UNIVERSITÀ DEGLI STUDI
DI TRENTO

CIMeC

Doctoral School in Cognitive and Brain Sciences
XXIX cycle

Audiotactile interactions in the perception of temporal duration

Lia Villanueva Villarreal

Advisor

Prof. Massimiliano Zampini

Acknowledgements

Firstly I would like to deeply thank my Professor, Massimiliano, for all his teaching, time and patience. Without his rigour and orientation, I would not have reached this point.

Also, I would like to thank Massimo Vescovi for all the technical support and orientation given for my experimental work. Any thanks to the students who helped in data collection, Irene, Vittoria, and very specially Sara and Enza whose help derived in such a nice friendship.

I also want to thank my Mexican friends in Rovereto… Alicia, Inti, Sofi, Elda and Yamila…thanks for being a family to me, such a nice one, you have changed the meaning of friendship….. las quiero mucho!

I also want to thank my friends from Cimec and Rovereto, for making this stay so pleasant and unforgivable: Alisha, Rachel, Rossella, Daniel and Irene and many more.

I also want to thank some friends back in Mexico, from whom I had constant support along the entire Phd process: Laura Ortega for helping me so kindly with Matlab, and also Efrén Avila, for always being so supportive and caring. Many thanks to Azareth for opening her house and heart for me every time I needed it.

Special thanks to my parents who have always been there for me; their care and education has made me arrive here.

And very specially, from the bottom of my heart I want to thank and dedicate this work to Ulises, my partner in life, for his touching devotion and love and from whom I have had unconditional support before and throughout my Phd. No words can explain all your help to me. I am truly blessed to have you by my side… Este logro también es tuyo!

I am also very grateful with the people of my country who are now going through very hard times. Their help allowed me to study abroad. I also want to thank them for the incredible, endless display of solidarity after the earthquake, along with the memory of those who are not with us any more. We will overcome this once more. Fuerza México!

This work was supported by Conacyt (Mexico), grant project No. 208372

Table of contents

Figures Index

Index of Tables

Thesis Outline

The following thesis project investigated the interactions between audition and touch in the perception of temporal duration. As will be seen throughout the thesis, different variables were explored to further understand these modalities interactions. Chapter 1 provides a review of the state of the art of multisensory integration, with an emphasis on studies addressing audiotactile interactions in perception. Chapter 2 offers an initial experimental exploration of the interactions between audition and touch, in a duration discrimination paradigm. The next chapters are dedicated to explore how different variables that have shown to play a key role in multisensory integration, affect audiotactile interactions in the perception of duration. For example, Chapter 3 describes a study that explored whether the intensities of the modalities of audition and touch affect their interaction in the perception of duration, allowing a further understanding of this particular modality interaction. Chapter 4 illustrates a study of how the spatial location of auditory and tactile stimuli affects their interaction in the same duration perception task employed in previous chapters. Chapter 5 presents an empirical demonstration that emerged from an interest to investigate how different spatial frames of reference affect tactile duration discrimination. Chapter 6 provides a description of an experiment that tested how the senses of audition, vision and touch interact with one another in the perception of duration, in an interest to find out the hierarchies between these senses when perceiving duration. A further question relating the nature or extent of the modulation of the auditory modality over the perception of tactile duration was addressed in Chapter 7. In this chapter, a description of an experiment testing possible elongation effects of the auditory modality over the tactile one in duration perception is provided. Chapter 8 offers a general discussion of the main findings described along the following thesis project.

Chapter 1. General Introduction

1.1 The study of multisensory integration

In everyday situations people are subject to continuous and simultaneous stimulation coming from different senses. Different sensorial information may provide related hints about the characteristics of an incoming event; for example, let's consider using a razor on the skin: we take into consideration both the sound produced by the razor on the skin, whilst feeling the roughness of the skin, along with visual feedback of the action itself. These sources of information are being processed and combined simultaneously, and it would be very difficult to imagine one shaving without listening to the sound (Foxe, 2009). It has been proposed that the general capacity to integrate information from different senses in a rapid and coherent manner represents a perceptual advantage for us to make sense of the events in the world (Calvert, Spence, & Stein, 2004). However, studies in sensory perception have been traditionally dedicated to learn and characterize each sensory modality in isolation (Hornbostel, 1938), not taking into account the fact that our everyday perception involves attending to different stimulations that may come from various senses. As suggested by Pavani, Murray and Schroeder (2006), the tradition of studying the senses in isolation can perhaps stem from the overall concept of modularity, which has been widely employed in cognitive science. Modules are regarded as cognitive units that are independent and process a specific type of stimuli only. Thus, the senses were considered to operate in this manner, with information from one sense being processed separately and independently (Pavani et al., 2006). Nevertheless, it is a fact that situations from the everyday world continuously present information that excites the different senses, and is accessed simultaneously by them. Increasingly, neurophysiological research has suggested that the senses operate in concert, rather than individually (Ghazanfar & Schroeder, 2006).

Understanding how we manage to make sense and organize simultaneous sensorial information is an intriguing question that has captured attention in research over the last years mainly (Stein, 2012), leading to the study of multisensory integration.

The area of multisensory integration studies the interaction between the different sensory modalities available to perception. Specifically, it addresses the interplay of information coming from different modalities that may sometimes derive in a single percept. These interactions have been addressed by different perspectives including neurological, physiological, and behavioural levels, psychophysically, and computational modelling.

1.2 Principles of multisensory integration at the single neuron level

Different findings in the study of multisensory integration have proposed an intriguing suggestion, which establishes the existence of common principles that dictate the way in which the different modalities can interact. These principles have been extracted from neurophysiological findings, suggesting that they may work on any combination of modalities, despite the nature of the modalities at play (Calvert et al., 2004). The following principles have been suggested to explain multisensory integration to occur in the superior colliculus at a neural level. However, these principles might also be generalized to behavioral phenomena (Holmes & Spence, 2005). These principles can be described as the following: proximity in space; proximity in time and inverse effectiveness rule.

Proximity in space. Multisensory integration is more likely to occur if the stimuli coming from different modalities share a location or are close to each other. The effects of close spatial proximity are reflected in facilitation effects in multisensory integration in a specific type of tasks that involve the spatial discrimination of stimuli. For example, Ro, Hsu, Yasar, Elmore and Beauchamp (2009) demonstrated that when presenting simultaneously auditory irrelevant stimuli on the same side as tactile stimulation (Experiment 2), leaded to a greater enhanced tactile detection as when compared to the opposite side appearance of the sound and tactile stimuli, as well as when compared with unimodal tactile stimuli detection. However, studies that involve temporal discrimination or the identification of target stimuli and also manipulate spatial coincidence of bimodal stimuli, have proven not to be affected by the spatial concurrence of different modalities, as highlighted by Spence (2012). This can be seen in for instance, in the study of the two-flash illusion. Here, participants usually report the number of visual events as a result of the number of presented concurrent auditory beeps, despite that the number of presented visual events is different (i.e. usually less visual flashes). Innes-Brown and Crewther (2009) used this illusion and controlled also the spatial congruence of both the flashes and the sounds. Their findings indicated that the illusion was present with congruent and also incongruent spatial coincidence between the flashes and the beeps, thus concluding that spatial proximity did not play a role in the illusion (Innes-Brown & Crewther, 2009).

Proximity in time. Closeness in the occurrence of stimuli from different modalities is an important factor for multisensory integration; when these are presented simultaneously or very close in time, multisensory integration is very likely to occur. There are different temporal windows that allow integration (Holmes & Spence, 2005). For example, findings reported by Meredith, Nemitz and Stein (1998) establish different temporal windows, with up to 1500 ms of separation between the occurrences of stimuli from different senses, and still allowing integration. The temporal windows of integration depend on the type of stimuli and modalities involved, since they obey to differences in processing and transduction times for each modality, for example, sometimes an auditory stimulus should precede a visual one (by various tens of ms, according to the type of stimuli), in order for integration to occur and allow synchrony perception (Van Eijk, Kohlrausch, Juola & van de Par, 2008; Spence, 2012).

Inverse effectiveness rule. This principle states that multisensory integration increases as the response to unimodal stimuli diminishes (Holmes, 2009). According to Stein, Stanford, Ramachandran, Perrault and Rowland (2009), the inverse effectiveness principle can be observed in two different ways: in enhancement or in lower enhancement of the multisensory response at a neuronal level. Enhancement in this case, suggests that stimuli from two modalities that are close in time and space generate more impulses than when presented individually (Stein et al., 2009). In lower enhancement, a stimulus from one modality can inhibit the neuronal response to a stimulus from the other modality (Jiang & Stein, 2009). Therefore, if unimodal stimuli provoke strong responses (in isolation), when they are presented in combination they tend to generate a lower multisensory enhancement. (Stein et al., 2009). Contrarily, when unimodal responses are weaker, the multisensory integration yields a more vigorous multisensory response, as registered by neurons from the superior colliculus (Jiang & Stein, 2003), resulting in a higher enhancement of the response.

Individual modality neurons respond to sensory events, generating a given number of spikes. Surprisingly, when these individual spikes are added, the number is always less than when a multisensory situation is present, so, there is a greater number of spikes when two modalities interact, (which has been termed the 'superadditivity effect'). That is, a multisensory situation will provide a greater number of spikes; however, the origin of these additional spikes is unknown (Holmes & Spence, 2005). This effect usually happens when one of the sensory modalities (in isolation) produces a low intensity or '*weak*' response, but could also occur with higher intensity stimuli.

1.3 Multisensory Integration at the behavioural and perceptual level

The interaction between information coming from different modalities may occur in different behavioural forms, according to some phenomena reported in research of multisensory integration. Some types of known interactions are the following:

a) Combination
b) Fusions
c) Cross modal illusions
d) Modality dominance

There is a degree of overlapping between these mechanisms of multisensory integration. However, for clarity, these mechanisms are explained separately below.

1.3.1 Combination and fusion

To best explain the processes of 'combination of information' and 'fusion', the McGurk effect represents a useful and interesting example. This effect was initially documented by McGurk and MacDonald (1976) and reports an interaction of visual and auditory information in speech perception. To determine the extent to which visual information like lip movements would alter the perception of speech, the authors made participants listen and watch a video of a woman pronouncing certain auditory syllables (i.e. /ba/) and dubbed the lip movements with a different syllable from the original (i.e. /ga/). Three conditions were tested: 1) looking at the video with incongruent visual information (lip movements) from the auditory input, 2) listening to the auditory input only, and 3) looking at the video with congruent auditory and visual information. Their results showed that when asking subjects to state the syllable they heard, they were very accurate when attending only to the auditory information as well as when they watched the congruent information video (providing the corresponding syllable for each condition), whereas looking at the incongruent video made subjects perceive a transformed or combined sound (i.e. /da/).

It seems clear that in the McGurk effect, differences in the resulting percept depend on the nature of the manipulation. For example, one finding in the study of McGurk and MacDonald (1976) shows that coupling an auditory disyllable /baba/ with a dubbed visual disyllable /gaga/, makes subjects perceive the auditory disyllable /dada/. Here, it is possible to notice a transformation in the final percept in relation to the initial visual and auditory disyllables, indicating an interaction from the two modalities that reflected in a new or transformed feature (the letter ‘d’) however, preserving also some initial features (‘a’); this represents an example of a fusion effect between the two modalities (Bertelson & de Gelder , 2004), in which old features of any of the modalities remain, but also new features emerge.

These authors also pinpoint a variation of the McGurk effect, where the final percept gathers all the information from the initial syllables in a unified result:

auditory /da/+ visual /ba/ = auditory/bda/

Here, the final percept represents the result of a combination of information, where information from the different modalities is all added or brought together in a single percept. While these processes suggest interactions in the form of combinations or fusions of the involved perceptual information, research has suggested there are other types of interplay between modalities.

1.3.2 Cross-modal illusions reflect an interplay of modalities

In some cases, the mechanisms of multimodal interaction described above, such as combination and fusion and modality dominance, can produce errors or distortions in perception, which are addressed in the literature as cross-modal illusions. Examples of these illusions are the Ventriloquist effect mentioned above (Howard & Templeton, 1966), where a visual image provokes the mislocation of an auditory event; and the McGurk effect (McGurk & MacDonald, 1976), where a new phoneme is produced, as previously described. In these illusions, the relationship between the senses may be so powerful that the multimodal interactions can generate distorted perceptions or lead to the perception of events or properties that do not exist.

The ventriloquist effect occurs due to the combination of various factors, which include spatial disparity between the visual and the auditory events (which should be small), timing between the events, along with ‘compellingness’ (i.e., the voice of the ventriloquist should be appropriate for the puppet’s appearance, and the ventriloquist has to make sure his lips do not move; Recanzone, 2009). The ventriloquist effect has been given different explanations. One of them suggests that, since the visual and the auditory events occur at approximately the same time and near the same space, the events tend to be considered as a single one by the perceptual system, because there is a previous assumption that an event is allowed to have one location only. Therefore, the difference in location between the visual and auditory events is usually solved by assuming the location of the visual event, because the visual system has a better spatial resolution than the auditory system (Recanzone, 2009; Vroomen & DeGelder, 2004). Another explanation is that there is a response competition and usually the visual location is the selected response (Vroomen & DeGelder, 2004).

The study of cross-modal illusions is interesting since it clearly depicts close relationships between different senses that could be conceived as acting separately. Numerous cross-modal illusions have been documented in various combinations of modalities, mostly involving vision and audition, but also between these modalities and touch.

1.3.3 Modality dominance

The term modality dominance refers to a mechanism of multisensory integration in which the information of an event coming from one modality totally impacts the perception, or partially modulates, what is perceived in another interacting modality (Bertelson & de Gelder, 2004). The modality dominance has been subject to recent research as an attempt to understand under what conditions one modality may influence the other, how this influence is manifested, and whether this sensory dominance can be reversed. The findings of some relevant studies indicate that modality dominance may be partial or total.

Total dominance: When having various stimulations coming from different senses, for instance in a bimodal situation, people sometimes tend to attend to only one of them, (e.g., the dominance of visual stimuli over auditory stimuli in Colavita, 1974)

Partial dominance: When having a bimodal situation, a modality may influence another modality by sometimes indicating what will be perceived in the other modality or altering the final percept in the target modality (e.g., visual flash illusion induced by sound, reported by Shams, Kamitani, & Shimojo, 2000, 2002)

Relevant examples of modality dominance have been documented with the study of visual dominance. The Colavita effect (Colavita, 1974) provides an experimental example of a situation where visual information dominates other interacting modalities. The author created an experimental situation where visual (e.g., a light) and auditory (e.g., a tone) stimuli were presented individually and, on few trials, simultaneously. Subjects were asked to press a key whenever they perceived a light and another key when they perceived a tone. When presented alone, subjects were very accurate and fast when responding to the presence of the stimuli, presenting similar reaction times for each modality. Moreover, under some conditions auditory events showed faster reaction times. On the trials where auditory and visual information was presented simultaneously, participants responded only to the visual situation and some of them reported even not having perceived the auditory stimuli at all (Colavita, 1974). The visual dominance found in the Colavita effect prevails despite differences in the intensities of the auditory and visual stimuli (e.g. auditory with a higher intensity), as well as matching intensities between the modalities; this visual dominance also extends to touch and is kept irrespective of the type of stimuli employed, or the position of the stimuli involved and also when attention is explicitly directed towards the other modalities (see Spence, Parise & Chen, 2012, for an extensive review of the studies documenting this dominance).

Spence et al. (2012) suggest that the Colavita visual dominance effect could be explained by the biased competition hypothesis proposed by Desimone and Duncan (1995), which suggest that brain systems compete with one another. Probably, the sensory system is itself of a competitive nature, where the most salient stimuli have a competitive lead that can direct attention to those stimuli, while inhibiting the neural activity from weaker stimuli (Desimone & Duncan, 1995; Spence et al., 2012). Therefore vision might represent a salient group of stimuli that impacts the perception of other modality stimuli (Spence et al., 2012). In addition to this explanation lies the fact that half of the human cortex is concerned with the processing of visual events (Sereno et al., 1995) thus suggesting that visual dominance might have a 'hardwired physiological basis' (Spence et al., 2012).

1.3.4 Theories and models explaining modality dominance

A number of explanations have been offered to explain the mechanisms underlying the modality dominance. Initially, Welch and Warren (1980) proposed the 'Modality appropriateness theory'. This theory suggests that, in a multisensory situation, the dominance of a given modality is dependent on its relevance to the task. Attentional mechanisms have been also proposed to explain modality dominance. Sinnett, Spence and Soto-Faraco (2007) suggested that the dominance of vision in the specific case of the Colavita effect could be explained by the attentional processes involved. The visual channel may be more 'available' when information from that stream is more frequent or sampled before auditory information. However, when trying to manipulate attention to other modalities to reverse the visual dominance, the results showed this was not possible (Sinnett et al., 2007).

Ernst and Banks (2002) further proposed the ‘Maximum Likelihood Estimation Model’. Given a bimodal situation, this model has proven to be a good predictor of which modality will tend to dominate. Based on computed variances from estimates of a given property (e.g., height) by both modalities, the model successfully predicts which one is the most reliable, as human performance would do. The overall idea is to simulate what the brain possibly does, which is to take into account these variances in estimates, and choose the modality whose percept is the most stable (Ernst & Banks, 2002). This model has extensively been recognized as a reliable method to understand sensory weightings and dominance effects. The model is also called optimal, since the final multisensory estimate is the product of the sum of unimodal information, and has the maximum consistency or reliability (Alais & Burr, 2004).

Since the main purpose of the current research project is to document the existence of audiotactile interactions, it is necessary to first illustrate that these modalities are related. Several works suggest that the basis for an interaction between these modalities stems from the similarities they both share.

1.4 Integration between audition and touch

This section reviews the literature discussing similarities and differences between audition and touch, along with works that document the interactions between both modalities.

1.4.1 Similarities and differences between audition and touch

As von Békésy (1959) has suggested, the first important similarity found in these senses is the fact that both modalities are receptive to the same type of physical stimulation, that is, mechanical vibrations. As Soto-Faraco and Deco (2009) indicate, there are some frequencies to which both senses are sensitive to and react to them at the same time, although the qualia experienced is very different. There is also another range of frequencies that only one modality in particular can process (Soto-Faraco & Deco, 2009); several studies document these ranges in both touch (from 3Hz, Talbot et al., 1968) and audition (from 20 Hz to 20 KHz, Goldstein, 2010).

Von Békésy (1959) draws another similarity regarding the vibratory patterns in the skin and the ear. When being stimulated by a vibratory wave, both the basilar membrane in the ear, and the skin tissue, have these waves travel along their surfaces (Von Békésy, 1959). In the skin, the sensation is felt only in close proximity to the surface that was stimulated by a given vibration, however, in the ear, this is dependent on the frequency of the vibration (Von Békésy, 1959). The author suggests that part of the similarities between these senses can come from evolutionary aspects: as the author states, the organ of Corti, located in the ear and with sensory cells that are subject to vibrations, has evolved from skin tissue in mammals. Von Békésy (1959) also suggested that the sense of touch could be used as an analogy to infer aspects of hearing that cannot be directly assessed.

An interesting and plausible explanation as to why these two senses can be in close relation is also exposed in Occelli, Spence and Zampini's (2011) review regarding the development of the senses in a prenatal stage, where it is highlighted that touch (being the first actual sense to develop) and audition are developed prior to vision (Moore & Persaud, 2008), making this synchronicity in development a key for audiotactile connections to exist (Lickliter & Bahrick, 2000). It has been demonstrated that early prenatal audiotactile interactions occur when both kinds of stimuli are presented to the human fetus; it has been seen that heart rate and body movements are more incremented than for unimodal presentations (Kisilevsky & Muir, 1991)

It is worth mentioning that information received from the vibratory frequencies from tactile stimulation is processed in regions that are proximate in the cerebral cortex (Ro et al., 2009), and that, according to a study by Foxe et al. (2002) using neuroimaging techniques, there are interactions between the auditory and the somatosensory cortex.

1.5 Interactions between audition and touch

Events that involve tactile stimulation often involve the generation of sounds. At the same time, audition and touch share a number of similarities. Although these modalities may provide different qualitative experiences, they both respond to the same type of physical stimulation, they have a common evolutionary origin and present interconnections as described earlier. Therefore, it may come as no surprise to consider the possibility that audition may have a participation in the perception of tactile events and viceversa. Although this type of multisensory interaction has not been very much documented, the current literature suggests some ways in which audition and touch may be related.

Multimodal interactions between vision, audition and other modalities have been thoroughly documented. In comparison, the study of interactions between touch and other senses has received less attention. The present research project investigates the interactions between touch and audition in time perception, specifically, duration. Studying these particular interactions represents part of the general interest for understanding how the information coming from different modalities is integrated.

In the next sections, few examples are drawn to exemplify influences of audition and touch over one another, in different aspects of perception, to name a few.

1.5.1 Influences of audition over touch

Specific examples of changes in tactile perception induced by sound are the studies reported by Bresciani et al., (2005) and Hötting and Röder (2004) where the number of auditory events influenced the number of perceived tactile events. The number of perceived tactile taps incremented in accordance to the number of presented tones.

Another tactile illusion induced by audition was reported by Jousmäki and Hari (1998) as ‘The parchment skin illusion’. These authors asked participants to listen to the sound of their hands rubbing. When the pitch and loudness of the rubbing hands sound was manipulated to be higher, people reported they perceived their own skin as more smooth/dry whereas a lower pitch sound generated the sensations of roughness/moisture. However, the method reported on this study is not very clear and lacks rigorous methods. For example, they used a scale that combined dimensions (smooth/dry-rough/moist), therefore it was unclear to what dimension where the participants actually responding, and also it is not theoretically justified why the authors link dimensions such as smoothness to dryness (in the unified scale). Nevertheless the results of this study were replicated by Guest, Catmur, Lloyd and Spence (2002), finding an illusion but in the opposite direction (Experiment 2). In Experiment 1, Guest et al. (2002) had the participants determine the roughness of different surfaces (i.e. a smooth and a rough sandpaper) that they touched for a short period. The sounds produced by the touching of the surfaces were sometimes altered (certain frequencies were attenuated, amplified or were not altered at all, as a control condition). Error rates were analyzed, showing that the attenuation of the high frequencies led to a 'smoother' perception of the texture, and therefore an increment in the high frequencies led to a 'rougher' perception of the textures (Guest et al., 2002). In Experiment 2, participants were asked to rub their hands and rate the dryness or moistness of their skin and the sounds generated by rubbing the hands were controlled in the same manner as Experiment 1. In this case, the authors employed separate scales for the dryness and moistness factors (improving Jousmäki & Hari, 1998 design). Results indicated that attenuating the high frequencies led to a sensation of dryness (which is consistent with Jousmäki & Hari, 1998 results), however increasing the high frequencies led to sensations of rough skin (consistent with the findings from Experiment 1, but opposite to Jousmäki & Hari, 1998). The differences in the findings of the works from Jousmäki and Hari (1998) and Guest et al. (2002) could be explained due to differences in the methods employed, where the latter authors clearly improved the design, as described earlier.

