimental Psychology: Human Perception and Performance*, 30(3):411.
- Sathian, K. (2005). Visual cortical activity during tactile perception in the sighted and the visually deprived. *Developmental Psychobiology*, 46(3):279–286.
- Savelsbergh, G., Whiting, H., Burden, A., and Bartlett, R. M. (1992). The role of predictive visual temporal information in the coordination of muscle activity in catching. *Experimental brain research*, 89:223–228.
- Scheidt, R. A., Dingwell, J. B., and Mussa-Ivaldi, F. A. (2001). Learning to move amid uncertainty. *Journal of neurophysiology*, 86(2):971–985.
- Scheidt, R. A. and Stoeckmann, T. (2007). Reach adaptation and final position control amid environmental uncertainty after stroke. *Journal of neurophysiology*, 97(4):2824–2836.
- Schmidt, R. A., Zelaznik, H., Hawkins, B., Frank, J. S., and Quinn Jr., J. T. (1979). Motor-output variability: A theory for the accuracy of rapid motor acts. *Psychological Review*, 86:415–451.
- Senna, I., Parise, C. V., and Ernst, M. O. (2015). Hearing in slow-motion : Humans underestimate the speed of moving sounds. *Nature Publishing Group*, pages 1–5.
- Senna, I., Parise, C. V., and Ernst, M. O. (2017). Modulation frequency as a cue for auditory speed perception. *Proc. R. Soc. B*, 284(1858).
- Senna, I., Piller, S., Gori, M., and Ernst, M. (2022). The power of vision: calibration of auditory space after sight restoration from congenital cataracts. *Proceedings of the Royal Society B: Biological Sciences*, 289(1984).
- Setti, W., Cuturi, L. F., Cocchi, E., and Gori, M. (2018). A novel paradigm to study spatial memory skills in blind individuals through the auditory modality. *Scientific Reports*, 8(1):1–10.
- Setti, W., Cuturi, L. F., Cocchi, E., and Gori, M. (2022a). Spatial Memory and Blindness: The Role of Visual Loss on the Exploration and Memorization of Spatialized Sounds. *Frontiers in Psychology*, 13(May):1–12.
- Setti, W., Cuturi, L. F., Engel, I., Picinali, L., and Gori, M. (2022b). The Influence of Early Visual Deprivation on Audio-Spatial Working Memory. *Neuropsychology*, 36(1):55–63.
- Shams, L., Kamitani, Y., and Shimojo, S. (2000). What you see is what you hear. *Nature*, 408(6814):788.
- Shams, L., Thompson, S., Shimojo, S., and Allman, J. (2002). Sound-induced illusory visual motion. *Journal of Vision*, 2(7):405–405.
- Shipley, T. (1964). Auditory flutter-driving of visual flicker. *Science*, 145(3638):1328–1330.
- Simon, O., Mangin, J.-F., Cohen, L., Le Bihan, D., and Dehaene, S. (2002). Topographical Layout of Hand, Eye, Calculation, and Language-Related Areas in the Human Parietal Lobe. *Neuron*, 33(3):475–487.

- Smeets, J. and Brenner, E. (2004). Curved movement paths and the hering illusion: Positions or directions? *Visual Cognition*, 11(2-3):255–274.
- Smeets, J. B. and Brenner, E. (1995). Perception and Action Are Based on the Same Visual Information: Distinction Between Position and Velocity. *Journal of Experimental Psychology: Human Perception and Performance*, 21(1):19–31.
- Soto-Faraco, S., Kingstone, A., and Spence, C. (2003). Multisensory contributions to the perception of motion. *Neuropsychologia*, 41(13):1847–1862.
- Soto-Faraco, S., Kingstone, A., and Spence, C. (2004). Cross-Modal Dynamic Capture: Congruency Effects in the Perception of Motion Across Sensory Modalities. *Journal of Experimental Psychology: Human Perception and Performance*, 30(2):330–345.
- Soto-Faraco, S., Lyons, J., Gazzaniga, M., Spence, C., and Kingstone, A. (2002). The ventriloquist in motion: Illusory capture of dynamic information across sensory modalities. *Cognitive Brain Research*, 14(1):139–146.
- 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(1):31–49.
- Stein, B. E., Huneycutt, W. S., and Meredith, M. A. (1988). Neurons and behavior: the same rules of multisensory integration apply. *Brain research*, 448(2):355–358.
- Stein, B. E. and Meredith, M. A. (1993). *The merging of the senses*. Cognitive neuroscience. The MIT Press, Cambridge, MA, US.