Zampini and Spence (2004) documented an effect of changing auditory inputs online when judging the crispness and staleness of potato chips. Here, the authors altered the frequencies and loudness of the auditory sounds produced when people bit crisps, and found that by increasing these auditory levels, people judged crisps as being fresher and crisper.

1.5.2 Influences of touch over audition

Gillmeister and Eimer (2007) found that auditory perception is enhanced when presenting tactile irrelevant stimuli. In their task (Experiment 1), participants were asked to detect the presence of an auditory stimulus that could be presented in one of two visually marked intervals. Tactile irrelevant stimuli were presented in both intervals and could be simultaneous or asynchronous with the auditory stimulus. Results showed that the detection of auditory stimuli was easier when tactile irrelevant stimuli were introduced (Gillmeister & Eimer, 2007). Additionally, subjects were asked to rate the perceived intensity of auditory stimuli (which varied in intensities from 64-78 dB; Experiment 2) whilst ignoring tactile stimuli (which were presented synchronously or asynchronously with the auditory stimuli). Findings indicated that the perceived loudness of tones was judged as higher when the distractors were present. The authors found that when using auditory stimuli with lower intensities, the effect of the distractor modality was higher than when employing higher intensities. This fact clearly illustrates the application of the inverse effectiveness rule explained earlier (Gillmesiter & Eimer, 2007).

Another area for exploring audiotactile interactions can be found in the study of time perception. Little has been documented regarding these sensory interactions in the perception of the important dimension of time. The following sections describe and illustrate the studies dedicated to further understand the relationship between these modalities in this feature.

1.6 Exploring audiotactile interactions through the study of time perception.

The question of how humans are able to perceive time is a topic that has intrigued researchers for long ago (Grondin, 2001). It is also true, that there are no specialized ‘receptors’ in humans dedicated to process and acquire the notion of time (Hasuo, Kuroda & Grondin, 2014). It is then through the information that arrives through our different senses, that temporal information is acquired; this is one of the reason for suggesting that temporal perception might be an amodal (Hasuo et al., 2014). Despite this lack of specialized receptors for time processing, however, there are experimental demonstrations that show that some sensory modalities are more reliable than others when perceiving the temporal dimension. For example, previous work has shown that the auditory modality provides good temporal resolution (Morein-Zamir, Soto-Faraco & Kingstone, 2003; Repp & Pennel, 2002), and that it is a more reliable and precise modality, at least in contrast to vision, for carrying out temporal tasks (Fujisaki, Kitazawa & Nishida, 2012; Ortega, Guzman-Martinez, Grabowecky & Suzuki, 2014; Romei, De Haas, Mok & Driver, 2011).

There is an ongoing debate regarding the existence of unitary clocks for each of the sensory modalities to perceive time (Gamache & Grondin, 2010) or whether there is a central internal clock in charge of this (e.g., the early model of Treisman, 1963, to name one). Few models have been suggested by the philosophical and psychological literature; some of them are further addressed in the section regarding 'Temporal perception models'.

1.7 Assessing time perception through the study of multisensory interactions

Time can be manifested in different 'instances', for example, synchronicity of events, the perception of the duration of an event, temporal order, sequence perception or rhythm, to name a few. These instances have been studied through different experimental tasks and different modality combinations, (i.e. bimodal situations) (Fujisaki et al., 2014); so far, we can find an important amount of work documenting audiovisual combinations in the study of temporal order judgments (Morein-Zamir et al., 2003; Zampini, Shore & Spence, 2003), synchronicity (Fujisaki & Nishida, 2005), simultaneity judgments (Zampini, Guest, Shore & Spence, 2005), and more recently, duration (Chen & Yeh, 2009; Klink, Montijn, & Wezel, 2010; Ortega et al., 2014; Romei et al., 2011).

Some of these temporal tasks are also documented for audiotactile combinations, for example in the study of temporal order judgments (Kitagawa, Zampini & Spence, 2005; Occelli, Spence & Zampini, 2008; Zampini et al., 2005), synchrony of events (Fujisaki & Nishida, 2009 and rhythm perception (Patel, Iversen, Chen & Repp 2005; Occelli et al. 2011, for a complete review of audiotactile interactions in temporal perception). Some of these works highlight the better temporal resolution of the audiotactile combinations in contrast to audiovisual combinations when studying simultaneity, temporal order, and synchronicity (Fujisaki et al, 2012; Fujisaki & Nishida, 2009).

In relation to the study of audiotactile combinations in temporal duration, there is one study by Mayer, Di Luca and Ernst (2014) that includes this particular modality combination, however it does not specifically address the study of sensory influences and dominance over one another in the perception of duration. This study, instead, explores the existence of single or multiple time clocks through the study of duration perception of empty and filled intervals. The onset and offsets of the empty intervals were marked by different modality combinations (audiovisual, audiotactile and visuotactile). The duration of the empty intervals ranged from 100 to 900 ms, (each of the onsets and offsets lasted 20 ms). Filled intervals were composed of a continuous sound which lasted less (30%) or more (from 100 up to 170%) of the duration of the empty interval. The task consisted on judging which of a pair of the empty and filled intervals was the shortest. It is also worth mentioning that the authors never used the same single modality marker for the empty intervals in their study, fact that is also different from our proposed research study. The main findings of Mayer et al. (2014) showed that stimuli with auditory onsets, were perceived as longer than stimuli with auditory offsets. Also, when assessing whether the data could be explained by a single clock model or by a multiple clocks model, they found that the pacemaker frequencies did not differ across modalities, suggesting the existence of a single clock time model that is used across all the modalities (Mayer et al., 2014).

1.8 Audiotactile interactions in temporal perception

Audiotactile interactions have been studied through temporal perception, in temporal aspects such as synchrony perception, temporal order and numerosity, mainly. Nevertheless, to the best of my knowledge, duration has been a temporal feature that has not been explored in regard to audiotactile interactions.

1.8.1 Synchrony perception

Multisensory synchrony perception refers to the moment in time where two events from different modalities are perceived to occur at the same time (Noel et al., 2015). There are modality differences in terms of transduction times (i.e. the time each modality information arrives to the brain; Noel et al., 2015; Schroeder & Foxe, 2004; Spence et al., 2001). The information coming from the different modalities reaches the sensory cortices at different times, because there are physical variances in the time of transmission of the different stimuli in the air, for example, the traveling of sound waves through air is slower than light (Navarra, Soto-Faraco & Spence, 2007; Spence, & Squire, 2003). Additionally, there are different biophysical transfer times; the transduction of sound at the ear is faster than the transduction of light at the retina (Navarra et al., 2007; Spence & Squire, 2003). An important measure in this kind of studies is the point of subjective simultaneity (PSS), which indicates the moment in time where a stimulus has to appear before the other in order to be seen as simultaneous (Occelli et al., 2011; Zampini et al., 2005). Despite the stimuli reach the brain at different times, we perceive events coming from different modalities as occurring simultaneously, all within a specific temporal window for integration. This time window is different according to each modality pairing (visuotactile, auditoactile, etc.) and across individuals (Noel et al., 2015). The just noticeable difference (JND) is relevant for these studies since it helps defining the participants' sensitivity to perceive temporal asynchronies. In particular, the JND is the smallest temporal interval between two stimuli needed for participants to be able to judge which one was presented first on 75% of trials. The JND might be used as a criterion for determining the temporal windows for integration (Navarra et al., 2005). In fact, the duration of the interval required for deciding the order of the stimuli could vary as a function of the duration of the temporal window for integration. For instance, it would require a longer interval in a TOJ task for deciding which stimuli came first (i.e., the JND should be larger) with a wider temporal window for integration.

Fujisaki and Nishida (2009) conducted a study to determine the temporal resolution or reliability of different modality pairings (AV, AT, VT, TT) finding that audiotactile pairings were the most reliable bimodal pairings when detecting synchrony between different bimodal stimuli (i.e., they provided better performance overall). The authors used a task where subjects had to determine whether the presented stimuli (in a repetitive pulse train format and single format) were synchronous or asynchronous. There were fixed time variations within the stimuli presentation (from 1.4 to 26. 7 Hz for repetitive pulse trains, and from 6.25 to 356.25 ms. for single presentation stimuli; Fujisaki & Nishida, 2009). In line with these findings are those reported by Noel et al. (2015), showing that audiotactile synchrony perception is more stable in contrast to other pairings, AT having the highest temporal resolution. Noel et al. (2015) also found that the physical time difference between auditory and tactile stimuli correlates with the perceived simultaneity in AT pairings.

An interesting question that arises in the interactions between the modalities lies in the amount of time necessary to generate the perception of synchrony between the information from different modalities. As previously mentioned, there are different processing and transduction times for each modality to occur in the brain (Spence et al., 2001), so a lag is necessary between the modalities to perceive synchronicity. Specifically, audiotactile combinations need a lag or separation of 80 ms. approximately, according to Zampini et al. (2005), irrespective of which modality was presented first. The tasks used to determine the gap needed to perceive synchronicity between modalities are temporal order judgment tasks.

1.8.2 Temporal order judgments

Temporal order judgments (TOJs) tasks are designed to determine temporal windows of integration among the different modalities, and consist on determining which of a pair of stimuli in two modalities comes first, when presented with different timings among the modalities.

The first studies by Hirsh and Sherrick (1961) suggested that there was a specific time gap for stimuli to be considered as occurring separately and that applied to every modality combination, however latter studies demonstrated that each pairing needs a different timing for correctly judging temporal order (Fujisaki & Nishida, 2009; Harrar & Harris, 2009; Zampini et al., 2005) and that this timing can vary between participants (Noel et al., 2015).

Harrar and Harris (2009) conducted a temporal order study with various modality pairings, separating the modalities in a range from 0, 50, 100, to 150 ms. The experiments included an exposition stage, as a way to determine the effects of learning or familiarization with the stimuli on correctly determining temporal order (in the case of audiotactile pairings, tactile stimulation was presented first), along with a test stage. Findings suggested that audiovisual pairings were influenced by the familiarization stages, showing changes in PSS, however audiotactile pairings showed a more stable and less susceptible to change PSS

(Harrar & Harris, 2009). This is in line with the previous mentioned findings documenting audiotactile pairings in synchronicity and temporal order judgments.

These kinds of studies have also sought whether the spatial proximity between modalities plays a role in determining the simultaneity of multisensory events (Harrar & Harris, 2008; Zampini et al., 2005). Zampini et al., 2005, showed however, that for audiotactile pairings, there is no effect of changing the spatial location of the pairings on performance of temporal order judgments. This finding is contrasting with studies using audiovisual (Zampini et al., 2003a, 2003b) and visuotactile temporal order judgments (Spence, Baddeley, Zampini, James & Shore, 2003), where these modality combinations have been reported to be sensitive to changes in the spatial location of the stimuli, in terms of a better performance when the bimodal stimuli are presented from different locations, as compared to the same locations (Spence, et al., 2003; Zampini et al., 2003a, 2003b). Given that this effect holds mainly for the pairings involving the visual modality (which is predominantly spatial sensible), it can perhaps suggest that audiotactile combinations are less sensitive to spatial factors (Zampini et al., 2005).

1.8.3 Numerosity

Temporal numerosity perception implies the ability to approximately determine the number of stimuli presented in succession over a short period of time. The study of this ability with a multimodal approach has produced a modest but informative number of insights for understanding the ways in which audition and touch might interact during temporal perception.

In terms of unimodal perception, audition has been identified to be the most reliable modality for temporal numerosity perception. An early study by Lechelt (1975) compared temporal numerosity discrimination across visual, auditory and tactile modalities. Participants were asked to count the number of items in a series of stimuli consisting of two to nine signals at rates ranging from 3 per second to 8 per second. Accuracy at counting visual stimuli was the lowest, resulting from an underestimation that increased noticeably after presentations rates of 3 seconds. This kind of underestimation and its increase as a function of the increase in presentation rate was also observed in tactile counting, although less dramatically. The counting of auditory stimuli was nearly perfect across rates. The advantages of audition and touch over vision in the perception of temporal numerosity have been replicated extensively, although there are mixed results in regards of which modality is more reliable. Some studies are consistent with Lechelt (1975) in reporting that audition is better than touch, for instance in terms of reliability (Bresciani, Dammeier, & Ernst, 2008), whereas other studies have reported better performance for touch (Philippi, van Erp & Werkhoven, 2008).

Multimodal research has revealed a number of conditions in which audition influences and is influenced by other modalities. The most common designs involve auditory, visual and tactile conditions. The focus here is on the findings related to the interaction between audition and touch. One research line has studied the relative advantages of merging auditory and tactile information. Philippi et al. (2008) compared whether the natural tendency to underestimate the numerosity of a stimuli set could be diminished by multimodal perception. They compared individuals' judgements about the temporal numerosity of *pulse trains*. In this case, pulse trains were composed by 2 to 10 flashes, beeps, or taps presented with an ISI of either 20, 40, 80, 160, or 310 ms., presented under unimodal visual, tactile and auditory conditions, as well as under bi-modal and tri-modal conditions. As the authors hypothesized, the lowest rates of underestimation were observed in the tri-modal condition produced. That is, individuals performed better than in the three unimodal and the two bimodal conditions. The authors concluded that multisensory perception has an advantage over unimodal perception in temporal numerosity judgements. Especially relevant for this review, the performance in the audiotactile condition was better than in the auditory condition, but not better than in the tactile one. This suggests that the accuracy of the auditory modality improved when interacting with tactile information. In contrast, the addition of auditory information did not add to any significant extent to the accuracy of the tactile perception. Another plausible explanation could be that performance was driven by the tactile stimuli only.

Another source of information for understanding the relation between audition and touch comes from research that studies the susceptibility of a given modality to be influenced by the other ones. Bresciani, Dammeier and Ernst (2008) asked participants to count the number of events (taps, flashes, and beeps) in a target modality (touch, vision or audition) whilst ignoring events in other background modality(ies). Each of the three modalities was target, with either one or both other two modalities acting as background. The target modalities presented two to four events per trial. In the background modalities the number of events could have been zero (target alone), one less (target events -1) and one more (target events + 1. In total, there were 12 experimental conditions. Overall, the background always biased individuals' responses to the target. However, vision was more susceptible to background bias than touch, and both of these modalities were more susceptible to bias than audition. Speaking specifically about audition and touch, when touch was the target the bias generated by vision and audition combined was stronger than the bias evoked by either of these modalities alone, although the bias provoked by audition was stronger. When audition was target, the bias evoked by vision and touch was stronger than the bias provoked by vision alone, but not stronger than the bias provoked by touch alone. This pattern of results was consistent with the output of prior research from the same authors that used similar paradigms and stimuli to assess the extent of the influence between audition and touch. One study reported that these modalities can influence each other, although the influence of audition over touch is stronger (Bresciani & Ernst, 2007) and that the influence of touch over audition only becomes noticeable when the stimuli from both modalities is similar (Bresciani et al., 2005).

Although it remains undefined whether audition is better than touch in the perception of temporal numerosity, there seems to be agreement in that audition has a more powerful influence on touch than the other way around. A recent study illustrates the practical applications of these findings. Bianchi, Oakley, and Kwon (2012) explored the possibilities of using numerosity in the design of computer interfaces for Personal Identification Number (PIN) mechanisms. They compared unimodal auditory and haptic cue-counting interfaces, as well as multi-modal combinations of visual and haptic or auditory information. The user evaluations indicated that unimodal haptic and auditory interfaces were more effective than multi-modal ones. This is consistent with the results of previous research, which results indicated that audition and touch are likely to bias the performance of each other.

To the best of our knowledge, there are no reported studies addressing how audition and touch modulate one another in the perception of duration, therefore, it seemed relevant to explore the question as to how each of these modalities can modulate one another in the perception of this temporal feature. Some of the experimental paradigms used to study audiovisual duration perception are addressed in the following section, since some of the methods have been employed in the current research project, which will be further described.

1.9 Duration studies

Duration can be defined as the amount of elapsed time (van Wassenhove, 2009). There are more reported studies addressing audiovisual pairings in the perception of duration than other type of modality combinations.

1.9.1 Duration studies in audiovisual combinations

Walker and Scott (1981) provided evidence that auditory information has a dominant role in the perception of the duration of visual events, being able to affect the perception of a shorter light when the sound was longer. The visual modality only produced an influence when the auditory stimulus intensity was low, in a paradigm that consisted on reproducing the perceived duration of the attended stimuli, in unimodal and bimodal conditions. Few years later, van Wassenhove, Buonomano, Shimojo and Shams (2008) showed a different tendency, where vision was able to drive the auditory duration judgments, and not the other way around, in an oddball paradigm, suing unimodal, bimodal congruent and bimodal incongruent conditions. Chen and Yeh (2009) later replicated this study using the same oddball paradigm, but however found similar results to those showed by Walker and Scott (1981), suggesting that vision had no influence in auditory comparative duration judgments, and audition was able to model visual duration judgments (Chen & Yeh, 2009). The differences in their findings have been explained as differences in the methods that they used, for example the type of stimuli employed, the experimental designs, as well as the durations employed. These findings, somehow controversial, suggested the need to further address this topic.

Romei, De Haas, Mok and Driver (2011) present two experiments which aim was to determine whether the duration of an auditory event (beeps) could influence the perceived duration of a visual event (flashes). Trials consisted on a simultaneous presentation of pairs of auditory and visual stimuli with a fixed duration stimulus in both modalities (55 ms) and a variable duration stimulus with different durations (in a range from 55 to 165 ms). Subjects had to decide which of the visual stimuli was the longest of the pair (ignoring the auditory information). Their paradigm had these conditions: congruent (where the presentations of durations both in audition and vision were the same), and incongruent (the order of durations was reversed, although the durations employed were the same), and a control condition (a unimodal visual condition). The synchrony of AV stimuli presentation was also controlled. The results showed that the number of mistakes increased significantly for the incongruent condition, and a facilitation effect was found for the congruent condition. This paradigm however, lacks a condition for testing whether the visual duration made an influence on the perceived auditory durations, as a way to compare the weight of each distractor to see which modality is more reliable in duration perception, and does not solve the debate raised by the opposite findings of Wassenhove et al. (2008) and Chen and Yeh (2009), as well as Walker and Scott (1981) regarding the dominance of audition or vision in the perception of duration.

Also, it is still not clear from their data, which exact member of the pair of divergent stimuli are they responding to (e.g., it could be a long auditory stimulus or short visual for example), so it is hard to know whether the participant’s perceived duration is really modulated by the auditory stimuli, and also, it is hard to determine whether there is a real perceptual change in the visual duration, or if participants are mainly attending to the auditory stimuli (due to its saliency and reliability in temporal domain) and ignoring the visual durations.

In line with this saliency phenomenon, Ortega et al. (2014) explored the effects of intensity of a distractor modality (i.e., the auditory modality in their study), over a target modality (i.e., the visual one), in a duration perception task. Here the intensity of the auditory modality changed in three ways in relation to the visual, *constant* intensity: a matching intensity, a higher intensity and a lower intensity. The authors found that the auditory information still dominated the perceived visual duration judgments when the intensity of the auditory stimulus was notoriously weaker, providing thus more evidence, that in temporal perception, the auditory modality could be the most reliable one.

Since the following review explores the interactions of modalities through the perception of time, it appears as relevant to explore some of the models that try to explain the perception of time through different modalities.

1.10 Temporal perception models

There have been few researchers in the domain of time perception proposing models of how humans and animals are able to estimate, perceive and reproduce time. One of the most popular models proposes the existence of a single internal clock for timing mechanisms (Treisman, 1963); the other proposes the presence of multiple clocks (Buhusi & Meck, 2009; Gamache & Grondin, 2010) that could be modality specific, and which could account for all the found differences in time perception, according to the modalities at play in a given temporal task. There is however, a lot of ongoing debate as to which model can successfully account for the perception of time (Grondin, 2010).

An early model of the single or central timing mechanism comes from Treisman (1963). In his model, there is a pacemaker and an accumulator. The pacemaker produces pulses that are counted or accumulated and this accumulation provides the experience of time (Grondin, 2001). Gibbon and Church (1984) later employed these components to build the Scalar Expectancy Model, which has been widely adopted in time perception research. This model is constituted of three different processes: clock, memory and a decision process. In the clock process, there is a switch that regulates the entry of ongoing produced pulses coming from the pacemaker, to the accumulator. In the memory processes, this accumulation of pulses is stored in a temporal (working memory) in order to be compared with a reference value stored in more permanent memory. Comparison happens in the decision processes. For a better understanding of the model, see Figure 1.

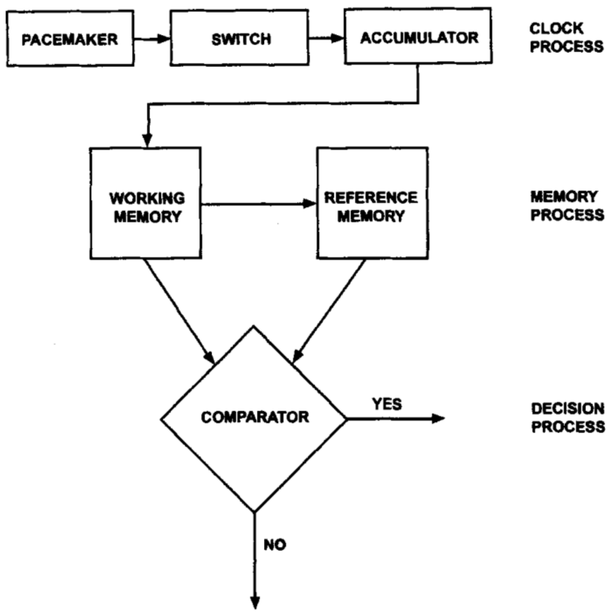


Figure 1 Time model by Gibbon and Church (1984). (From 'From Physical Time to the First and Second Moments of Psychological Time' by S. Grondin, (p.28) in *Psychological Bulletin* (2001)

Later models with some variants of this main model have been proposed. For example, among other differences, they include the presence of multiple oscillators instead of a single pacemaker, that are sensitive to different time scales, from subsecond to seconds and/or *modality* related processes (Grondin, 2001; Church & Broadbent, 1990; Mayer et al., 2014; Treisman & Brogan, 1992) and constitute the multiple modality specific clock models.

Gamache and Grondin (2010) explain the difficulty of solving the debate regarding the existence of a single or a multiple time mechanism, since, as they explain, neurophysiological evidence has highlighted the activation of sensory areas in the processing of time, bringing attention to brain specialization, rather than a single, general mechanism; nevertheless, the possibility of comparing temporal information among modalities also suggests an amodal representation of time. Up until now it is not possible to make any final conclusions.

Some of the findings described in this literature review can provide an account of the different interactions of mainly the auditory and the visual modalities, in relation to the perception of duration. Most of them agree on their findings, suggesting that audition could be the most reliable modality when perceiving duration. However, some of these studies lack experimental conditions previously described that prevent from observing the effects or dominance of vision over the duration of auditory events, etc. Also, comparison between these studies can be difficult since they often use different methodological approaches and paradigms. Given these aspects in mind, along with the lack of studies regarding audiotactile interactions in the perception of duration, the following research thesis is proposed and the experimental stages adopted to study this are described in the next section.

1.11 Research question and aims

1.11.1 Research question

The general research question of the following thesis addressed the study of how audition and touch interact with one another when perceiving duration. The purpose of the thesis was to document if and how audition and touch modulate one another (both in terms of interference and/or facilitation) when perceiving duration. A series of studies are reported here to fully address this question. They explore different variables such as spatial location and intensity of the stimuli, which have shown to play a key role in multisensory integration, in a further attempt to learn how they affect audiotactile interactions in the perception of duration.

Solving this question can add to a further understanding of the interactions between audition and touch in the perception of temporal variables. It can also provide evidence that could support some of the different models of time perception. The findings can also help to learn how the senses relate to each other in the processing of magnitudes. In more general terms, the findings could offer a greater description of multisensory processing overall, thus allowing a better and more complete understanding of human perception.

1.11.2 General hypothesis

Previous studies have shown that audition and touch share a close relation with one another and that this relation seems to be more balanced in different aspects of temporal perception, rather than one modality always dominating the other (for a review, see Occelli et. al, 2011). Therefore, based on this previous evidence it is believed that audition and touch will modulate one another similarly when perceiving duration.

1.11.3 Aims of the different studies

Study 1. Studying the interactions between audition and touch in a duration perception paradigm

The purpose of the following study was to learn how the modalities of audition and touch interact with one another (in terms of facilitation or interference effects) in the perception of duration.

Study 2. Exploring the role of intensity of the distractor modalities in an audiotactile duration paradigm

This study was designed to understand whether the use of different intensities of the auditory and tactile modalities play a role in their interactions when perceiving duration. For example, it could allow knowing whether different intensities modify the nature (e.g. symmetry) of the interaction between audition and touch in this particular temporal feature.

Study 3. Exploring the role of spatial location in an audiotactile duration perception task

The goal of this experiment was to explore how changes to the location of auditory and tactile stimuli (e.g. same/different location) affect the nature of their interaction, when perceiving duration.

Study 4. The role of arms position in a tactile duration discrimination task

This experiment represented a side study of the current thesis, and aimed to explore how different spatial frames of reference (i.e. non-crossed and crossed arms position) influence the discrimination of the duration of tactile stimuli. The motivation behind this study was to see whether the durations are perceived as longer in the crossed arms position, as a result of the updating and remapping process that occurs in the brain when the body adopts this kind of body posture. Since this process takes time, assuming a crossed arms body posture might affect the perceived duration.

Study 5. Hierarchies between the senses of Audition, vision and touch in a duration perception task

Considering previous findings showing that some modalities are better and more reliable than others for certain aspects of perception, it seemed important to address the role of the modalities of vision, audition and touch in the perception of duration. The purpose of these experiments was twofold: to test which modality was better for duration discrimination, in order to establish a hierarchy between the senses for this particular feature. Also, it seemed relevant to find which modality exerted greater influence over the other between the different modality combinations (e.g. comparing audiovisual, audiotactile, visuotactile pairings) in a duration perception task. A hierarchy pattern of these modulations could allow a further understanding of the interactions between the different modalities when perceiving duration.

Study 6. Exploring elongation effects of the auditory modality in the perception of tactile durations

The purpose of this experiment was to understand the extent of the modulatory effects of the auditory modality over the tactile one in the perception of tactile duration. More specifically, this study tested whether the perceived tactile duration was expanded by auditory durations. The aim was to see whether audition could truly change the perceived duration of subsequent tactile stimuli, after exposure of auditory longer-than-tactile durations.

For a clearer view of the studies and their relation to the general research question, see Figure 2.

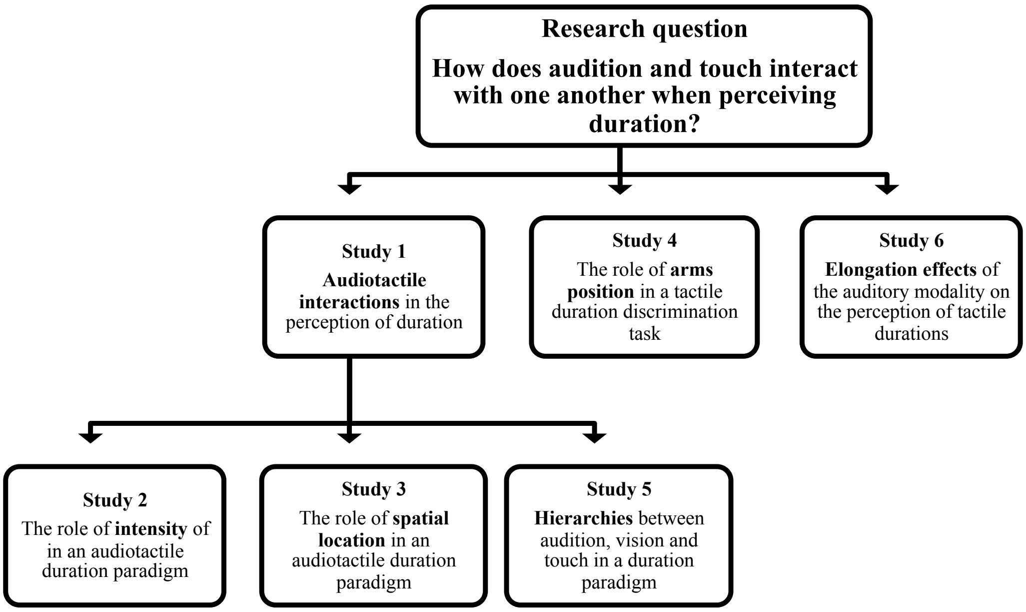


Figure 2 Diagram showing how the individual studies relate to the general question and to each other

References

Alais, D. & Burr, D. (2004). The Ventriloquist Effect results from near optimal bimodal integration. *Current Biology, 14*, 257-262.

Bertelson, P., & de Gelder, B. (2004). The psychology of multimodal perception. In C.Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 141-177). New York: Oxford University Press.

Bianchi, A., Oakley, I., & Kwon, D. S. (2012). Counting clicks and beeps: Exploring numerosity based haptic and audio PIN entry. *Interacting with Computers*, *24*, 409–422.

Bresciani, J.-P., Dammeier, F., & Ernst, M. O. (2008). Tri-modal integration of visual, tactile and auditory signals for the perception of sequences of events. *Brain Research Bulletin, 75*, 753–760.

Bresciani, J.-P., & Ernst, M. O. (2007). Signal reliability modulates auditory–tactile integration for event counting. *Neuroreport*, *18*, 1157–1161.

Bresciani, J. P., Ernst, M. O., Drewing, K., Bouyer, G., Maury, V., & Kheddar, A. (2005). Feeling what you hear: Auditory signals can modulate tactile tap perception. *Experimental Brain Research, 162*, 172-180.

Buhusi, C.V. & Meck, W.H. (2009). Relativity theory and time perception: single or multiple clocks?. *PLoS One, 4*, e6268.

Calvert, G., Spence, C., & Stein, B. E. (2004). *The Handbook of Multisensory Processes*. Massachusetts: MIT Press.

Chen, K.M & Yeh, S.L. (2009) Asymmetric cross-modal effects in time perception. *Acta Psychologica, 130*, 225-234.

Church, R.M.& Broadbent, H.A. (1990). Alternative representations of time, number and rate. *Cognition, 37*, 55-81.

Colavita, F. B. (1974). Human sensory dominance. *Attention, Perception, & Psychophysics, 16*, 409-412.

Desimone, R. & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193-222.

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

Foxe, J.J. (2009) Multisensory Integration: Frequency Tuning of Audio Tactile integration. *Current Biology*, *19*, 373-375.

Foxe, J. J., Wylie, G. R., Martinez, A., Schroeder, C. E., Javitt, D. C., Guilfoyle, D., Ritter, W., & Murray, M. M. (2002). Auditory-somatosensory multisensory processing in auditory association cortex: An fMRI study.

Journal of Neurophysiology, 88, 540-543.

Fujisaki, W., Kitazawa, S. & Nishida, S. (2012). Multisensory Timing. In B.E Stein (Ed.), *The New Handbook of Multisensory Processes* (pp. 302-317). Massachusetts: MIT Press.

Fujisaki, & Nishida, S. (2005). Temporal frequency characteristics of synchrony asynchrony discrimination of audio-visual signals. *Experimental Brain Research, 166*, 455-464.

Fujisaki, & Nishida, S. (2009). Audio-tactile superiority over visuo-tactile and audiovisual combinations in the temporal resolution of synchrony perception. *Experimental Brain Research, 198*, 245-259.

Gamache, P.L. & Grondin, S.(2010). Sensory-specific clock components and memory mechanisms: Investigation with parallel timing. *European Journal of Neuroscience, 31*, 1908-1914.

Ghazanfar, A.A. & Schroeder, C.E. (2006). Is neocortex essentially multisensory?. *Trends in Cognitive Sciences, 10*, 278-285.

Gibbon, J. & Church, R.M. (1984). Sources of variance in an information processing theory of timing. In H.L. Roitblat, T.G. Bever, & H.S. Terrace (Eds.), *Animal Cognition* (pp.465-488). Hillsdale, NJ: Erlbaum.

Gillmeister, H. & Eimer, M. (2007). Tactile enhancement of auditory detection and perceived loudness. *Brain Research, 1160*, 58-68.

Goldstein, E.B. (2010). *Sensation and Perception*. California: Wadsworth.

Grondin, S. (2001). From Physical Time to the First and Second Moments of Psychological Time. *Psychological Bulletin*, *127*, 22-44.

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception & Psychophysics, 72*, 561-582.

Guest, S., Catmur, C., Lloyd, D., & Spence, C. (2002). Audiotactile interactions in roughness perception. *Experimental Brain Research, 146*, 161-171.

Hasuo, E., Kuroda, T. & Grondin, S. (2014). About the time-shrinking illusion in the tactile modality. *Acta Psychologica, 147*, 122-126.

Harrar, V. & Harris, L. (2008)The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Experimental Brain Research, 186*, 517–524.

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

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

Holmes, N. P. & Spence, C. (2005) Multisensory Integration : Space , Time and Superadditivity. *Current Biology, 15*, 762–764.

Hornbostel, E. (1938). The unity of the senses. A source book of Gestalt psychology. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 210-216). London: Kegan Paul, Trench, Trubner & Company.

Hötting, K., & Röder, B. (2004). Hearing cheats touch, but less in congenitally blind than in sighted individuals. *Psychological Science, 15*, 60.

Howard, I. P., & Templeton, W. B. (1966). *Human spatial orientation*. London: Wiley.

Innes-Brown, H. & Crewther, D. (2009). The impact of spatial incongruence on an auditory-visual illusion. *PLoS One, 4*, e6450.

Jiang, W. & Stein, B.E. (2003). Cortex controls multisensory depression in superior colliculus. *Journal of Neurophysiology, 90*, 2123-2135.

Jousmäki, V., & Hari, R. (1998). Parchment-skin illusion: sound-biased touch. *Current Biology, 8*, 190.

Kisilevsky, B. S., & Muir, D. W. (1991). Human fetal and subsequent newborn responses to sound and vibration. *Infant Behavior and Development, 14*, 1–26.

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

Klink, P.C., Montijn, J.S. & van Wezel, R.J.A. (2011). Crossmodal duration perception involves perceptual grouping, temporal ventriloquism, and variable internal clock rates. *Attention, Perception & Psychophysics*, *73*, 219-236.

Lechelt, E. C. (1975). Temporal numerosity discrimination: Intermodal comparisons revisited. *British Journal of Psychology*, *66*, 101–108.

Lickliter, R., & Bahrick, L. E. (2000). The development of infant intersensory perception: Advantages of a comparative convergent-operations approach. *Psychological Bulletin, 126*, 260–280.

Mayer, K.M., Di Luca, M. & Ernst, M.O. (2014). Duration perception in crossmodally –defined intervals. *Acta Psychologica*, 147, 2-9.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature,* 264, 746 – 748.

Meredith, M.A., Nemitz, J.W., Stein, B.E. (1987) Determinants of Multisensory Integration Neurons . I . Temporal Factors in Superior Colliculus. *The Journal of Neuroscience*, *7*, 3215–3229.

Moore, K. L., & Persaud, T. V. N. (2008). *The developing human: Clinically* *oriented embryology*. Philadelphia: Saunders.

Morein-Zamir, S., Soto-Faraco, S. & Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Cognitive Brain Research Brain Research, 17,* 154-163.

Navarra, J., Soto-Faraco, S. & Spence, C. (2007). Adaptation to audiotactile asynchrony. *Neuroscience Letters, 413*, 72-76.

Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W. & Spence, C. (2005). Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Cognitive Brain Research, 25*, 499-507.

Noel, J.P., Wallace, M.T., Orchard-Mills, E., Alais, D. & Van der Burg, E. (2015). True and perceived synchrony are preferentially associated with particular sensory pairings. *Scientific Reports, 5*, 1-8.

Occelli, V., Spence, C., & Zampini, M. (2008). Audiotactile temporal order judgments in blind and sighted individuals. *Neuropsychologia, 46*, 2845–2850.

Occelli, V., Spence, C. & Zampini, M.(2011). Audiotactile interactions in temporal perception. *Psychonomic Bulletin Review, 18,* 429–454.

Ortega, L., Guzman-Martinez, E, Grabowecky, M & Suzuki, S. (2014). Audition dominates vision in duration perception irrespective of salience attention and temporal discriminability. *Attention, Perception & Psychophysics, 76,* 1485-1502.

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, *163*, 226–238.

Pavani, F., Murray, M.M. & Schroeder, C.E. (2006). Rethinking mind, brain and behaviour through a multisensory perspective. *Neuropsychologia, 45*, 467-468.

Philippi, T. G., van Erp, J. B. F., & Werkhoven, P. J. (2008). Multisensory temporal numerosity judgment. *Brain Research*, *1242*, 116–125.

Recanzone, G. (2009). Interactions of auditory and visual stimuli in space and time. *Hearing Research, 258*, 89-99.

Repp. B.H., & Pennel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. Journal of Experimental Psychology: *Human Perception and Performance, 28*, 1085-1099.

Ro, T., Hsu, J., Yasar, N. E., Elmore, C. L., & Beauchamp, M. S. (2009). Sound enhances touch perception. *Experimental Brain Research, 195*, 135-143.

Romei, V., De Haas, B., Mok, R.R. & Driver, J. (2011) Auditory stimulus timing influences perceived duration of co-occurring visual stimuli. *Frontiers in Psychology, 2,* 1-8.

Schroeder, C. E., & Foxe, J. J. (2004). Multisensory convergence in early cortical processing. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The Handbook of Multisensory Processes* (pp. 295– 309). Cambridge: MIT Press.

Sereno M.I, Dale A. M, Reppas J. B, Kwong K. K, Belliveau J. W, Brady T. J, Rosen B. R & Tootell R. B. H. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268*, 889-893.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature, 408*, 788.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research, 14*, 147-152.

Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: The Colavita effect revisited. *Perception & Psychophysics, 69*, 673.

Soto- Faraco, S. & Deco, G. (2009). Multisensory contributions to the perception of vibrotactile events. *Behavioural Brain Research, 196*, 145-154.

Spence, C. (2012). Multisensory Perception, Cognition and Behavior: Evaluating the factors Modulating Multisensory Integration. In B.Stein (Ed.), *The New Handbook of Multisensory Processing*. Massachusetts: MIT Press.

Spence, C., Baddeley, R., Zampini, M., James, R. & Shore, D. (2003). Multisensory temporal order judgments: when two locations are better than one. *Perception & Psychophysics, 65*, 318-328.

Spence, C., Parise, C. & Chen, Y.C. (2012). The Colavita Visual Dominance Effect. In M.M. Murray & M.T. Wallace (Eds.), *The Neural Bases of Multisensory Processes* (pp. 529-556). Florida: CRC Press.

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

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

Stein, B. E. (2012). *The New Handbook of Multisensory Processes*. Massachusetts: MIT Press.

Stein, B.E., Stanford, T.R., Ramachandran, R., Perrault, T.J.Jr. & Rowland, B.A. (2009). Challenges in quantifying multisensory integration: alternative criteria, models, and inverse effectiveness. *Experimental Brain Research, 198*, 113-126.

Talbot W.H., Darian-Smith I., Kornhuber H.H. & Mountcastle V.B. (1968). The sense of flutter-vibration: comparison of the human capacity with response patterns of mechanoreceptive afferents from the monkey hand. *Journal of Neurophysiology, 31*, 301–34.

Treisman, M. (1963). Temporal indiscrimination and the indifference interval: Implications for a model of the 'internal clock'. *Psychological Monographs, 77,* 1-31.

Treisman, M. & Brogan, D. (1992). Time perception and the internal clock: Effects of visual flicker on the temporal oscillator. *European Journal of Cognitive Psychology, 4*, 41-70.

Van Eijk, R.L.J., Kohlrausch, A., Juola, J.F. & Van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Perception & Psychophysics, 70*, 955-968.

van Wassenhove, V., Buonomano, D.V., Shimojo, S. & Shams, L. (2008). Distortions of subjective time perception within and across senses. *PLoS ONE, 3*, e1437.

van Wassenhove, V. (2009). Minding time in an amodal representational space. *Philosophical Transactions of the Royal Society B, 364*, 1815-1830.

von Békésy, J. (1959). Similarities between hearing and skin sensations. *Psychological Review, 66*, 1-22.

Vroomen, J. & DeGelder, B. (2004). Perceptual effects of cross-modal stimulation. In G. Calvert, C. Spence & B.E. Stein (Eds). *Handbook of Multisensory Processes* (pp. 141-150). Massachusetts: MIT Press.

Walker, J.T & Scott. K.J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones and gaps. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 1327-1339.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88*, 638.

Zampini, M., Brown, T., Shore, D., Maravita, A., Röder, B. & Spence, C. (2005). Audiotactile temporal order judgments. *Acta Psychologica, 118*, 277-291.

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

Zampini, M., Shore, D.I. & Spence, C. (2003a). Audiovisual temporal order judgments. *Experimental Brain Research, 152,* 198-210.

Zampini, M., Shore, D.I. & Spence, C. (2003b). Multisensory temporal order judgments: the role of hemispheric redundancy. *International Journal of Psychophysiology, 50,* 165-180.

Zampini, M., & Spence, C. (2004). The role of auditory cues in modulating the perceived crispness and staleness of potato chips. *Journal of Sensory Studies, 19*, 347-363.

Chapter 2. Reciprocal Interferences between Audition and Touch in the Perception of Duration(Multisensory Research, in press)

DOI:10.1163/22134808-00002583

Reciprocal Interferences between Audition and Touch in the Perception of Duration

Lia Villanueva[1] and Massimiliano Zampini[1,2]

[1]CiMeC Center for the Mind/Brain Sciences, University of Trento, Corso Bettini 31, Rovereto (Trento), Italy; [2] Department of Psychology and Cognitive Science, University of Trento, Corso Bettini 31, Rovereto (Trento), Italy

---

[1] To whom correspondence should be addressed. E-mail: l.villanueva@unitn.it, massimiliano.zampini@unitn.it

Abstract

Audition and touch interact with one another and share a number of similarities, however, little is known about their interplay in the perception of temporal duration. The present study intended to investigate whether the temporal duration of an irrelevant auditory or tactile stimulus could modulate the perceived duration of a target stimulus presented in the other modality (i.e., tactile or auditory) adopting both a between-participants (Experiment 1) and a within-participants (Experiment 2) experimental design. In a two-alternative forced choice task, participants decided which of two events in a target modality was longer. The simultaneously presented distractor stimuli were presented either with a congruent or incongruent duration from the target. Results showed that both the auditory and tactile modalities affected duration judgments in the incongruent condition, decreasing performance in both experiments. Moreover, in Experiment 1, the tactile modality enhanced the perception of auditory stimuli in the congruent condition, but audition did not facilitate performance for the congruent condition in the tactile modality; this tactile enhancement of audition was not found in Experiment 2. To the best of our knowledge, this is the first study documenting audiotactile interactions in the perception of duration, and suggests that audition and touch might modulate one another in a more balanced manner in contrast to audiovisual pairings. The findings support previous evidence as to the shared links and reciprocal influences when audition and touch interact with one another.