- Stein, B. E., Meredith, M. A., Huneycutt, W. S., and McDade, L. (1989). Behavioral indices of multisensory integration: orientation to visual cues is affected by auditory stimuli. *Journal of Cognitive Neuroscience*, 1(1):12–24.
- Stein, B. E. and 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., and Rowland, B. A. (2009). The neural basis of multisensory integration in the midbrain: Its organization and maturation. *Hearing Research*, 258(1-2):4–15.
- Stevens, A. A. and Weaver, K. (2005). Auditory perceptual consolidation in early-onset blindness. *Neuropsychologia*, 43(13):1901–1910.
- Stevenson, R. A., Fister, J. K., Barnett, Z. P., Nidiffer, A. R., and Wallace, M. T. (2012). Interactions between the spatial and temporal stimulus factors that influence multisensory integration in human performance. *Experimental Brain Research*, 219(1):121–137.
- Stevenson, R. A., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., Kurela, L. A. R., Siemann, J. K., James, T. W., and Wallace, M. T. (2014). Identifying and Quantifying Multisensory Integration: A Tutorial Review. *Brain Topography*, 27(6):707–730.
- Striem-Amit, E., Dakwar, O., Reich, L., and Amedi, A. (2012). The large-scale organization of "visual" streams emerges without visual experience. *Cerebral Cortex*, 22(7):1698–1709.

- Surkys, T. (2021). An auditory time perception illusion analogous to the visual Müller-Lyer illusion. *European Journal of Neuroscience*, 54(7):6512–6519.
- Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., and Leahy, R. M. (2011). Brainstorm: A user-friendly application for MEG/EEG analysis. *Computational Intelligence and Neuroscience*, 2011.
- Talsma, D., Doty, T. J., and 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., and Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, 14(9):400–410.
- Talsma, D. and Woldorff, M. G. (2005). Selective Attention and Multisensory Integration : Multiple Phases of Effects on the Evoked Brain Activity. *Journal of Cognitive Neuroscience*, pages 1098–1114.
- Teder-Sälejärvi, W., McDonald, J. J., Russo, F. D., and Hillyard, S. A. (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Cognitive Brain Research*, 14:106–114.
- Teder-Sälejärvi, W. A., Di Russo, F., McDonald, J. J., and Hillyard, S. A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *Journal of Cognitive Neuroscience*, 17(9):1396–1409.
- ten Oever, S., Romei, V., van Atteveldt, N., Soto-Faraco, S., Murray, M. M., and Matusz, P. J. (2016). The COGs (context, object, and goals) in multisensory processing. *Experimental Brain Research*, 234(5):1307–1323.
- Thinus-Blanc, C. and Gaunet, F. (1997). Representation of space in blind persons: Vision as a spatial sense? *Psychological Bulletin*, 121(1):20–42.
- Thoroughman, K. A. and Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature*, 407(6805):742–747.
- Tinelli, F., Cioni, G., and Purpura, G. (2017). Development and implementation of a new telerehabilitation system for audiovisual stimulation training in hemianopia. *Frontiers in Neurology*, 8(NOV):1–10.
- Tinelli, F., Gori, M., Beani, E., Sgandurra, G., Martolini, C., Maselli, M., Petri, S., and Purpura, G. (2022). Feasibility of audio-motor training with the multisensory device ABB: Implementation in a child with hemiplegia and hemianopia. *Neuropsychologia*, 174(August 2021):108319.
- Tinti, C., Adenzato, M., Tamietto, M., and Cornoldi, C. (2006). Visual experience is not necessary for efficient survey spatial cognition: Evidence from blindness. *Quarterly Journal of Experimental Psychology*, 59(7):1306–1328.
- Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience*, 7(9):907–915.

- Tomassini, A., Gori, M., Burr, D., Sandini, G., and Morrone, M. C. (2011). Perceived duration of visual and tactile stimuli depends on perceived speed. *Frontiers in Integrative Neuroscience*, 5(September):1–8.
- Tonelli, A., Campus, C., and Gori, M. (2020). Early visual cortex response for sound in expert blind echolocators, but not in early blind non-echolocators. *Neuropsychologia*, 147(September):107617.
- Tresilian, J. R. (1999). Visually timed action: time-out for ‘tau’? *Trends in cognitive sciences*, 3(8):301–310.
- Tresilian, J. R. (2004). The accuracy of interceptive action in time and space. *Exercise and Sport Sciences Reviews*, 32(4):167–173.
- Tresilian, J. R. and Houseman, J. H. (2005). Systematic variation in performance of an interceptive action with changes in the temporal constraints. *The Quarterly Journal of Experimental Psychology Section A*, 58(3):447–466.