Keywords

audition; touch; duration perception; multisensory integration

2.1 Introduction

Temporal perception is a fundamental aspect of human behavior and perception (see Friedman 1990; Meck et al., 2012, for reviews). Various aspects of the temporal flow of information can provide important cues for understanding the mechanisms of multisensory interaction. A large amount of studies have investigated the perception of temporal synchrony between different sensory modalities (Fujisaki & Nishida, 2005; Fujisaki & Nishida, 2009; Kitagawa et al., 2005; Morein-Zamir et al., 2003; Occelli et al., 2008; Repp & Penel, 2002; Zampini et al., 2003; Zampini et al., 2005b). Indeed, the perception of temporal synchrony is an important issue given that the maximal multisensory integration seems to happen when the constituent unimodal stimuli occur at approximately the same time or fall within the 'temporal window' of integration (Meredith et al., 1987; Spence & Squire, 2003). However, the ability of detecting the synchrony of multisensory inputs is relatively informative because it does not provide any cue about their temporal durations that might serve as another important basis for the perception of their coherence.

Perception of temporal duration is one of the 'elementary time experiences' (Pöppel, 1978) that can be estimated by our various senses. In particular, it requires the ability of determining the time that passes by. Temporal duration can be considered an amodal characteristic that can be determined independently in different modalities and, for this reason, can be estimated even in the absence of any temporal synchrony information (e.g., even if the multisensory stimuli are not presented simultaneously, their respective duration can be perceived). Focusing on the range of seconds or subseconds durations, research contrasting individual unimodal performance on temporal duration has revealed some intriguing differences across sensory modalities. Most of the studies on temporal duration have focused on the auditory and visual modalities, and report that auditory events are perceived as being longer than visual events that have the same objective duration (Penney et al., 2000; Walker & Scott, 1981). Moreover, audition is a more reliable sense than vision in duration discrimination tasks (Grondin et al., 1998; Grondin, 2003).

More often, studies using audiovisual combinations usually suggest that audition modulates the perception of the duration of visual events, and vision does not modify the perceived duration of auditory events (Chen & Yeh, 2009; Klink et al., 2011). In Chen and Yeh's (2009) study, participants had to compare the duration of an oddball stimulus (either auditory or visual, which was always accompanied by a distractor stimulus in the opposite modality) with a standard duration stimulus. Results showed an influence of the auditory modality as a distractor, expanding the perceived visual duration, but not the opposite (i.e., no influence of the visual distractor on the auditory perceived duration; Chen & Yeh, 2009). However, a study by van Wassenhove et al. (2008) showed that vision dominated the auditory modality in duration perception. One could explain these different results by considering the absence of controls that ensured matches in intensity, in the various studies (see van Wasshenove, 2009). In various experiments, Walker and Scott (1981) reported a vast dominance of audition over visual duration perception, except for one condition, where vision dominated audition when the intensity of the auditory stimulus was lower (Experiment 3).

A possible explanation for the dominance of a particular modality over another is that the sensory modality that provides the more reliable information is the one that modulates perception in the other sensory modality (Ernst & Banks, 2002; Welch & Warren, 1980). Hence, vision has the best spatial acuity, and so modulates audition when considering spatial tasks (e.g. stimuli localization; Bertelson & Radeau, 1981). By contrast, when individuals are involved in temporal judgments tasks (e.g. duration evaluation), they trust more on their auditory modality because it has much better temporal resolution than the visual (Grondin, 1993). However, Ortega et al. (2014) reported that when temporal discriminability was equal for audition and vision (or audition was weaker than vision), audition still dominated vision. Ortega et al. (2014) results seem to suggest that duration perception is strongly linked to auditory processing even when audition does not provide the more discriminable information.

The role of tactile information in multisensory studies on temporal duration perception has been less investigated. Tomassini et al. (2011) examined the effects of combining vision and touch in duration discrimination. In this type of comparisons, the authors found that duration reproduction in touch was more accurate and reliable (i.e. more close to the physical duration) than vision (Tomassini et al., 2011). Also, bimodal stimuli were no more accurate than unimodal reproductions, thus showing no benefits from a multisensory integration perspective. Rach and Diederich (2006) show another example of visuotactile combinations in duration perception. In their study, however, they only investigated the effects of tactile duration on visual duration perception in a visual detection task. Here, participants had to detect whenever a visual stimulus appeared, whilst ignoring tactile distractors. The authors found that facilitation or enhancement effects were greater when the durations of the tactile stimuli were shorter (e.g. 50 ms) than larger (500 ms), suggesting an inverse effectiveness effect when weak stimuli are presented (Rach & Diederich, 2006). These findings indicate the existence of facilitation effects under certain conditions where bimodal situations are introduced.

Temporal aspects of audiotactile interactions have been studied in the perception of synchrony (Kitagawa et al., 2005; Occelli et al., 2008; Zampini et al., 2005a), and highlight their better temporal resolution in contrast to audiovisual or visuotactile combinations when perceiving this temporal aspect (Fujisaki et al., 2012; Fujisaki & Nishida, 2009). However, to the best of our knowledge, there are no reported studies on temporal duration using audiotactile combinations. The study of this particular modality combination is interesting because these modalities share a great number of aspects; for example, they are sensitive to the same type of physical stimulation and seem to share a common evolutionary origin (Occelli et al., 2011; Soto-Faraco & Deco, 2009; Von Békésy, 1959). Also, there seems to be a synchronicity in the development of touch (being the first actual sense to develop) and audition in a prenatal stage, both developed prior to vision (Moore & Persaud, 2008), perhaps being a key for audiotactile connections to occur (Lickliter & Bahrick, 2000). In addition, Foxe et al. (2002), suggest that there are interactions between the auditory and the somatosensory cortex: information received from the vibratory frequencies from tactile stimulation is processed in regions that are proximate in the cerebral cortex (Ro et al., 2009). Therefore, the present study aimed to investigate the possible influence of the tactile and the auditory modalities over one another when perceiving duration. In particular, in a two-alternative forced choice paradigm, participants were presented with two target stimuli of the same sensory modality (either auditory or tactile) in succession and had to judge which one was longer. In bimodal trials, distractor stimuli in the non-target modality were also presented. Target and distractor stimuli could have the same duration (congruent condition) or different duration (incongruent condition). One group of participants performed the auditory duration task and the other group the tactile duration task. A between-participants design was adopted as a way to avoid a possible carry over effect arising from doing the two tasks with the two possible modalities as targets, a bias that could be manifested in confounding which target modality to attend, and which one to ignore. It was expected that the relation between the auditory modality and the tactile modality would be more symmetrical than audiovisual or visuotactile combinations (see e.g., Chen & Yeh, 2009; Penney et al., 2000; Tomassini et al., 2011; van Wassenhove et al., 2008; Walker & Scott, 1981), since audition and touch have shown to have a reasonable number of links (e.g., Occelli et al., 2011), along with experimental evidence showing these symmetrical relation in time perception, for example in event counting, where both modalities seem to impact the number of perceived events in the other modality (Bresciani & Ernst, 2007).

2.2 Experiment 1. Audiotactile duration task. Between-subjects design

2.2.1 Participants

A total of 30 participants from the University of Trento took part in the first experiment. 15 participants (11 F; mean age 27 years, range from 22 - 30 years) completed the ‘tactile target task’, while other 15 participants (7 F, mean age 26 years, range from 21 - 29 years) completed the ‘auditory target’ task. The number of participants was decided on the basis of the sample size adopted in previous studies addressing similar research questions (Romei et al., 2011). Both the experiments reported in the present study were conducted in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki (most recently amended in 2013, Fortaleza), as well as the ethical guidelines laid down by the University of Trento. All participants gave their informed consent prior to their inclusion in the study and took part on a voluntary basis, or to obtain course credits. Participants were not informed as to the purpose of the experiment, and all reported normal hearing and tactile sensitivity.

2.2.2 Apparatus and Stimuli

The tactile stimuli were delivered by means of an Oticon BC 461-1 (100 Ohm, Oticon, UK) bone conductor vibrator (1.4 cm X 2.4 cm). Auditory stimuli were delivered through headphones (Sennheiser HD 25-C11, Germany). White background external noise was sent to external loudspeakers (Dell A215, USA) located in a lower position from the monitor, parallel to the external corners of the computer used to deliver the task, and was controlled through an mp3 player (ipod Nano, Apple, USA). The experiment was delivered through a PC laptop (Dell Windows 7) using PsychoPhysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) for MATLAB R2012b. Tactile stimuli consisted on 200 Hz vibrations (0.085 N) delivered to the left hand's index finger. The tactile stimulator was placed on top of a foam rectangle located at a distance of 40 cm from the response keyboard. Auditory stimuli consisted of pure tones with a frequency of 200 Hz and a constant intensity of 50 dB (SPL). The intensities used were previously agreed between the experimenters as matching one another (touch and auditory intensities perceptually, that is, intensities that seemed to be equivalent). External white noise was presented to cover the sound made by the tactile stimulator, with an intensity of 91 dB (SPL) as measured by an audiometer, which still allowed listening to the auditory stimuli coming from the headphones as verbally reported by the participants during a practice session. Both tactile and auditory stimuli were presented in a fixed constant duration, that is, a standard duration, as well as in different variable durations, which were termed probe durations. The standard duration was of 90 ms, whereas the probe durations consisted of 50, 60, 80, 100, 120 and 130 ms each, so they could be either shorter or longer than the standard duration.

2.2.3 Design and Procedure

The design of the experiment consisted of a between-participants factor of Target (Auditory/Tactile) along with Stimulus presentation (Bimodal Congruent, Bimodal Incongruent and Unimodal) and Duration (50, 60, 80, 100, 120 and 130 ms) as within-participants factors. A 2AFC task was employed with one of the two possible modality targets. The task consisted on attending to a target modality (i.e. auditory or tactile) while trying to ignore duration information on the distractor modality (tactile or auditory, correspondingly) that could be either consistent or inconsistent with the duration of the target. Participants needed to determine which of the pair of presented tones (for the auditory as target condition), or vibrations (for the tactile as target condition) was the longest, either the first stimulus, by pressing 1, or the second stimulus, by pressing 2 on the response keyboard. The instructions were provided on the computer screen and were also further explained by the experimenter. Each trial was composed of a pair of events. Each event consisted of a simultaneous presentation of a stimulus in the target modality, and a stimulus in the distractor modality. Each trial could either have a unimodal or a bimodal presentation. For bimodal trials, both the target and distractor modality presented a fixed standard duration and a variable probe duration (detailed in the previous section). However, the manner or order in which the standard and probe duration were presented varied according to the experimental conditions. For the *bimodal congruent condition*, the target and distractor modality presented the same durations. For the *bimodal incongruent condition*, the events had the opposite duration information (i.e., when one of the target modality stimulus was presented in the standard duration, the distractor was presented in the probe duration, and viceversa for the other stimulus). The *unimodal condition* for each target modality was included to serve as a baseline for comparing the other conditions. Unimodal trials kept the same structure of events from the previous conditions (i.e., a standard and a probe duration stimulus). For a clearer view of the conditions, see Figure 3. The order of presentation of the target standard and probe durations was counterbalanced across the two events of a trial, for all the conditions (e.g., standard – probe, probe – standard). The factor Stimulus presentation was presented randomly. A total of 360 trials were presented, with 120 trials for each condition and duration. There were 20 trials for each duration. The experiment took around 50 minutes to be completed, and a practice session was always presented before starting the experiment, and was composed of 20 trials with all the stimuli and conditions included in the experiment. Three pauses with a fixed duration (3 min) were introduced among the experiment. Participants sat in the front of a laptop computer and had to wear headphones and place their left hand index finger resting over the tactile stimulator, attached by a velcro tape. The finger pressure was not controlled mechanically, however participants were instructed to always keep the finger in the same position indicated by the experimenter and to avoid making pressure over the tactile stimulator. Participants were tested separately in a dimly illuminated soundproof booth and were asked to keep their gaze at the fixation point in the center of the screen. A between-subjects design was adopted as a way to avoid possible practice effects, like having participants always responding to the same modality despite its role as target or distractor for example, since the aim was to focus only on the target modality whilst ignoring the distractor.

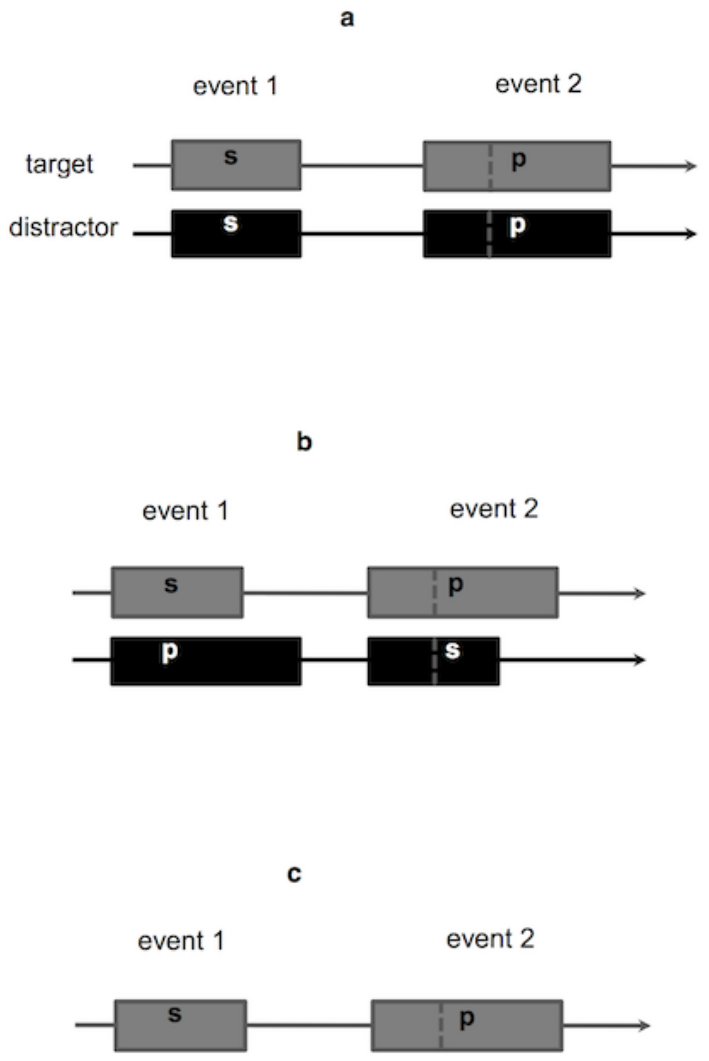


Figure 3 Trial structure examples across conditions. a) Trial structure example for the Bimodal Congruent Condition with the corresponding durations both in the target and distractor modality. b). Trial example structure for the Bimodal Incongruent Condition. c. Trial example structure for the Unimodal Condition. Rectangle bars indicate the durations; ‘s’ stands for standard duration, ‘p’ for probe duration. Dotted vertical lines along the probe durations indicate that durations could be either shorter or longer than the standard duration.

2.2.4 Results and Discussion

The proportion of correct responses was computed for each individual, according to the controlled experimental conditions, and was employed to perform a mixed ANOVA with TARGET (auditory/tactile) as the between-participants factor, and STIMULI PRESENTATION (bimodal congruent/bimodal incongruent/unimodal) and DURATION (50, 60, 80, 100, 120 and 130 ms) as the within-participants factors. The TARGET condition, provided no significant main effects, *F*(1, 28) = .05, NS, *p* = .812. Since sphericity assumption was violated, Greenhouse Geisser values are reported in some cases. A main effect of STIMULUS PRESENTATION was obtained *F* (1.22, 34.40) = 87.34, *p <.001* (Greenhouse Geisser corrected, X^2 = 26.70, *p* < .001). Posthoc pairwise comparisons revealed that the bimodal incongruent condition had the lowest performance (M = .628, SE =.024) compared to the bimodal congruent (M=.862, SE = .011, *p* <.001) and the unimodal conditions (M = .821, SE = .014, *p* <.001). Overall performance in the bimodal congruent condition was significantly greater than for the unimodal condition (*p* <.001). A main effect of DURATION was also found *F*(3.70, 101.85) = 71.60, *p* <.001 (Greenhouse Geisser corrected, X^2 = 26, *p* = .027). Posthoc pairwise comparisons revealed that durations of 50, 60, 120 and 130 ms (M = .853, SE = .018, M = .831, SE = .017, M = .805, SE = .017, M = .831, SE = .018 respectively) were not significantly different from one another, but however provided better overall performance, when compared to the range of durations of 80 and 100 ms, where performance is significantly lower (M = .645, SE = .016, M = .647, SE =.013 respectively, *p* <.001). The latter durations did not reveal differences between one another in terms of accuracy. Because of their proximity to the standard duration (90 ms), participants made more errors for this range of durations since it was hard to differentiate them from the standard referent. The interaction STIMULUS PRESENTATION x DURATION was significant *F*(5.60, 166.70) = 5.50, *p* <.001 (Greenhouse Geisser corrected, X^2 = 92.60, *p* = .001). The interaction STIMULUS PRESENTATION x TARGET did not prove significant, *F*(2, 28) = 2.00, NS, *p* = .144, neither the interaction DURATION x TARGET *F*(5, 28) = 1.17, NS, *p* = .326. A 3-way interaction of

STIMULUS PRESENTATION x DURATION x TARGET was found, *F*(10, 28) = 2.05, *p* = .029. To facilitate data interpretation, the interaction was split for further analysis in the following way: a within-participants ANOVA was performed for each target (auditory, tactile) with STIMULUS PRESENTATION and DURATION as the within-participants factors. For data simplification purposes, we only report here the main effect that we consider accounts for the found differences. The ANOVA for the tactile target group of participants revealed a main effect of STIMULUS PRESENTATION *F*(1.15, 16.19) = 45.33, *p* <.001. Posthoc pairwise comparisons revealed that performance in the bimodal incongruent condition (M = .613, SE = .037) was significantly lower than the bimodal congruent condition (M = .863, SE = .016, *p* <.001) and unimodal (M = .844, SE = .018, *p* <.001), however bimodal congruent and unimodal conditions were not different from one another in terms of accuracy. A main effect of STIMULUS PRESENTATION was also found in the ANOVA run for the auditory target group of participants *F*(1.33, 18.68) = 43.70, *p* <.001. Here, the accuracy for the bimodal incongruent (M = .642, SE = .029) was smaller than the bimodal congruent condition (M = .861, SE = .014, *p* <.001), and unimodal condition (M = .798, SE = .022, *p* <.001). Also, unlike the tactile target condition, the accuracy for the bimodal congruent condition was higher than the unimodal condition (*p* = .001) in the auditory target task. This difference could be partly driving the 3-way interaction. All posthoc pairwise comparisons were made using Bonferroni adjustments. For a broader view of the results, see Figure 4 for performance by stimulus presentation across the target modalities.

The results obtained from Experiment 1 reveal a significant decrement in performance for the incongruent conditions, suggesting that the conflicting information on stimuli durations presented in the distractor modalities, (either tactile or auditory), interfered with the participants' ability to determine target stimuli durations. Moreover, since both groups of participants were similarly influenced by the different distractor modalities in the incongruent condition (*see Fig 2*), it is reasonable to conclude that, both touch and audition seem to affect one another in the same way in this particular duration task. The bimodal congruent condition provided the best performance, and was significantly better than for the unimodal condition. This finding is interesting, suggesting perhaps an overall facilitation effect when coherent duration information from both the target and the distractors is presented, improving performance. The result is in line with multisensory integration principles, which suggest that when redundant information from two sensory inputs is given, performance becomes greater than with unimodal presentations (Forster et al., 2002; Lovelace et al., 2003; Stein et al., 1996). Previous evidence has highlighted the superiority of the auditory modality over the visual when performing unimodal temporal judgments, where auditory temporal discrimination has proven to be more accurate and reliable (Grondin et al., 1998; Repp & Penel, 2002). Although the role of the tactile modality in multisensory temporal duration tasks has been less studied, it has been shown that the tactile modality might be better than the visual modality (Tommassini et al., 2011). To the best of our knowledge, this is the first study comparing temporal duration accuracy between audition and touch. The fact that there were no differences in performance between the auditory and the tactile modality in the present study, seems to support the idea that touch and audition are both better than vision (see e.g., Grondin et al, 1998; Repp & Penel, 2002; Tommasini et al., 2011) and might share similar mechanisms in the temporal domain.