- Tresilian, J. R., Oliver, J., and Carroll, T. J. (2003). Temporal precision of interceptive action: Differential effects of target size and speed. *Experimental Brain Research*, 148(4):425–438.
- Tresilian, J. R. and Plooy, A. (2006). Systematic changes in the duration and precision of interception in response to variation of amplitude and effector size. *Experimental brain research*, 171(4):421–435.
- Tresilian, J. R., Plooy, A. M., and Marinovic, W. (2009). Manual interception of moving targets in two dimensions: Performance and space-time accuracy. *Brain Research*, 1250:202–217.
- Ungerleider, L. G. and Desimone, R. (1986). Cortical connections of visual area mt in the macaque. *Journal of Comparative Neurology*, 248(2):190–222.
- Vallesi, A., Binns, M. A., and Shallice, T. (2008). An effect of spatial–temporal association of response codes: Understanding the cognitive representations of time. *Cognition*, 107(2):501–527.
- Valzolgher, C., Campus, C., Rabini, G., Gori, M., and Pavani, F. (2020). Updating spatial hearing abilities through multisensory and motor cues. *Cognition*, 204(December 2019):104409.
- Van Atteveldt, N., Murray, M. M., Thut, G., and Schroeder, C. E. (2014). Multisensory integration: Flexible use of general operations. *Neuron*, 81(6):1240–1253.
- van Beers, R. J. (2009). Motor Learning Is Optimally Tuned to the Properties of Motor Noise. *Neuron*, 63(3):406–417.
- Van Der Lubbe, R. H., Van Mierlo, C. M., and Postma, A. (2010). The involvement of occipital cortex in the early blind in auditory and tactile duration discrimination tasks. *Journal of Cognitive Neuroscience*, 22(7):1541–1556.

- van der Stoep, N., Postma, A., and Nijboer, T. C. (2017). *Multisensory Perception and the Coding of Space*. Elsevier Inc.
- Van der Stoep, N., Spence, C., Nijboer, T. C., and Van der Stigchel, S. (2015). On the relative contributions of multisensory integration and crossmodal exogenous spatial attention to multisensory response enhancement. *Acta Psychologica*, 162:20–28.
- van Wassenhove, V. and Grzeczkowski, L. (2015). Visual-induced expectations modulate auditory cortical responses. *Frontiers in Neuroscience*, 9(FEB).
- VanMarle, K. and Wynn, K. (2006). Six-month-old infants use analog magnitudes to represent duration. *Developmental science*, 9(5):F41–F49.
- Veerman, M. M., Brenner, E., and Smeets, J. B. (2008). The latency for correcting a movement depends on the visual attribute that defines the target. *Experimental Brain Research*, 187(2):219–228.
- Vercillo, T., Burr, D., and Gori, M. (2016). Early visual deprivation severely compromises the auditory sense of space in congenitally blind children. *Developmental Psychology*, 52(6):847–853.
- Vetter, P., Smith, F. W., and Muckli, L. (2014). Decoding sound and imagery content in early visual cortex. *Current Biology*, 24(11):1256–1262.
- Vicario, C. M., Pecoraro, P., Turriziani, P., Koch, G., Caltagirone, C., and Oliveri, M. (2008). Relativistic compression and expansion of experiential time in the left and right space. *PloS one*, 3(3):e1716.
- Voss, P., Lassonde, M., Gougoux, F., Fortin, M., Guillemot, J.-p., Lepore, F., and Postale, C. (2004). Early- and Late-Onset Blind Individuals Show Supra-Normal Auditory Abilities in Far-Space. *Current Biology*, 14:1734–1738.
- Voss, P., Tabry, V., and Zatorre, R. J. (2015). Trade-off in the sound localization abilities of early blind individuals between the horizontal and vertical planes. *Journal of Neuroscience*, 35(15):6051–6056.
- Wada, Y., Kitagawa, N., and Noguchi, K. (2003). Audio–visual integration in temporal perception. *International Journal of Psychophysiology*, 50(1):117–124.
- Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, 7(11):483–488.
- Waltemate, T., Senna, I., Hülsmann, F., Rohde, M., Kopp, S., Ernst, M., and Botsch, M. (2016). The impact of latency on perceptual judgments and motor performance in closed-loop interaction in virtual reality. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, 02-04-Nov:27–35.
- Wandell, B. A., Dumoulin, S. O., and Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron*, 56(2):366–383.
- Wardle, S. G. and Alais, D. (2013). Evidence for speed sensitivity to motion in depth from binocular cues. *Journal of Vision*, 13(1):1–16.