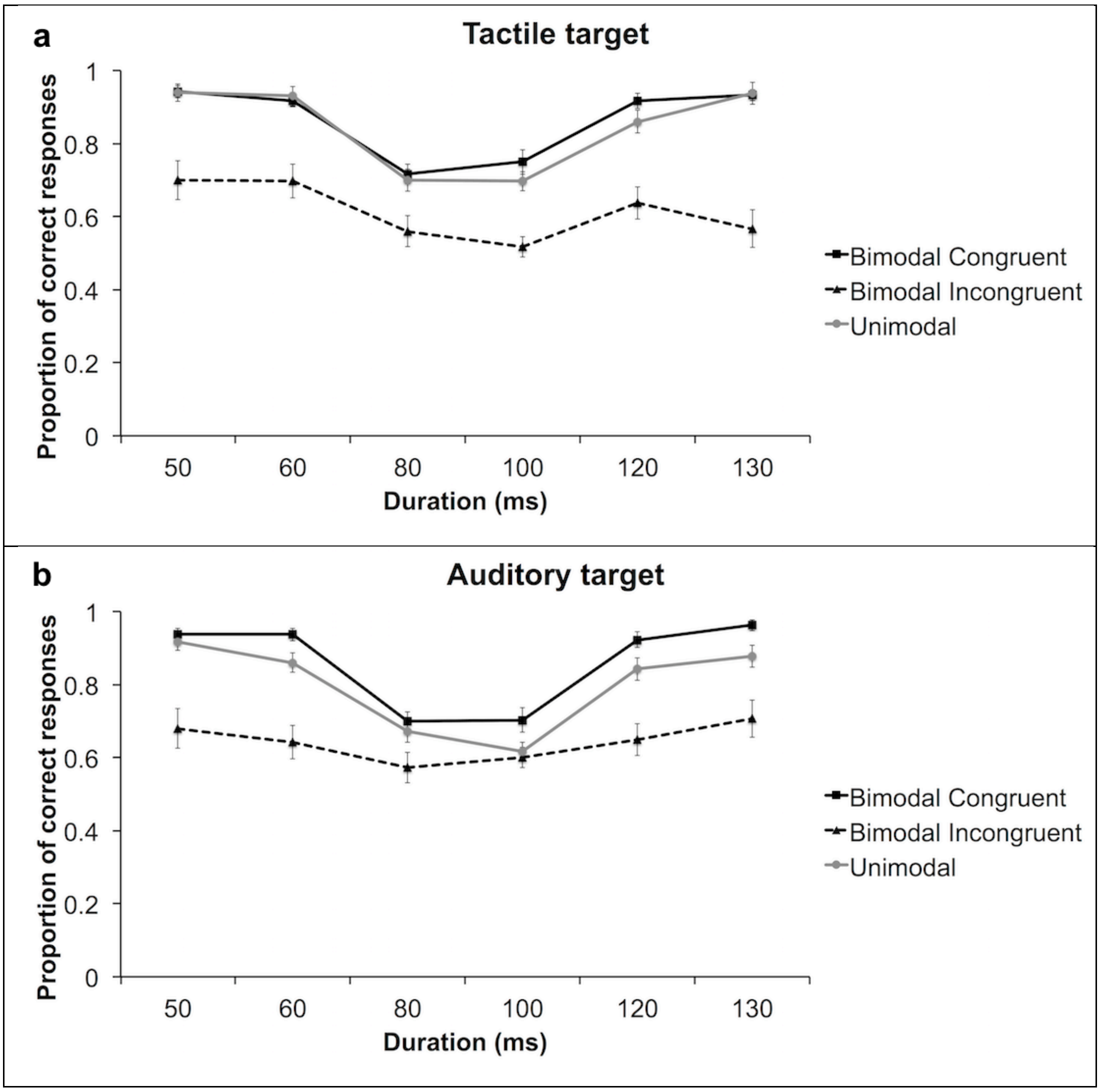


Figure 4 Performance by stimulus presentation and Duration across the a) tactile and b) auditory target modalities in Experiment 1. Lines indicate mean values of proportion of correct responses. Error bars illustrate the standard errors. Black line corresponds to the Bimodal Congruent condition; dotted line indicates the Bimodal Incongruent condition; gray line corresponds to the Unimodal condition.

Finally, another interesting finding emerged when splitting the 3-way interaction. In particular, participants' performance for the auditory target group was better in the congruent condition than in the unimodal condition, suggesting perhaps a facilitation effect. Conversely, auditory distractor information did not seem to improve performance for the tactile target congruent condition in contrast to the unimodal condition. This difference in enhancement is interesting, since previous evidence has shown that audition can act as an enhancer in discrimination sensitivity, at least in audiovisual duration judgments, when the visual stimulus is the target and duration information between vision and audition is coherent (e.g., Romei et al., 2011). A possible explanation for this difference could be also a result of having different participants carrying out a different target task. For example, there could have been a relative perceptual 'preference or easiness' to perceive tactile stimuli of the group performing the auditory target task, in comparison to the group performing the tactile target task, the latter being less affected or susceptible to auditory distractors. With this pattern of results, we considered important to verify whether part of the obtained data was not a result of individual differences in performance, therefore, a within-subjects design was introduced in Experiment 2, as a way to control the possible differences among individuals when performing both the auditory and tactile as target tasks.

2.3 Experiment 2. Audiotactile duration task. Within-subjects design.

2.3.1 Participants

A total of 20 participants from the University of Trento participated in the second experiment in exchange for course credit. Age range was 20-25 years old (mean age 21.6, 15 F). The number of participants was decided on the basis of the sample size adopted in previous studies addressing similar research questions (Romei et al., 2011). Order of task completion was counterbalanced across participants; 10 completed the auditory-then-tactile target order, while other 10 did the tactile-then-auditory tasks. All participants gave written informed consent and were naïve to the aim of the experiment. All reported normal hearing and tactile sensitivity.

2.3.2 Apparatus & Stimuli

The apparatus and stimuli employed for this experiment were the same as the ones used in Experiment 1.

2.3.3 Design and Procedure

The task design and procedure was identical to the first experiment, however, participants completed each different target modality task with one week separation from each other, in order to avoid long sessions and fatigue effects, as well as to minimize possible carry over effects.

2.3.4 Results and Discussion

The proportion of correct responses was obtained for each participant in relation to the experimental conditions employed. These values were used to perform a repeated measures ANOVA with TARGET (tactile/auditory), STIMULUS PRESENTATION (bimodal congruent/bimodal incongruent/unimodal) and DURATION (50, 60, 80, 100, 120 and 130 ms) as the within-subjects factors. Results indicated that the factor TARGET was not significant $F(1, 19) = 2.88$, NS, $p = .106$. A main effect of STIMULUS PRESENTATION was observed $F(2, 38) = 187.70$, $p <.001$. Posthoc pairwise comparisons revealed that performance in the bimodal incongruent condition was significantly lower (M = .591, SE = .018) and smaller than the bimodal congruent (M = .820, SE = .017, $p <.001$) and unimodal conditions (M = .805, SE = .018, $p <.001$); however the latter two were not statistically different from one another. For a clearer description of the results, see Figure 5. The factor DURATION showed a main effect $F(5, 95) = 81. 88$, $p <.001$, and posthoc pairwise comparisons revealed that the shortest duration, 50 ms, provided the best performance (M = .840, SE =.018) followed by the other range of durations 60, 120 and130 ms (M = .802, SE =.019, $p = .003$, M = .760, SE = .021, $p = .001$, M =

.779, SE = .023, p = .005 respectively), which did not differ from one another but were significantly different from durations 80 and 100 ms (M = .642, SE = .011, p <.001, M = .608, SE = .016, p <.001).The latter two provided the worst performance and were not statistically different from one another. The low performance in 80 and 100 ms durations can be explained by the difficulty of differentiating these durations from the standard duration of 90 ms. In some cases, sphericity assumption was not met, so Greenhouse Geisser values are reported. The interaction STIMULUS PRESENTATION x DURATION $F(5.25, 100)$ = 8.78, p <.001 (Greenhouse Geisser corrected, X^2 = 87.33, p = .004) was found significant, as well as the interaction TARGET x DURATION $F(5, 95)$ = 2.47, p =.037. The interactions TARGET x STIMULUS PRESENTATION $F(1.26, 24)$ = 1.29, NS, p = .286 (Greenhouse Geisser corrected, X^2 = 16, p < .001) did not show significance. The 3-way interaction of STIMULUS PRESENTATION x DURATION x TARGET was not significant either F(4.60, 87.56) = 1.34, NS, p = .211 (Greenhouse Geisser corrected, X^2 = 103.39, p < .001). All post hoc pairwise comparisons were made using Bonferroni adjustments.

The results found in Experiment 2 confirm the decrement of performance in the incongruent condition illustrated in Experiment 1. When both audition and touch acted as distractors, they affected performance in the same manner in this condition, decrementing the accuracy of responses. Both modalities were equivalently influential to one another (i.e. to the corresponding target), as there were no significant differences according to the target factor. However, unlike Experiment 1, there were no effects of facilitation in the congruent condition. One of the reasons for this difference could be that the within-participants manipulation did serve here to control for differences in the sensitivities to the distractor modalities, and in this case the participants showed a more uniform or less sensitivity to any of the distractor's influence. Perhaps the facilitation found in Experiment 1 was driven by the effective influence of touch on audition of that specific group of participants in the congruent condition when performing the auditory task. Nevertheless, despite these differences in sensitivities, the effect of the incongruent condition was still strong enough to disrupt performance in a balanced manner on both targets.

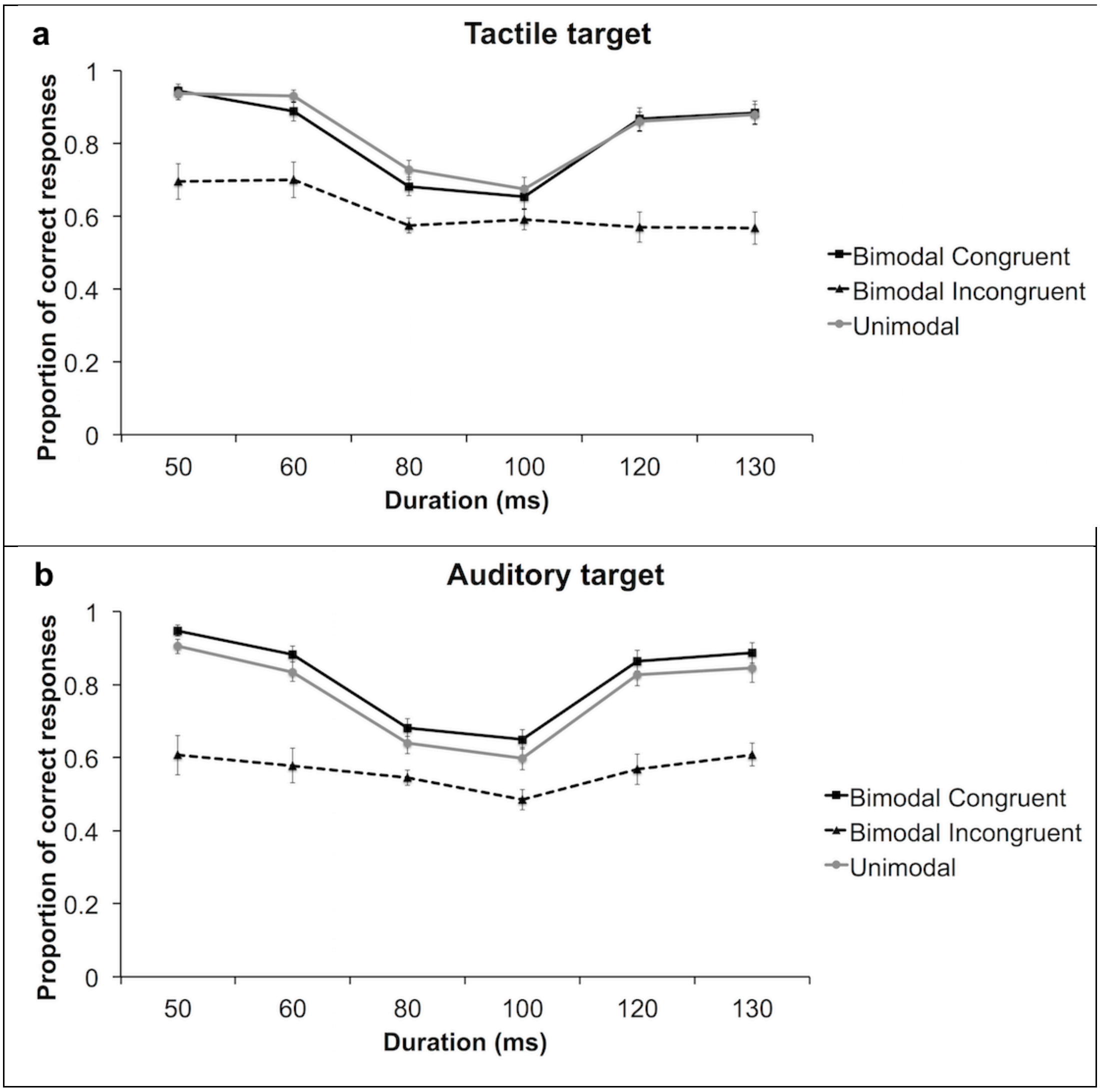


Figure 5.Performance by stimulus presentation and Duration across the a) tactile and b) auditory target modalities in Experiment 2. Lines indicate mean values of proportion of correct responses. Error bars illustrate the standard errors. Black line corresponds to the Bimodal Congruent condition; dotted line indicates the Bimodal Incongruent condition; gray line corresponds to the Unimodal condition.

2.4 General discussion

The present study explored the influences of the modalities of touch and audition over one another in a 2AFC duration task. Experiment 1 explored this influence using a between-participants design; each group of participants performed a different target modality task, with the aim of avoiding a possible bias for executing both tasks with the two modalities as targets (i.e., confounding which target modality to attend). Experiment 2 explored this influence but in this case, using a within-subjects design in order to replicate the findings from Experiment 1, in a design that would allow discarding any possible influence of individual differences in performance according to the target modalities at play. Overall findings from Experiments 1 and 2 indicated that both modalities can interfere with performance when the duration information is incongruent or different between the target and distractor modalities. Experiment 1 showed a facilitation effect when participants performed the auditory as target task. The facilitation effect consisted of an overall better performance in the bimodal congruent condition, in contrast to the unimodal condition. This finding suggests a possible help or integration coming from the tactile modality; however, a facilitation effect in the auditory target condition was not found in Experiment 2.

To the best of our knowledge, this study is the first one addressing the question of how audition and touch interact with one another in the perception of duration. Experiments 1 and 2 showed that the overall performance in both of the tasks (tactile target/ auditory target) did not change significantly according to the distractor modalities. This finding could suggest that there are no differences in terms of dominance of one modality over the other. Other modality combinations (e.g. audiovisual) addressing the study of duration perception, have shown that it is usually the auditory modality the most influential when presenting conflicting duration information (Chen & Yeh, 2009; Klink et al., 2011; Walker & Scott, 1981). In the present study, data suggests that both audition and touch seem to influence one another in a balanced manner.

Both experiments showed a main effect of stimulus presentation, providing additional evidence to the previous literature that suggests that audition and touch influence one another (Bresciani & Ernst, 2007; Occelli et al.,2009; for a review on audiotactile interactions, see Occelli et al., 2011), and that these two modalities can disrupt the perception of duration information over one another in the same manner, since the direction of the found incongruency effects follows the same pattern in the two experiments (i.e. a poor, disrupted performance in the incongruent bimodal condition) for both modalities. The incongruency effect seems to be a robust finding, since it is replicated in both experiments. This phenomenon also highlights the difficulty of separating or disentangling the information in both modalities, despite the explicit request to ignore information in the distractor modality.

The similarity on the obtained effects could also stem from the fact that audition and touch are closely linked and share some similarities (Occelli et al., 2011; von Békésy, 1959), therefore they might as well modulate one another in a similar way. Although these modalities may provide different qualitative experiences, they both respond to the same type of physical stimulation, they have a common evolutionary origin and present interconnections, as there are interactions between the auditory and the somatosensory cortex (Foxe et al., 2002; Occelli et al., 2011; von Békésy, 1959).

An effect that could perhaps represent a facilitation phenomenon was seen only for the auditory as target task in Experiment 1; this effect was not seen in Experiment 2. It is worthwhile asking why touch facilitated audition in Experiment 1 and not the opposite. A possibility that could account for this influence and that could also explain the overall pattern of results could likewise lay in a methodological issue related to the intensities employed to present the auditory and tactile stimuli. The intensities used were previously agreed between the experimenters as matching one another (touch and auditory intensities). However, it could have been the case that the tactile intensity employed was perceived as more intense or salient in general, and therefore impacted the results, seeing no facilitation of the auditory modality over the tactile (since the former could have been relatively weak), etc. It becomes difficult to establish similarities in intensities of stimuli of different nature and this has been a debatable and an unsolved question (Spence, 2011), nevertheless, Ortega et al. (2014) show there is an ability to match intersensory intensities, as well as agreement between participants, for example in an audiovisual setting. Further work could try to establish these cross-modal matchings' of intensities from auditory and tactile stimuli, as well as explore the role of changing the intensities of the distractor modalities. This would allow observing whether the same patterns of sensory modulation are kept in the interaction of these particular modalities in a duration perception task, or whether they change as a function of the saliency of the distractor modality. A speculative explanation for the differences in the facilitation effect found only in Experiment 1 could be that participants in the tactile target group, were more 'tactile sensible' and therefore, more susceptible to tactile influences on the perception of auditory stimuli, in a sense that touch *helped* audition. The impact of individual differences and possible inhomogeneity across the groups, however, were controlled in Experiment 2 in the within-participants design, and could explain why these effects and differences were not found here. Perhaps a future study could be implemented in order to control for the participants' auditory and tactile sensitivities, by checking individual thresholds first.

The findings of a mutual reciprocal influence between audition and touch in this particular temporal duration task could be in line with the central clock model, which operates supramodally (Grondin, 2010; Levitan et al., 2015; Treisman, 1963), thus suggesting the existence of a central timing mechanism that is shared by the modalities, allowing at some point some sort of modulation or temporal crosstalk between the senses. Perhaps this can explain why these two modalities influenced one another in the same manner on this particular set of durations. There is however, some debate regarding time perception models. Some suggest the existence of modality specific clocks (Morrone et al., 2005) that explain differences in perceived durations by the different senses (Grondin, 2010), nevertheless it is still unclear how these modality specific clocks account for evidence of temporal crosstalk or modulation of time perception between the senses.

Another question rises regarding the nature of the final percept or judgment made by participants. Do they really perceive the duration of the targets as longer or shorter in accordance to the distractor? If so, this could imply a real multisensory integration phenomenon. Or is this percept the result of a mere attentional bias or saliency effect, where participants pay more attention to the distractor modality and only respond according to these distractor stimuli? This question is frequent in this kind of paradigms and remains unsolved. A possible way to explore the extent of the distractor's influence could be by using aftereffect paradigms (Recanzone, 2002), where the effect of the distractor could be seen in more detail, for example, if the duration of the target stimulus is 'expanded or shortened' even after the distractor stimulus is not present. In conclusion, our results suggest that tactile information is also a reliable estimator for duration perception, and could act in a similar way as audition, in time perception. It also shows that touch can influence audition too, a modality that has traditionally been found to be the most reliable one for perceiving time.

Acknowledgments

We would like to thank Irene Susini for her kind help in collecting data for Experiment 2. This project (208372) was funded by the National Council for Science and Technology, Mexico.

References

Bertelson, P. & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics, 29*, 578–584.

Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Bresciani, J.P. & Ernst, M.O. (2007). Signal reliability modulates auditory-tactile integration for event counting. *Neuroreport, 18*, 1157–1161.

Chen, K.M. & Yeh, S.L. (2009). Asymmetric cross-modal effects in time perception. *Acta Psychologica, 130*, 225–234.

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

Forster, B., Cavina-Pratesi,C., Aglioti, S.M. & G. Berlucchi, G.(2002). Redundant target effect and intersensory facilitation from visual–tactile interactions in simple reaction time. *Experimental Brain Research, 143*, 480–487.

Foxe, J. J., Wylie, G. R., Martinez, A., Schroeder, C. E., Javitt, D. C., Guilfoyle, D., Ritter, W., & Murray, M. M. (2002). Auditory-somatosensory multisensory processing in auditory association cortex: An fMRI study. *Journal of Neurophysiology, 88*, 540–543.

Friedman, W. J. 1990. *About time: Inventing the Fourth Dimension*. MIT Press, Cambridge, MA., USA.

Fujisaki, W., Kitazawa, S. & Nishida, S. (2012). Multisensory Timing. In B.E Stein (Ed.), *The New Handbook of Multisensory Processes*, pp. 302-317, MIT Press: Massachusetts, USA.

Fujisaki, W. & Nishida, S. (2005). Temporal frequency characteristics of synchrony asynchrony discrimination of audio-visual signals. *Experimental Brain Research, 166*, 455–464.

Fujisaki, W. & Nishida, S. (2009). Audio-tactile superiority over visuo-tactile and audio-visual combinations in the temporal resolution of synchrony perception. *Experimental Brain Research, 198*, 245–259.

Grondin, S. (1993). Duration discrimination of empty and filled intervals marked by auditory and visual signals. *Perception & Psychophysics, 54,* 383–394.

Grondin, S. (2003). Sensory modalities and the temporal processing . In H.Helfrich (Ed.), *Time and Mind II*, pp.61–77, Hogrefe & Huber, Göttingen, Germany.

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception & Psychophysics, 72*, 561–582.

Grondin, S., Meilleur-Wells, G., Ouellette, C., & Macar, F. (1998). Sensory effects on judgments of short time-intervals. *Psychological Research, 61*, 261–268.

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

Kleiner, M., Brainard, D., & Pelli, D.(2007). "What's new in Psychtoolbox-3?" *Perception, 36,* 1–16.

Klink, P.C., Montijn, J.S. & van Wezel, R.J.A. (2011). Crossmodal duration perception involves perceptual grouping, temporal ventriloquism, and variable internal clock rates. *Attention, Perception & Psychophysics, 73,* 219–236.

Levitan, C.A., Ban, Y.H.A., Stiles, N.R.B. & Shimojo, S. (2015). Rate perception adapts across the senses: evidence for a unified timing mechanism. *Scientific Reports, 5*, 1–5.

Lickliter, R., & Bahrick, L. E. (2000). The development of infant intersensory perception: Advantages of a comparative convergent-operations approach. *Psychological Bulletin, 126*, 260–280.

Lovelace, C.T., Stein, B.E. & Wallace, M.T. (2003). An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Cognitive Brain Research, 17*, 447–453.

Meck, W. H., Doyère, V., & Gruart A., (2012) Interval Timing and Time-Based Decision Making. *Frontiers in Integrative Neuroscience, 6*, 13.

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience, 10*, 3215–3229.

Moore, K. L., & Persaud, T. V. N. (2008). *The Developing Human: Clinically Oriented Embryology*. Saunders, Philadelphia, USA.

Morein-Zamir, S., Soto-Faraco, S. & Kingstone, A. (2003). Auditory capture of vision : examining temporal ventriloquism. *Cognitive Brain Research, 17*, 154–163.

Morrone, M.C., Ross, J. & Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. *Nature Neuroscience, 8*, 950–954.

Occelli, V., Spence, C. & Zampini, M. (2008). Audiotactile temporal order judgments in sighted and blind individuals. *Neuropsychologia, 46*, 2845–2850.

Occelli, V., Spence, C. & Zampini, M. (2009). The effect of sound intensity on the audiotactile crossmodal dynamic capture effect. *Experimental Brain Research, 193*, 409–419.

Occelli, V., Spence, C., & Zampini, M. (2011). Audiotactile interactions in temporal perception. *Psychonomic Bulletin & Review, 18*, 429–554.

Ortega, L., Guzman-Martinez, E, Grabowecky, M & Suzuki, S. (2014). Audition dominates vision in duration perception irrespective of salience attention and temporal discriminability. *Attention, Perception & Psychophysics, 76,* 1485–1502.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Penney, T.B., Gibbon, J., & Meck, W.H. (2000). Differential effects of auditory and visual signals on clock speed and temporal memory. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1770–1787.

Pöppel, E. (1978). Time Perception. In R. Held, H. W. Leibowitz, & W. L. Teuber (Eds.), *Handbook of Sensory Physiology, Vol. 8.,* pp. 713–729, Springer-Verlag, Berlin, Germany.

Rach, S. & Diederich, A. (2006). Visual-tactile integration: does stimulus duration influence the relative amount of response enhancement? *Experimental Brain Research, 173*, 514–520.

Recanzone, G.H. (2002). Auditory influences on visual temporal rate perception. *Journal of Neurophysiology, 89***,** 1078–1093.

Repp, B.H. & Penel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology. Human Perception and Performance, 28*, 1085–1099.

Ro, T., Hsu, J., Yasar, N. E., Caitlin Elmore, L., & Beauchamp, M. S. (2009). Sound enhances touch perception. *Experimental Brain Research, 195*, 135–143.

Romei, V., De Haas, B., Mok, R.R. & Driver, J. (2011) Auditory stimulus timing influences perceived duration of co-occurring visual stimuli. *Frontiers in Psychology, 2,*1–8.

Soto-Faraco, S. & Deco, G. (2009). Multisensory contributions to the perception of vibrotactile events. *Behavioural Brain Research, 196*, 145–154.

Spence, C. (2011). Cross modal correspondences: a tutorial review. *Attention, Perception & Psychophysics, 73*, 971–995.

Spence, C., & Squire, S. (2003). Multisensory integration: Maintaining the perception of synchrony. *Current Biology, 13*, 519–521.

Stein., B.E., London, N., Wilkinson, L.K. & Price, D.D. (1996). Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. *Journal of Cognitive Neuroscience, 8*, 497–506.

Tomassini, A., Gori, M., Burr, D., Sandini, G. & Morrone, M.C. (2011). Perceived duration of visual and tactile stimuli depends on perceived speed. *Frontiers in Integrative Neuroscience, 5,* 1–8.

Treisman, M. (1963). Temporal indiscrimination and the indifference interval: Implications for a model of the 'internal clock'. *Psychological Monographs, 77,* 1-31.

van Wassenhove, V. (2009). Minding time in an amodal representational space. *Philosophical Transactions, 364,* 1815-1830.

van Wassenhove, V., Buonomano, D.V., Shimojo, S. & Shams, L. (2008). Distortions of subjective time perception within and across senses. *PLoS ONE,* 3, e143.

Von Békésy, J. (1959). Similarities between hearing and skin sensations. *Psychological Review,* 66, 1–22.

Walker, J.T. & Scott, K.J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones and gaps. *Journal of Experimental Psychology. Human Perception and Performance, 7,* 1327–1339.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88*, 638–667.

Zampini, M., Brown, T., Shore, D., Maravita, A., Röder, B. & Spence, C. (2005a). Audiotactile temporal order judgments. *Acta Psychologica, 118*, 277–291.

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

Zampini, M., Shore, D.I. & Spence, C. (2003). Multisensory temporal order judgments: the role of hemispheric redundancy. *International Journal of Psychophysiology, 50*, 165–180.

Chapter 3. Exploring the role of changing the intensities of the distractor modalities in an audiotactile duration paradigm

In the previously reported experiments (from chapter 2), the intensity of the distractor and target stimuli of the different modalities were not controlled. In order to see whether the intensity of the distractors might play a role in the audiotactile interactions of a duration task, it seemed necessary to conduct further experiments. Therefore, the following chapter addresses the study of audiotactile interactions in the perception of duration, when changing the intensity of the distractor stimuli. The following section introduces some of the studies that have previously explored the effects of the intensities of the stimuli from different modalities on situations that involve multisensory integration.

Different degrees of intensity from stimuli of bimodal situations have been shown to exert an impact in multisensory integration. Starting with the idea that two stimuli from different modalities are better than one, in terms of faster reaction times (known as the ‘intersensory facilitation’ effect; Diederich, Colonius, Bockhorst & Tabeling, 2003), it has also been shown that weak intensities from bimodal stimuli generate facilitation and enhancement effects in multisensory integration, (principle known as the inverse effectiveness rule; Stein & Meredith, 1993).

Diederich and Colonius (2004) report a paradigm that tested the intensity of different modality stimuli in order to check for the inverse effectiveness rule. The task consisted on the presentation of unimodal, bimodal and trimodal stimuli of visual, auditory and tactile stimuli, which varied in intensities (three different intensities were employed, and only changed for the auditory and tactile stimuli, whilst the visual modality kept a constant intensity). Additionally, taking into account the different processing times required for each sensory channel to reach activation, the timing between the different modalities’ stimuli was varied (and could be either the same timing or variable), as a way to allow an overlap among the modalities. In the task, participants had to respond whenever they perceived a stimulus. The findings indicated that reaction times for trimodal stimuli were the faster, followed by bimodal stimuli and lastly, unimodal stimuli. Also, the unimodal stimuli with higher intensities presented faster reaction times than stimuli with lower intensities (Diederich & Colonius, 2004). More interestingly, the stimuli that presented the best facilitation enhancement were the ones with 'low-low' intensities in bimodal situations; additionally, a 'constant - low intensity' pair, as well as a 'high-low' intensity pair was better than a 'high-high' intensity pair (Diederich & Colonius, 2004).

In a further study, Rach and Diederich (2006) studied the enhancement effects of duration (which was either short or long), were duration was considered as some *sort* of intensity factor in the task for visuo-tactile pairings. In the task, participants needed to react to visual stimuli, whilst ignoring the tactile ones. The findings indicated that the shorter stimulus pairings (i.e. in this case, the authors considered the short duration as the less intense) provided greater facilitation (i.e. shorter saccadic reaction times) than the longer stimulus, effect that the authors termed 'inverse effectiveness of stimulus duration' (Rach & Diederich, 2006).

There are studies that have used the inverse effectiveness rule to study the intersensory facilitation effect, and have however, shown that not always the 'weak-weak' bimodal combinations yield faster reaction times. Bernstein, Chu, Briggs and Schurman (1973) found that a weak target and an intense non-target revealed more facilitation effects (as cited in Rach & Diederich, 2006, p.515). The reason behind this is explained by Colonius and Diederich (2004), whom reflect on the fact that maybe a stronger non-target could increment or trigger a response, and a very intense target would impede the influence of the non–target, since the former previously attained the necessary response activation (Colonius & Diederich, 2004). So perhaps a weak target could better take advantage of the effectiveness of the strong non-target (Rach & Diederich, 2006).

Some studies have expanded on the effects of the intensity of stimuli from different modalities, in order to understand its relation and effects on sensory dominance in the perception of duration and event counting. For example, Walker and Scott (1981) report an auditory dominance effect in audiovisual pairings in the perception of duration, employing different intensities for the auditory stimuli. The authors adopted a duration reproduction task, and presented unimodal durations (auditory/visual) along with bimodal pairings. The findings indicated that the reproduction of the bimodally perceived durations were always closer to the auditory unimodal durations, than to the visual. This effect happened when the auditory intensity of the stimulus was higher in Experiment 1. Conversely, in Experiment 3, the intensity of the auditory stimulus was lower and a visual dominance was found, but only for the shortest duration of the experimental set of durations (Walker & Scott, 1981). A problem with this study is that the intensities were not tested initially as to make sure they were really perceived as one being more intense than the other.

The study of Ortega, Guzman-Martinez, Grabowecky and Suzuki (2014) also addressed the role that the saliency of auditory stimuli could play in the perception of duration, when presenting audiovisual pairings. The authors were interested to find whether the reduction or increment of the intensity of an auditory stimulus could alter the visual duration percept, for example, by making it seem to last longer than the actual visual duration (auditory dominance). They employed a series of temporal bisection tasks, where the intensity of an auditory stimulus was set to a lower, matching intensity, or higher intensity than the visual stimulus in bimodal presentations (Experiment 1). The task consisted on a reference phase that helped participants experience the short and long reference durations (which could be only visual, or only auditory), as well as a test phase, where participants decided whether the presented stimulus (visual/auditory/audiovisual with changing intensities) was close to the short or long reference. The different auditory intensities were defined by a previous calibration task where the proportion of auditory more intense responses was fitted to a psychometric function to obtain the matching intensity (PSE) in relation to the visual, along with the indexes of lower and higher intensities, as perceived by all the participants the 100% of times (Ortega et al., 2014). This result itself is interesting since it demonstrates that participants are able to compare stimuli from different modalities, and establish a hierarchy in intensities among stimuli that are totally different in nature.

The results by Ortega et al., 2014 showed that audition dominated the perception of duration, even in the cases where its intensity was weaker (and therefore, vision more salient). In sum, the duration of the audiovisual events from the test phase stimuli were perceived to last according to the duration of the auditory unimodal stimuli from the test stage. These findings were replicated in different tasks where the auditory test stimuli with low intensities were longer than the visual test stimuli and presented in *asynchrony* (Experiment 2a, 2b), as well as with *spatial separation* from the visual stimuli (Experiment 2c), factors that the authors established as selective attention (Ortega et al., 2014).

In other kind of tasks such as event counting, Bresciani and Ernst (2007) controlled the intensities of auditory stimuli (i.e. loud and quiet) in order to test the effects of this variable in the perceived number of events in audiotactile bimodal presentations. In this task, audition and touch served both as distractors and targets, correspondingly. The results showed that the 'quiet' auditory beeps resulted in a weaker effect of this modality over the tactile, as compared to the loud beeps, which indeed altered in a greater fashion the perceived number of tactile taps (Bresciani & Ernst, 2007).

To the best of my knowledge there are no reported studies addressing the impact of the different intensities in the perception of duration for audiotactile pairings.

The experiments reported in chapter two demonstrated an influence of audition over touch, and touch over audition, in a condition where duration information was inconsistent between the two modalities, decrementing accuracy or performance. This might suggest a modulatory effect of both modalities over one another; however, with the employed paradigm it is hard to rule out the fact that participants could have attended more to the distractor modality than the target modality. Therefore, the motivation behind the next experiments was to test whether the intensity of the distractor stimuli play a role in the interactions of audition and touch in the same duration paradigm from chapter 2. This exploration allows to check whether the same interference effects are still present when the intensities between the distractor and target are clearly matched, and clearly contrasting from one another (high/low intensities). This was explored using different approaches described in the following experiments.

3.1 Experiment 1. Varying the intensity of the distractor stimuli in an audiotactile duration paradigm.

In order to define which intensities to employ in the duration paradigm, it was first considered necessary to design a task in which participants previously rated different intensities available from the auditory and tactile modalities, to later use these individual values to set the distractor's intensities in the audiotactile duration paradigm. Therefore, a cross-modal matching task was implemented.

3.1.1 Cross-modal matching task

Here, a series of different intensities for the auditory and the tactile modalities were obtained to test them against a constant referent from the opposite modality. The aim was to obtain values or indexes to get low, medium and high intensities for the two modalities.

3.1.1.1 Participants

A total of 18 participants from the University of Trento collaborated in this task (12F, mean age 24, age range from19- 36 years old), in exchange for course credit or reimbursement. All participants gave their informed consent prior to the experiment and were not informed as to the purpose of the experiment. All reported normal hearing and tactile sensitivity. All the experiments reported in this chapter were conducted in accordance with the ethical standards established by the 1964 Declaration of Helsinki, as well as the ethical guidelines from the University of Trento.

3.1.1.2 Apparatus and stimuli

A total of 11 intensities were employed for each modality to be compared to the constant opposite modality. The stimuli with these variable intensities were generated through Matlab R2012b, with different ranges of intensities for both modalities, including for example: 46 dB, 50 dB, 58 dB, 71.6 dB (SPL) (for the auditory modality), .45Vpp, 4.40 Vpp, 12Vpp and 15 Vpp (Volt pico pico) (for the tactile modality).

These values were compared to a constant auditory of 50 dB, when the varying intensities were the tactile, and a constant tactile intensity of 12 Vpp when the auditory intensities were the changing ones.

The auditory stimuli were delivered via headphones (Sennheiser HD 25-C11, Germany), and the tactile stimuli were delivered by means of an Oticon BC 461-1(100 Ohm, Oticon, UK) bone conductor vibrator (1.4 cm x 2.4 cm). White background external noise was sent to external loudspeakers at the corners of the computer. The experiment was implemented using a PC laptop (Dell Windows 7) using PsychoPhysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) for MATLAB R2012b.

3.1.1.3 Design and Procedure

The task consisted on determining which intensities of a tactile stimuli were perceived as equal, more intense or less intense than a constant (intensity) auditory stimulus (i.e. a referent), and viceversa, that is, determining which intensities of an auditory stimuli were perceived as equal, more, or less intense than a constant (intensity) tactile stimulus.

The rationale of comparing between modalities was used to allow establishing similarities and differences between them.

The constant intensity of a modality was compared to 11 different intensities on the other modality. The presentation was done in pairs, and the task consisted on deciding which of the stimuli in the pair was the most intense, either the tactile or the auditory, by pressing the corresponding key (T for tactile, A for auditory). The stimuli lasted 1 second each.

Each stimulus comparison was repeated a total of 12 times. A total of 264 trials were introduced. A block consisted of 22 trials, and was composed of half of the trials with auditory changing intensity, and the other half with the tactile trials as the variable ones, and were all randomized. A fixed pause of 3 minutes was introduced in the middle of the experiment. The session lasted approx. 50 minutes.

3.1.1.4 Results and Discussion

The proportion of "tactile/auditory more intense" responses data was employed to calculate a Psychometric function fitting for each participant, and a different fitting was made for each modality that changed in intensity, (i.e., one fitting for the auditory varying in intensity, and one fitting for the tactile variable modality) in every participant.

A logistic function was used to calculate the fitting, using the following formula:

$$F_L(x;\propto,\beta) = \frac{1}{1 + \exp\left(-\beta(x-\alpha)\right)}$$

Where x represents the stimulus intensity, β establishes the value of the slope and $\propto$ corresponds to the threshold: $F_L\ (x = \propto\ ; \propto, \beta)$. The data was computed through Palamedes Toolbox in MATLAB (Prins & Kingdom, 2009).

The purpose of this fitting was to obtain the values at 25%, 50% and 75% of the times perceived as tactile/auditory (accordingly) more intense. It was considered that these indexes would provide individual thresholds of low, medium and high intensities. The fitting provided these values, in terms of the *intensities* that represented these percentages. Figure 6shows an example, which is the fitting of participant 6 for the auditory more intense responses, along with the 25%, 50%, and 75% values.

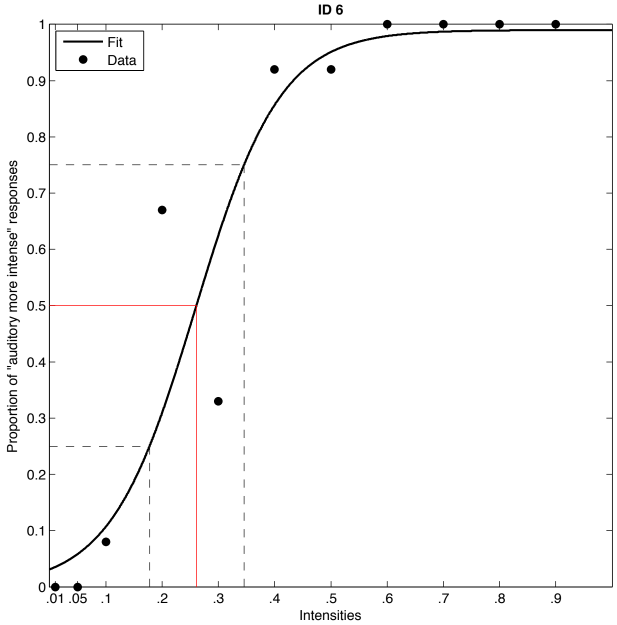


Figure 6. Psychometric Function Fitting for Participant 6 at the Auditory Intensity changing modality. Black dotted lines indicate thresholds at 25% and 75%, whereas the red line indicates the 50% or PSE.

Unfortunately, the fitting was not successful for 8 of the 18 participants, so we discarded their data.

The obtained values or individual indexes of intensities allowed to give the next step, where the individual intensities were tested in an audiotactile duration paradigm and is further described.

Some participants reported it was hard to compare intensities of modalities of different nature, so they somehow found the task hard, however, despite this impression, the main finding is that some participants were *able* to establish similitudes or differences between the magnitudes of stimuli that come from a different nature. Perhaps this result suggests that intensity or magnitude could be seen as an amodal feature (Lewkowicz & Turkewitz, 1980) that allows subjects to compare intensities of stimuli from different modality nature.

3.1.2 Audiotactile duration perception task with different intensities of the distractor modalities

With the previous individual indexes of 25%, 50% and 75%, allowed the implementation of the task described here. This task was the same as the audiotactile duration paradigm explained in chapter one. The aim was to determine whether the balanced incongruence effect or modulatory effects by both audition and touch found in chapter one, still holds for the distractor modalities despite changes in its intensity. This exploration can provide a more robust or complete answer as to the nature of the interaction between audition and touch for this particular duration task.

In line with the previous reported findings from chapter one, here it was expected to find a balanced influence from audition and touch, between one another. However, it was expected to see changes in the effects of the distractor modalities, according to the intensity at play; perhaps a stronger effect could be seen when the distractors were at a low intensity, than a high intensity, enabling more facilitation for instance, at the congruent condition.

3.1.2.1 Participants

10 participants taken from the previous cross-modal matching task sample were asked for participation on this task (mean age 25.18, age range 19-36 years old). All agreed to collaborate and gave prior informed consent before starting the experiment. Participants had to come to two different sessions with approx. one

week separation from each session. They were not informed as to the purpose of the experiment.

3.1.2.2 Apparatus and stimuli

The auditory and tactile stimuli were delivered the same via as described in the previous section. The procedure was similar to the experiments reported in chapter one, so the same physical setup and apparatus were employed, along with the white noise used to cover the sound made by the tactile stimulator.

There were three different intensities of the distractor modalities (tactile/auditory): low, medium and high, and were different for each participant, according to their previous individual psychometric fitting. However, the intensities employed for the target modalities remained constant all across the experiments, and were the same employed in experiments from chapter one. For auditory target stimuli, intensity was 50 dB (SPL), for tactile target stimuli it was 12 Vpp.

Both tactile and auditory target stimuli were presented in standard and probe durations, as in chapter one. The standard duration was 90 ms. Durations for the probe stimuli were 50, 80, 100 and 130 ms.

3.1.2.3 Design and procedure

The design of the experiment comprised the within participants factors of: Target (Auditory/Tactile) along with Stimulus Presentation (Bimodal Congruent, Bimodal Incongruent, Unimodal), Distractor's Intensity (Low, Medium, High) and Duration (50, 80, 100 and 130 ms.). A 2AFC task was employed (the same task used in chapter one), where participants decided which of a pair of target stimuli was the longest, whilst ignoring distractor stimuli in the other modality, that could either have the same duration information (Bimodal Congruent), or the opposing durations (Bimodal Incongruent) as the target modality. The unimodal condition was included to maximize the possibility for the participants to focus on the target modality.

A summary of the types of stimuli that could be found is the following:

a) Bimodal Congruent Low Intensity Distractor

b) Bimodal Congruent Medium Intensity Distractor

c) Bimodal Incongruent High Intensity Distractor
d) Bimodal Incongruent Low Intensity Distractor
e) Bimodal Incongruent Medium Intensity Distractor

Each stimulus under each condition was repeated a total of 10 times. A cycle consisted on the presentation of all the possible stimuli combinations. All the stimuli presentations within a cycle were randomized. There were 10 blocks and all the possible stimuli were randomized in every block. There were a total of 280 trials. A fixed pause with duration of 30 seconds was presented at the end of every block. Each session corresponded to a given target, either tactile or auditory. Order of target task completion was counterbalanced among participants. Instructions were the same as in chapter one, and were given orally as well as in a written form on the screen.

3.1.2.4 Results and Discussion

The proportion of correct responses was obtained for each participant. A within participants ANOVA was computed with the factors Target (Auditory/Tactile), Stimulus presentation (bimodal congruent/bimodal incongruent), Intensity (low/medium/high) and Duration (50, 80, 100 and 130 ms). The Unimodal condition was omitted from this analysis, because it was not possible to compare it with the other factors because it did not present changes in the intensity levels. Table 1 summarizes the results.

A main effect of stimulus presentation was found, indicating that the incongruent condition had a worse performance (M = .621, SE = .026), in contrast to the congruent condition (M = .820, SE = .012, $p = .000$).

The results showed that there was not a significant effect of the factor intensity, which can suggest that the fact of varying the intensity did not interfere in the effects of the distractor on the performance of the incongruent condition. Also, the factor target was not significant, neither the interaction of stimulus presentation x target, so there were no differences in performance between the conditions according to the modality. The lack of a significant interaction can suggest a balanced interference from the distractor modalities between one another, as has been found in the experiments from chapter 2. Therefore, audition and touch modulated one another in the same way, which is reflected in a worse performance when the condition is incongruent.

Table 1. Values for within participants ANOVA.

Factor	df	F	p
Intensity	2	1.07	.36
Stimulus presentation	1	71.08	.00**
Duration	3	40.15	.00**
Target	1	3.63	.08
Intensity * Stimulus presentation	2	.16	.85
Intensity * Duration	6	.68	.65
Duration * Target	3	4.36	.01*
Stimulus presentation * Duration	3	5.50	.00**
Intensity * Target	2	1.15	.33
Stimulus presentation * Target	1	2.33	.161
Intensity * Stimulus presentation * Duration	6	.55	.76
Intensity * Stimulus presentation * Target	2	.28	.75
Intensity * Duration * Target	6	.48	.82
Stimulus presentation * Duration * Target	3	1.93	.14
Intensity * Stimulus presentation * Duration * Target	6	.99	.43

*Indicate significance at p <.05, **p <.01

The following figures summarize the overall performance by stimulus presentation and intensity (Figure 7), as well as the performance along the different modality targets (Figure 8 and Figure 9).

The factor duration, along with its significant interactions is not addressed here, because there was not a particular hypothesis of how the different levels of duration would affect our results. Indeed, it was necessary to employ a range of durations in order to avoid making the task predictable.