- Warren, D. H., Welch, R. B., and McCarthy, T. J. (1981). The role of visual-auditory "compellingness" in the ventriloquism effect: Implications for transitivity among the spatial senses. *Perception & Psychophysics*, 30(6):557–564.
- Warren, J. D., Zielinski, B. A., Green, G. G., Rauschecker, J. P., and Griffiths, T. D. (2002). Perception of sound-source motion by the human brain. *Neuron*, 34(1):139–148.
- Watson, J. D., Myers, R., Frackowiak, R. S., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., Shipp, S., and Zeki, S. (1993). Area v5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral cortex*, 3(2):79–94.
- Weeks, R., Horwitz, B., Aziz-Sultan, A., Tian, B., Wessinger, C. M., Cohen, L. G., Hallett, M., and Rauschecker, J. P. (2000). A positron emission tomographic study of auditory localization in the congenitally blind. *Journal of Neuroscience*, 20(7):2664–2672.
- Welch, R. B. and Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological bulletin*, 88(3):638–667.
- Wenke, D. and Haggard, P. (2009). How voluntary actions modulate time perception. *Experimental Brain Research*, 196(3):311–318.
- Witt, J. K., Proffitt, D. R., and Epstein, W. (2004). Perceiving distance: A role of effort and intent. *Perception*, 33(5):577–590.
- Witt, J. K., Proffitt, D. R., and Epstein, W. (2005). Tool use affects perceived distance, but only when you intend to use it. *Journal of Experimental Psychology: Human Perception and Performance*, 31(5):880–888.
- Wolbers, T., Zahorik, P., and Giudice, N. A. (2011). Decoding the direction of auditory motion in blind humans. *NeuroImage*, 56(2):681–687.
- Wuerger, S., Hofbauer, M., and Meyer, G. (2003a). The integration of auditory and visual motion signals at threshold. *Perception & Psychophysics*, 65:1188–1196.
- Wuerger, S., Hofbauer, M., and Meyer, G. (2003b). The integration of auditory and visual motion signals at threshold. *Perception and Psychophysics*, 65(8):1188–1196.
- Wuerger, S., Meyer, G., Hofbauer, M., Zetsche, C., and Schill, K. (2010). Motion extrapolation of auditory-visual targets. *Information Fusion*, 11(1):45–50.
- Xuan, B., Zhang, D., He, S., and Chen, X. (2007). Larger stimuli are judged to last longer. *Journal of Vision*, 7(10):2.
- Zäch, P. and Brugger, P. (2008). Subjective Time in Near and Far Representational Space. *Cognitive and Behavioral Neurology*, 21(1).
- Zago, M., McIntyre, J., Senot, P., and Lacquaniti, F. (2009). Visuo-motor coordination and internal models for object interception. *Experimental Brain Research*, 192(4):571–604.
- Zahorik, P. (2001). Estimating sound source distance with and without vision. *Optometry and Vision Science*, 78(5):270–275.

- Zanchi, S., Cuturi, L. F., Sandini, G., and Gori, M. (2022). How much i moved: Robust biases in self-rotation perception. *Attention, Perception, & Psychophysics*, pages 1–14.
- Zangaladze, A., Epstein, C. M., Grafton, S. T., and Sathian, K. (1999). Involvement of visual cortex in tactile discrimination of orientation. *Nature*, 401(6753):587–590.
- Zatorre, R. J. (2007). There’s more to auditory cortex than meets the ear. *Hearing Research*, 229(1-2):24–30.
- Zigiotto, L., Damora, A., Albin, F., Casati, C., Scrocco, G., Mancuso, M., Tesio, L., Vallar, G., and Bolognini, N. (2021). Multisensory stimulation for the rehabilitation of unilateral spatial neglect. *Neuropsychological Rehabilitation*, 31(9):1410–1443.
- Zihl, J., Von Cramon, D., Mai, N., and Schmid, C. (1991). Disturbance of movement vision after bilateral posterior brain damage: Further evidence and follow up observations. *Brain*, 114(5):2235–2252.
- Zimmer, U., Lewald, J., Erb, M., Grodd, W., and Karnath, H. O. (2004). Is there a role of visual cortex in spatial hearing? *European Journal of Neuroscience*, 20(11):3148–3156.
- Zimmermann, E. and Lappe, M. (2016). Visual space constructed by saccade motor maps. *Frontiers in Human Neuroscience*, 10(MAY2016):1–11.
- Zwiers, M. P., Van Opstal, A. J., and Cruysberg, J. R. (2001). A spatial hearing deficit in early-blind humans. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 21(9):RC142–RC145.