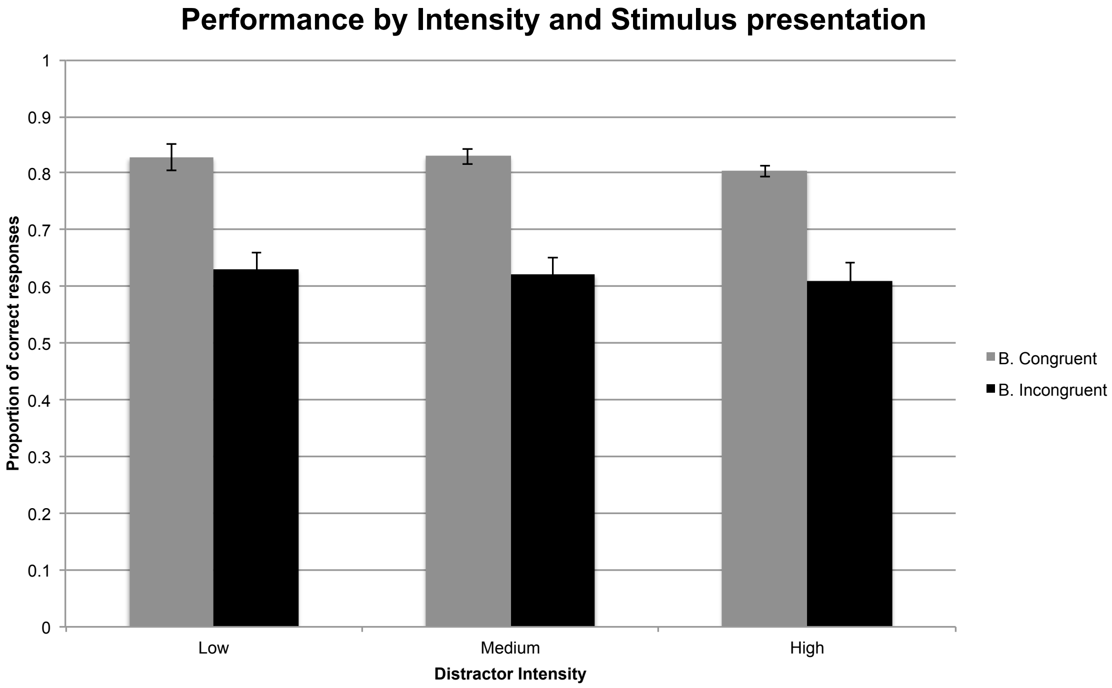


Figure 7 Accuracy performance of the intensity x stimulus presentation factors.

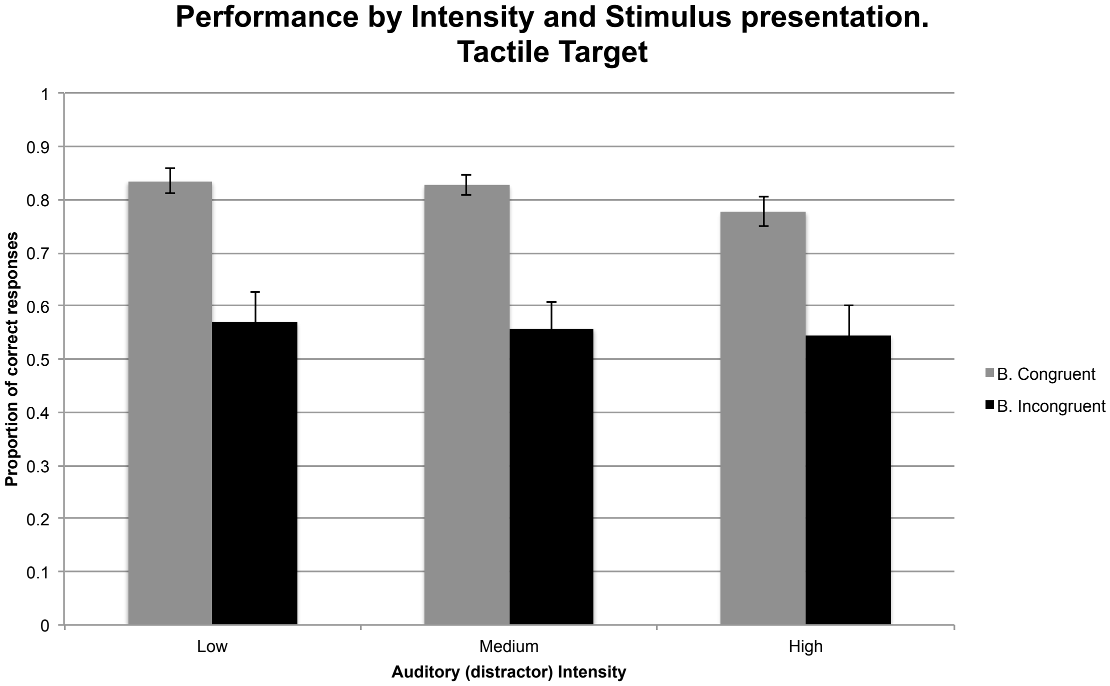


Figure 8. Performance of the intensity x stimulus presentation factors of the Tactile target, with Audition as distractor.

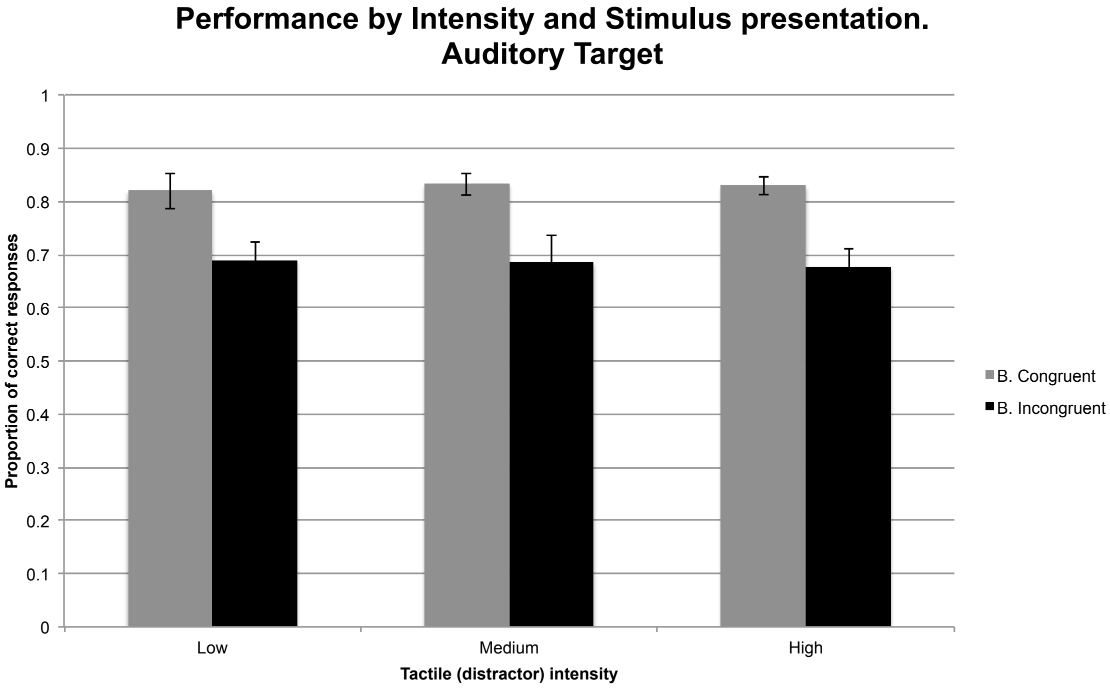


Figure 9. Performance of the intensity x stimulus presentation factors of the Auditory target, with Touch as distractor.

The results presented here showed no effect of changes in the intensities over the audiotactile duration task. One reason could possibly lie on the fact that perhaps the durations employed could have been too short to notice the changes in the intensity of the distractors, that is, they could have been rather imperceptible. Also, the number of repetitions for each stimulus type was very short: only 10 repetitions, and could have impacted the non- significances found for the variables controlled. Some of these flaws on the experiment design motivated the next couple of experiments, where some parameters changed in order to better explore an effect of the distractor's intensities.

Another interpretation could be that maybe the different intensities employed could have been perceived as not so different from one another, and perhaps (despite the indexes provided by the psychometric fitting) the intensities were felt as being very close to one another, so the low intensity might have been heard/felt as not that importantly different from the high intensity, etc. Moreover, the sample size of participants collaborating in this study might have been small (10), which could also represent a reason for the lack of expected results.

3.2 Experiment 2. Exploring the role of the distractors' intensities in an audiotactile duration task with a previous detection task

This experiment comprised two stages. First, a detection task was necessary to obtain intensity thresholds. The second stage involved the testing of the audiotactile duration paradigm. The differences in relation to Experiment 1 lied on the way to establish the intensity thresholds, along with some modifications on the parameters of the audiotactile duration task previously described. Also, the participants used in the detection task were different from the ones collaborating in the audiotactile task.

3.2.1 Determining the stimuli intensities through a Detection task

A different approach was used here to obtain the necessary values to introduce the intensities. Due to the difficulty of establishing similarities between stimuli of different modalities in the previous cross-modal matching task, an alternative approach was considered to avoid these kinds of confusion and explore more directly the possible different intensities. This was done in individual auditory/tactile sensory detection tasks, where participants had to establish whether they perceived or not a stimulus, in a train of stimuli presentation.

3.2.1.1 Participants

A total of 10 participants (14 F, mean age of 24.4, age range of 22-31 years old) from the University of Trento participated in this task in exchange for course credits or reimbursement. Participants gave their prior written consent and were not informed as to the purpose of the experiment.

3.2.1.2 Apparatus and Stimuli

The stimuli were delivered using the same procedure and equipment as the above sessions.

A total of 14 different intensities were employed in the detection task. The intensities were very low with slight variations between them, for example, indexes such as 46.15 dB, 47 dB (SPL) (for the auditory) or .45 Vpp (for the tactile) stimuli. The duration of the stimuli was always of 200 ms.

The reason for using low variations between the intensities was to avoid possible ceiling effects, since a preliminary pilot task, which used bigger variations showed this kind of effects.

3.2.1.3 Design and procedure

A trial consisted on a stimulus (tactile/auditory) that was presented in a given intensity.

Participants could either hear or not a sound for the auditory detection task, or feel or not a vibration, for the tactile detection task. The task consisted on stating whether they heard (or felt) the sound (vibration) or not, by pressing a 'yes/no' key accordingly. The two modality tasks were performed separately, but in a single session. Each stimulus was presented 20 times, so a total of 300 trials were presented for each modality. Each task presented 3 fixed pauses with a duration of 1 minute, and were presented every 5 blocks of stimuli. All trials were randomized, and task completion order was counterbalanced among participants.

3.2.1.4 Results and Discussion

The proportion of 'yes responses' was computed for each target modality accordingly, to then fit the data to a Psychometric Function Fitting for each of the 10 participants. After obtaining the individual's fittings, they were then averaged and submitted to another fitting for each modality, correspondingly (Figure 10 and Figure 11)

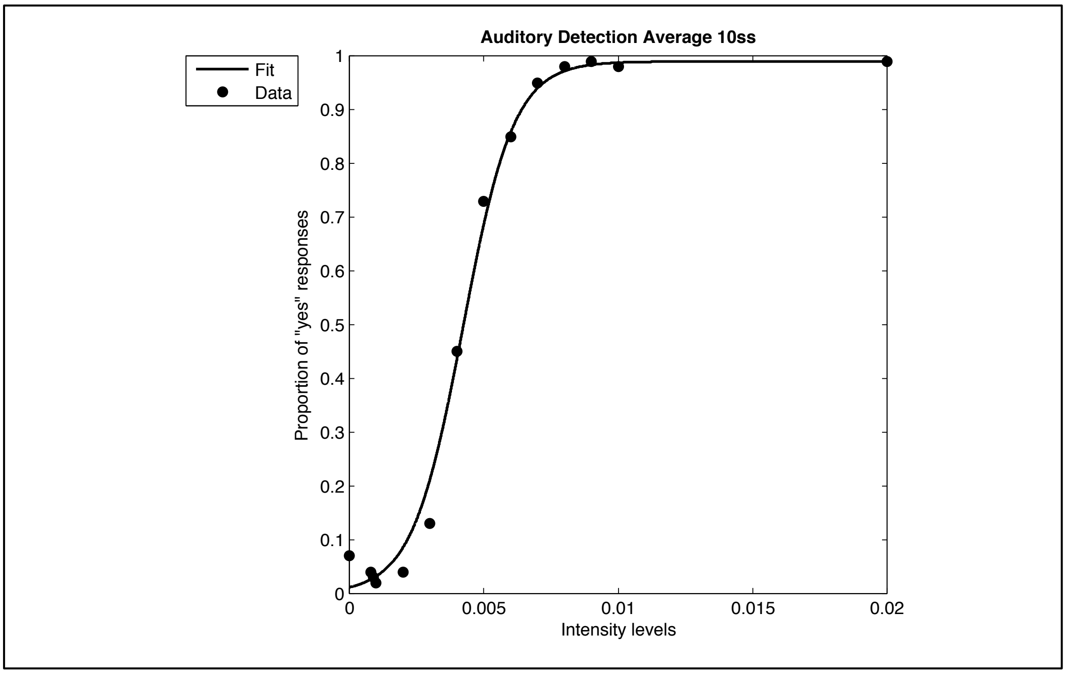


Figure 10. Psychometric Function Fitting for 10 participants at the Auditory Detection Task

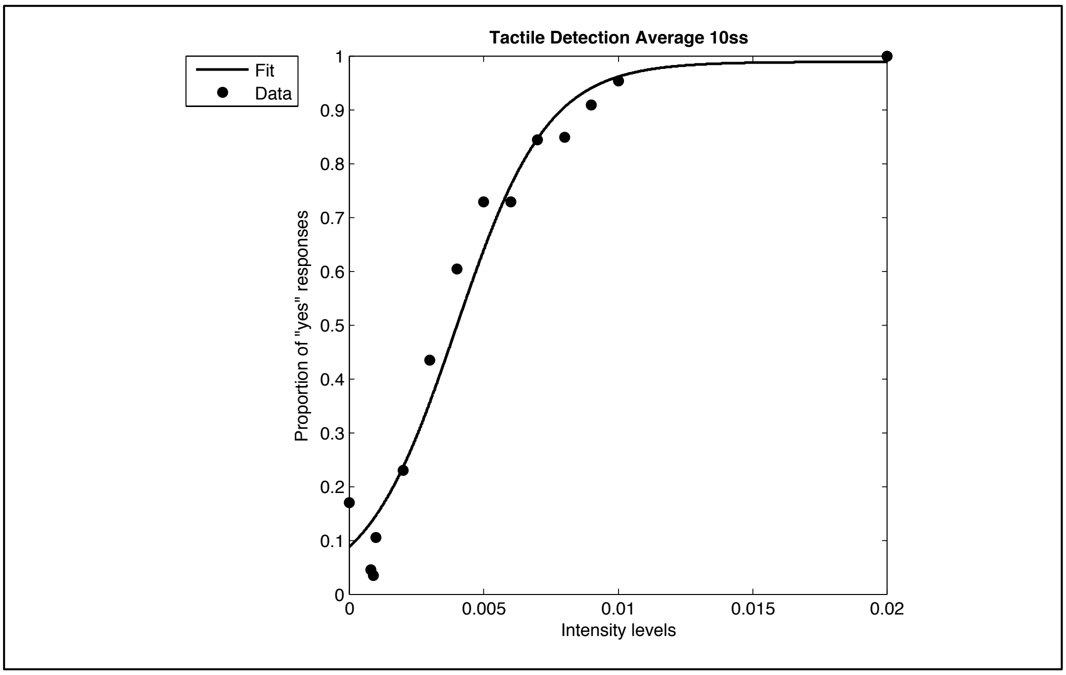


Figure 11. Psychometric Function Fitting for 10 participants at the Tactile Detection Task

Through this fitting, the intensity values were obtained. The stimulus established as the low intensity, was the one that was perceived 98% of the times.

This value was gradually increased to define the medium and high intensity value in the following way: an increase of 1400% from the 98% value was done to obtain the ‘medium value’ to set the standard intensity in the next task. Furthermore, the high value was obtained by increasing the value from the 98% to 7000%. The increment steps of 1400% and 7000% were defined in this way because they seemed to produce noticeable intensity changes.

3.2.2 Audiotactile duration perception task with general intensity values for the distractor modalities.

The purpose of this experiment was to explore the possible influence of changing the saliency or intensity of the modality distractors in an audiotactile duration paradigm. In this experiment, some parameters were changed such as the number of durations introduced, as well as the stimulus presentation conditions, as a way to improve the experiment and thoroughly explore whether there is or not, an effect of the intensity variable.

In the following task, it was not considered necessary to use the same participants as in the detection task because the thresholds obtained from the detection task came from an average among the participants, and the three experimenters took care of this matter and agreed that the low and high intensities were contrasting enough.

3.2.2.1 Participants

A total of 18 participants from the University of Trento (14 F, 20.4 mean age, 19 – 29 age range) took part in this experiment, in exchange for course credit or reimbursement. Participants were not previously informed as to the purpose of the experiment. Since the experiment involved two separate sessions, participants came twice to the lab with one week separation from each session approx.

3.2.2.2 Apparatus and stimuli

The stimuli were delivered via the same procedure and same apparatus as described above.

In this experiment we presented longer durations as a way to make them more perceivable than the shorter durations from the previous task. Also, the number of durations was incremented to 6, in order to augment the number of trials and allow a psychometric data modelling. The durations employed were 200, 210, 230, 250, 270 and 280 ms. and the standard duration had a length of 240 ms.

This time, only low and high intensities were employed. The high and low indexes were taken from the previous average psychometric fitting and were used to produce the stimuli with these intensities on the Matlab script.

Tactile as target task. Here, there were changes in the stimuli of the auditory distractors intensities. The auditory low consisted of 46.15 dB and the auditory high of 71.65 dB, whilst the tactile constant intensity was of 4.40 Vpp and consisted on the middle value obtained in the previous task.

Audition as target task. Here, there were changes in the stimuli of the tactile distractors intensities. The tactile low intensity consisted of .45 Vpp and the tactile high of 15 Vpp, whilst the auditory constant volume consisted of 58.2 dB, which consisted on the middle value obtained in the previous task.

3.2.2.3 Design and Procedure

The stimuli presentation condition presented only the bimodal congruent and bimodal incongruent condition. This time, the unimodal condition was excluded to reduce the length of the experiment and also, because it was not of our interest to vary the intensities of the unimodal condition, making the comparisons against the other presentations (congruent, incongruent) imbalanced.

Therefore, the stimulus presentation and intensity combination was the following:

- Bimodal congruent, low intensity
- Bimodal congruent, high intensity
- Bimodal incongruent, low intensity
- Bimodal incongruent, high intensity

Each stimulus was presented 20 times, so a total 480 trials was presented in each task.

A cycle consisted on the presentation of all the possible stimuli combinations. All the stimuli presentations within a cycle were randomized. A fixed pause of 180 ms. was given every 5 cycles. Order of task completion (tactile or auditory target) was counterbalanced across participants.

As in experiment 1, the procedure employed was the same. Participants needed to decide which of a pair of events in the target modality was the longest (by pressing the no. 1 or 2 key), whilst ignoring the stimuli in the distractor modality.

3.2.2.4 Results and Discussion

The proportion of correct responses from all the participants as computed for both the tactile and auditory modality as targets. This data was used to perform a repeated measures ANOVA with Target (Tactile, Auditory), Stimulus presentation (Bimodal Congruent, Bimodal Incongruent), Intensity (Low, High) and Duration (200, 210, 230, 250, 270 and 280 ms.) as the within participants factors. The results are summarized in Table 2.

Table 2. Values for within participants ANOVA.

Factor	df	F	*p*
Target	1	43.99	.00**
Stimulus presentation	1	48.75	.00**
Intensity	1	27.81	.00**
Duration	5	25.85	.00**
Target * Stimulus presentation	1	26.12	.00**
Target * Intensity	1	14.46	.00**
Stimulus presentation * Intensity	1	30.26	.00**
Target * Duration	5	5.28	.00**
Intensity * Duration	5	3.23	.01*
Stimulus presentation * Duration	5	15.03	.00**
Target * Stimulus presentation * Intensity	1	11.75	.00**
Target * Stimulus presentation * Duration	5	4.51	.00**
Target * Intensity * Duration	5	2.17	.06
Target * Stimulus presentation * Intensity * Duration	5	3.02	.01*

*Indicate significance at $p < .05$, **$p < .01$

Since not all of the main effects and interactions were of our interest, only the effects that are important for the aim of the experiment are discussed in this section.

An initial inspection of the results indicated that there was a main of effect of target, showing that performance was overall better when the auditory modality played as target (M = .694, SE = .015) than when the tactile modality was the target (M = .583, SE = .021, *p* = .000). Moreover, the stimulus presentation condition showed the overall pattern found along the reported here experiment, that is, the incongruent condition (M = .579, SE = .014) had an overall worse performance than the congruent condition (M = .697, SE = .014, *p* = .000).

The intensity factor had a main effect, suggesting that performance was better when the distractor was low (M = .665, SE = .020), than when it was at a high intensity (M = .611, SE = .013, *p* = .000).

The interaction stimulus presentation x target revealed that the congruent condition was better in the auditory as target modality (M = .718, SE = .025) than the tactile target (M = .677, SE = .019, *p* = .000). The same pattern holds for the incongruent condition, with auditory better (M = .669, SE = .020) than tactile performance (M = .489, SE = .020, *p* = .000). In both targets, the congruent condition (tactile: M = .677, SE = .019; auditory: M= .718, SE = .025) had always a significant better performance than the incongruent one (tactile: M = .489, SE = .020; auditory: M= .669, SE = .020). In addition, performance in the low intensity was always poorer for the tactile modality (M = .628, SE = .020) than the auditory (M = .702, SE = .025, *p* = .003); the same case was seen for the high intensity, where the tactile target modality reflected inferior accuracy (M = .538, SE = .012), than the auditory target (M = .685, SE = .019, *p* = .000).

Posthoc pairwise comparisons of the interaction target x intensity revealed that performance in the low intensity condition (M = .628, SE = .020) was better than in the high intensity (M = .538, SE = .012, *p* = .000) but only for the tactile modality as target.

Another interesting interaction showing significance was the Stimulus presentation x Intensity; poshtoc pairwise comparisons (Bonferroni corrected) revealed that the intensity factor played a role only for the incongruent condition, where performance was better in the low intensity condition (M = .641, SE = .020) than in the high intensity (M = .518, SE = .013, *p* = .000). Expectedly, the high intensity distractor had a greater impact, disrupting performance in a bigger fashion.

A facilitation effect was found for the low intensity distractor in the congruent condition (M =.689), when compared to the low incongruent condition (M = .641). Perhaps this finding can be associated to the inverse effectiveness rule of multisensory integration, where bimodal stimuli with low intensities (or at least in one of the modalities), tend to be more integrated. There was however, a facilitation effect also for the high intensity distractor: congruent condition (M = .705) provided better performance than the incongruent condition (M = .518).

For a clearer view of the results, see Figure 12 for a summary of the performance of Intensity by Target interaction, along with the figures representing stimulus presentation by intensity performance, according to the two target modalities (Figure 13 and Figure 14).

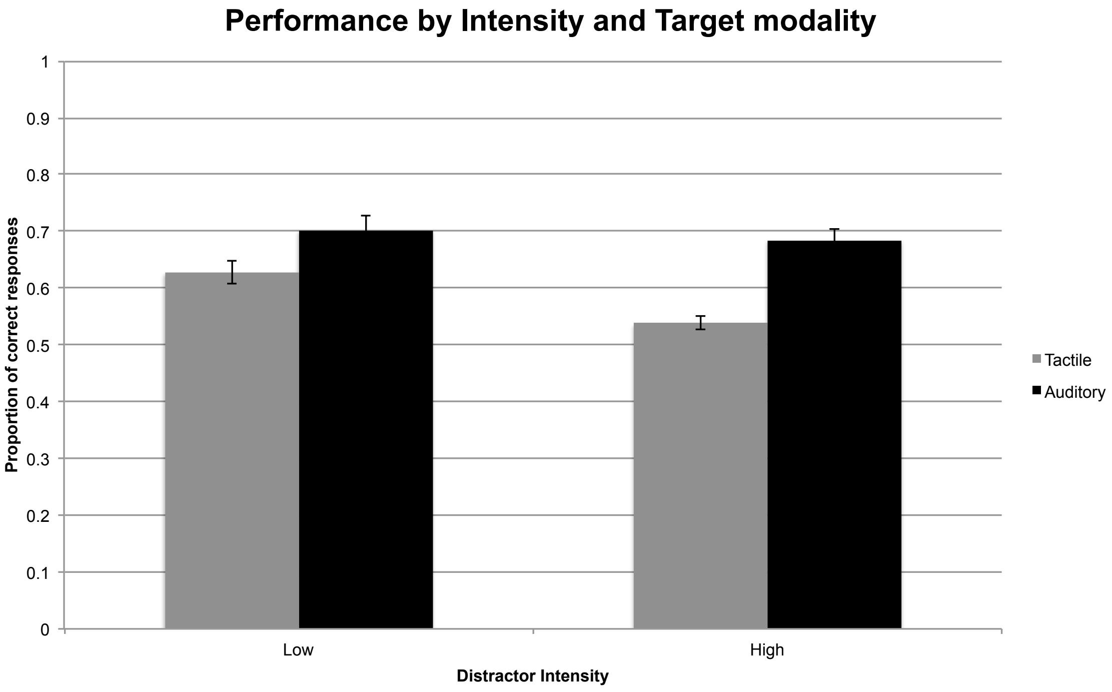


Figure 12. Overall performance of the intensity levels across the different target modalities

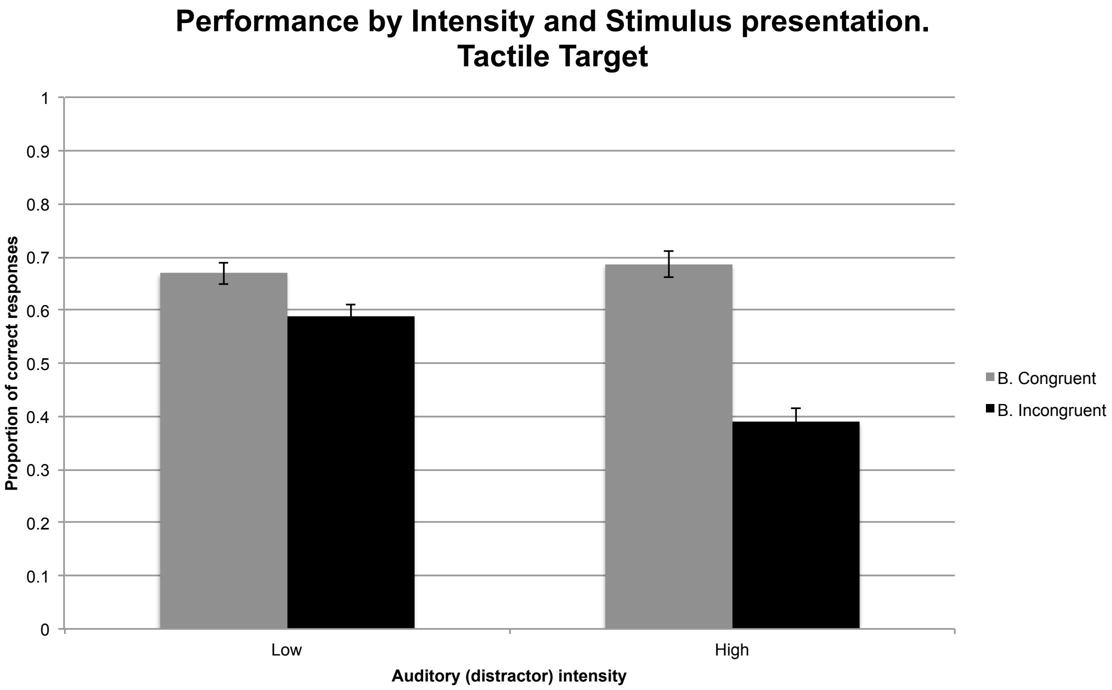


Figure 13. Performance of the intensity x stimulus presentation factors of the Tactile target, with Audition as distractor.

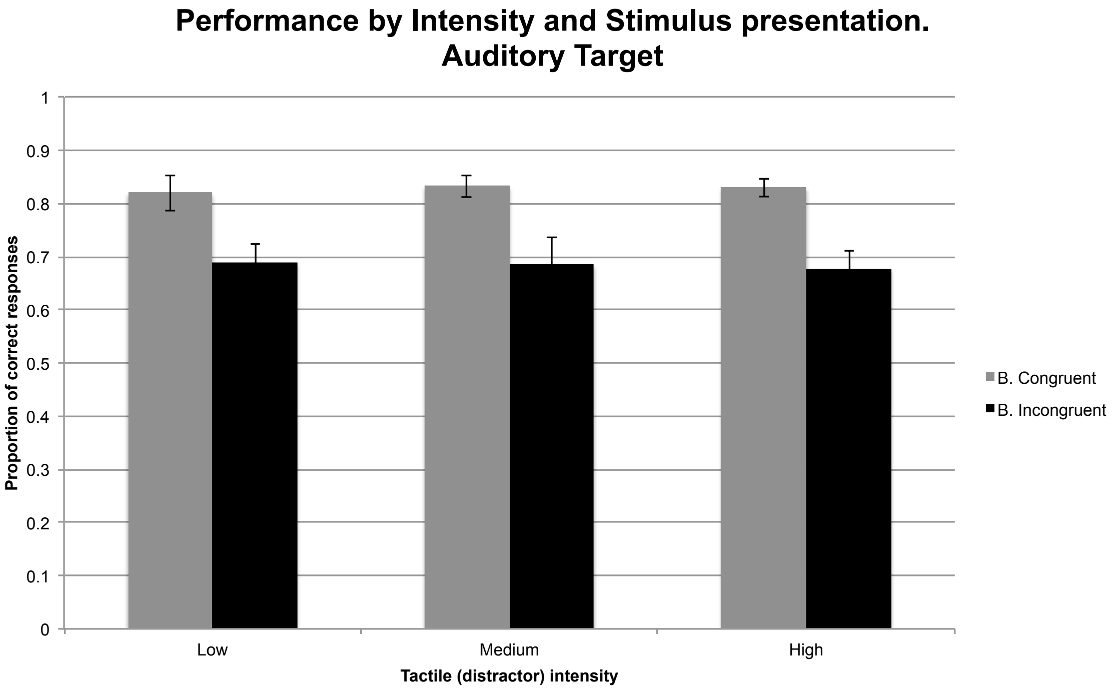


Figure 14. Performance of the intensity x stimulus presentation factors of the Auditory target, with Touch as distractor.

Both low and high intensities were able to generate facilitation effects, when comparing the congruent versus the incongruent condition. Therefore, in this case, both kinds of distractor intensities triggered a facilitation effect, at least for the tactile as target modality. For the auditory modality, however, this facilitation was only appreciated when the intensity of the distractor modality (i.e. tactile) was set to a high level.

With these findings in mind, it could be possible to adhere them to the principle of inverse effectiveness, in which a ‘constant – low’, and also a ‘constant – high’ intensity pairings are both able to produce enhancement effects (as in Diederich & Colonius, 2004 and Bernstein et al., 1973).

The findings reported here are somewhat similar to Ortega et al. (2014), in the sense that all the different intensities adopted by the auditory (or distractor) modality were capable of exerting an influence in vision (in their study), therefore, dominance still prevailed despite using a low and a matched intensity (with the visual), as some findings from our paradigm seem to suggest.

The data also suggests that the overall performance in the low incongruent intensity was better than the incongruent high intensity, suggesting that the high intensity stimuli disrupted performance in a bigger manner.

The overall pattern of results showed that audition in general, had a better performance and therefore, generated the pattern of results described above, according to the experimental conditions. Another interpretation though, could be explained by the following. First, data suggests that the balanced interference between audition and touch is no longer present in this experiment, since the auditory incongruent condition had better performance than the tactile incongruent. Perhaps this suggests that audition as a distractor was *altering* performance more strongly when touch played as target and this generated the difference in performance. This finding might suggest that the direction of the interaction between audition and touch is sensible to the intensities of the stimuli at play. In this case, audition seemed to be more influential to touch than the opposite, when the intensities of the auditory stimuli were truly contrasting (as seen in Experiment 2, reported in this chapter). Another finding was that the auditory congruent condition also had better performance than the tactile congruent, and could be in line with previous literature highlighting the great temporal resolution of audition (Grondin, 1993). This could be possibly read as touch *helping* more greatly audition, generating a facilitation effect when it acted as a distractor. From the pattern of results found in Experiment 2, it can be seen that they are different from the findings of Experiment 1 (where possibly the intensities were not highly contrasting), and from the experiments reported in Chapter 2. These differences could be caused by the methodological differences from each experiment. For example, it could be the case that the method employed in Experiment 2 was more sensible to finding contrasts between the intensities than the methods from Experiment 1. Nevertheless, further work or replications should be made as a way to find out whether the found auditory dominance is dependent on the contrasts between the intensities of the stimuli employed from both audition and touch.

This last experiment demonstrated a direct influence of the intensity of the distractor modalities on performance, according to the different conditions; this is different from Experiment 1, which failed to show an effect of this variable, but however, revealed a balanced interaction between audition and touch, in the incongruent condition, for example. In contrast, Experiment 2 shows differences in modulation.

The difference between these findings could lie in the different methodologies employed, but it is believed that Experiment 2 represents an improved version of the first experiment. For example, in this time, we ensured that the intensity indexes employed were clearly contrasting from one another along each of the modalities. Also, more repetitions of the stimuli and a bigger sample size were introduced to allow revealing the possible effects. Likewise, the range of durations employed was different, so this could also be a variable that could explain the different pattern of results.

Additionally, from this series of experiments, it can also be concluded that participants are able to compare between intensities of different modalities, as shown in the cross-modal matching task from Experiment 1. Despite the phenomenological differences that each modality-intensity might bring, there might be an amodal component or mechanism that allows participants to match and compare them.

References

Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433-436.

Bresciani, J.P. & Ernst, M.O. (2007). Signal reliability modulates auditory-tactile integration for event counting. *Neuroreport*, *18*, 1157-1161.

Colonius, H. & Diederich, A. (2004). Multisensory interaction in saccadic reaction time: a time-window-of-integration model. *Journal of Cognitive Neuroscience, 16*, 1-10.

Diederich, A. & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics*, *66*, 1388-1404.

Diederich, A., Colonius, H., Bockhorst, D. & Tabeling, S. (2003) Visual-tactile spatial interaction in saccade generation. *Experimental Brain Research*, *148*, 328-337.

Grondin, S. (1993). Duration discrimination of empty and filled intervals marked by auditory and visual signals. *Perception & Psychophysics, 54*, 383-394.

Kleiner, M., Brainard, D., & Pelli, D. (2007). "What's new in Psychtoolbox-3?" *Perception*, 36, ECVP Abstract Supplement.

Lewkowicz, D.J. & Turkewitz, G. (1980). Cross-Modal Equivalence in Early Infancy: Auditory-Visual Intensity Matching. *Developmental Psychology, 6*, 597-607.

Ortega, L., Guzman-Martinez, E, Grabowecky, M & Suzuki, S. (2014). Audition dominates vision in duration perception irrespective of salience attention and temporal discriminability. *Attention, Perception & Psychophysics, 76,* 1485–1502.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437-442.

Prins, N., & Kingdom, F.A.A. (2009). Palamedes: MATLAB routines for analyzing psychophysical data. http://www.palamedestoolbox.org .

Rach, S. & Diederich, A. (2006). Visual-tactile integration: does stimulus duration influence the relative amount of response enhancement? *Experimental Brain Research*, *173*, 514-520.

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

Walker, J.T. & Scott, K.J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones and gaps. *Journal of Experimental Psychology. Human Perception and Performance,* 7, 1327–13

Chapter 4. Exploring the role of spatial location in an audiotactile duration perception task.

4.1 Introduction

As has been previously described in the main introduction, the literature regarding multisensory integration often establishes some basic principles that have been seen to influence the occurrence of multisensory interactions. One of these main principles is the spatial rule, which has been explored at a neural level, but however also tested at a behavioural level (Stein & Meredith, 1993). This rule establishes that multisensory interactions can occur whenever there is a spatial alignment or even overlap of the receptive fields of the sensory modalities. As a result, multisensory interactions can be reflected in for example, facilitatory effects or enhancement of multimodal responses (Murray et al., 2004). These interactions can occur even when the bimodal stimuli do not coincide spatially in exterior locations or referents, as long as the neurons involved present large receptive fields that can still be stimulated by the given position information (Wallace & Stein, 2007). With this rule in mind, the following introduction describes some of the previous work done at a behavioural level (mostly), which document the advantages of presenting bimodal stimuli within an alignment of spatial location (or not).

These advantages and evidence of multisensory binding when showing spatial co-occurrence of stimuli from different modalities, can be seen in different measures used in multisensory integration. For example, the measures of the just noticeable difference (JND) in temporal order judgments tasks. This measure represents the smallest amount of necessary time between stimuli, needed to provide an accurate response on 75% of the times in these kind of tasks (Zampini et al., 2005). As documented, when the bimodal stimuli are distant in spatial locations, just noticeable differences are clearly shorter, indicating that it is easier to distinguish the occurrence of the stimuli. When presented in close spatial proximity, these judgments are harder to accomplish, at least for audiovisual and visuotactile bimodal stimuli, and reflected in larger JND indexes (Zampini, Shore & Spence, 2003a, b; Spence, Shore & Klein, 2001). A plausible explanation is that when stimuli come from the same position, there is no redundant information as to their location, making it harder to identify them, as opposed to the multiple or redundant information provided by having stimuli coming different locations (Zampini et al., 2005).

Moreover, other kinds of tasks, such as stimulus detection have shown facilitation effects when addressing the role of alignment in spatial location. For example, Harrington and Peck (1998) documented the existence of facilitation effects when visual and auditory stimuli were presented simultaneously within the same location, where participants had to make saccades to bimodal (audiovisual) and unimodal (auditory or visual) stimuli. The findings revealed an improvement in saccadic reaction times when the stimuli presentation was bimodal and coincided spatially (in contrast to unimodal presentations). However, this first experiment lacks a condition in which the bimodal stimuli are separated. In their second experiment, they went further to explore the degree of distance separation that these bimodal audiovisual stimuli could allow, in order to still show a facilitation effect. They found that up to 25º degrees of separation could still allow an enhancement effect.

The findings by Caclin, Soto-Faraco, Kingstone and Spence (2002) also illustrate a facilitation effect when controlling the spatial location of bimodal stimuli (Experiment 1 & 2). Participants were required to localize auditory stimuli, whenever there was a tactile simultaneous presentation of stimuli, occurring from separate spatial locations (i.e. tactile stimuli at the center, auditory stimuli occurring at the left or right of the tactile stimulators). Their results are discussed in terms of a tactile capture effect of audition, where the tactile location influenced the perceived location of an auditory stimulus, i.e. the position of the latter stimulus was perceived according to the location of the tactile one (Caclin et al., 2002).

Additionally, there are some studies that address the role of the spatial location of audiotactile interactions in temporal perception.

4.1.1 The study of the spatial location factor in audiotactile studies of temporal perception

Zampini et al. (2005) tested the effects of spatial location (same, different) on the precision of audiotactile temporal order judgments. The authors employed different stimulus onset asynchronies (SOA) between vibrotactile stimuli and white noise bursts. The two different modalities' stimuli could come from either the same position (the loudspeakers located immediately behind the tactile stimulators) or a different spatial location (with a separation of 26 cm distance from one another, from left to right locations). The data was analysed in terms of both JND and point of subjective simultaneity (PSS). The PSS refers to the amount of time needed for a stimulus to precede another one, in order to perceive both of them as occurring simultaneously (Zampini et al., 2005). The data revealed that there were no significant differences between the JND's of the same versus different spatial location of the stimuli from the different modalities, neither for the PSS of same vs. different location. The direction of these findings were reported both for inexperienced and experienced psyschophysical observers (Zampini et al., 2005). The findings suggest that this particular modality pairing is not benefited by the fact of placing stimuli in different spatial locations in order to improve precision in temporal order judgments, as opposed to audiovisual and visuotactile combinations where redundant spatial cues seem to aid stimuli identification (Zampini, Shore & Spence, 2003a, b; Spence, Shore & Klein, 2001). The lack of effects found for the audiotactile pairing could suggest somehow that these two modalities are particularly less sensitive to spatial relations, as opposed to vision (Zampini et al., 2005), thus showing no advantage from different spatial location information.

There are some other different studies that also document the effects of aligning and misaligning audiotactile stimuli. A few of them are described in the following section.

4.1.2 Studying the effects of spatial location in audiotactile interactions in other kinds of perceptual tasks

Lloyd, Merat, McGlone and Spence (2003) designed a study to address the effects of spatial attention in the perception of auditory and tactile stimuli. In their task (Experiment 2), participants had to discriminate between the elevation of auditory and tactile stimuli, when they came from the same spatial location, or from a different location, and also when they had an expectation or not as to where the stimuli would appear. Their results indicated that when participants had expectations that bimodal stimuli would come from the same location (and actually came from the same locations), their responses were faster (expected location condition). However, when stimuli were expected and presented in opposite positions, it was harder for participants to attend to the different locations, since it became harder to split the attention to different locations (Lloyd et al., 2003).

These findings were discussed in relation to theories of attention. For example, the modality specific resource model (Duncan et al., 1997), suggests that perhaps each modality has an attentional spatial system of their own, thus allowing to attend to different locations simultaneously. However, this model could not account for the found data, since if there were different attentional systems for each modality, it would have been equally easy to perform the task and split attention when attending to different modalities in different locations, than to attend to a single location, and this was not the case (Lloyd et al., 2003). The authors explain that the pattern of their results is more in line with a model of 'separate-but-linked attentional systems' for each modality, which specifies that each modality has an attentional system of its own, but the cross-modal links and crosstalk that exists between the modalities (which influences one another at some point) can sometimes make it harder to direct attention to the different modalities in different locations (Spence & Driver, 1996).

In an interest to understand whether the spatial coincidence factor impacts early audiotactile interactions in the human cortex, Murray et al. (2004) also tested the effects of spatially aligning audiotactile stimuli, by means of electrophysiological and behavioural recordings. The task consisted on the speeded detection of unimodal and bimodal audiotactile stimuli (white noise and tactile vibrations) that could coincide or not in spatial location. Interestingly, reaction times were not significantly different for the same/different spatial location of the stimuli, indicating no effect of this variable, however, the bimodal stimuli yielded faster responses than unisensory (Murray et al., 2004). These findings are consistent with Zampini et al. (2005), were no differences in the effects of spatial location have been found for audiotactile interactions.

Zampini, Torresan, Spence and Murray (2007) conducted a similar study in which participants had also to detect unimodal and bimodal audiotactile stimuli, with aligned or misaligned spatial location of the bimodal stimuli. Nevertheless, for the different location condition, instead of placing the stimuli on the opposite side as in previous studies (i.e. right or left), they were placed behind, that is, on the back side of the participants in an interest to understand whether there was a difference in terms of facilitation, when the stimuli were not presented in front of the participant, as other studies document. The motivation behind using a rear position lies in some claims that there might be different mechanisms involved in processing audiotactile stimuli located at the front and back part of the body, because in the back part, visual cues become unavailable and therefore auditory and tactile cues come to be more dominant (Kitagawa, Zampini & Spence, 2005; Zampini et al, 2007).

Zampini et al. (2007) replicated Murray et al. (2004) findings, showing no differences in terms of reaction times with same/different bimodal locations, and obtaining facilitatory effects on both conditions.

In sum, some studies demonstrated that for audiovisual pairings, there is a stronger effect of the spatial coincidence (Zampini 2003a,b) than for audiotactile pairings. One of the explanations for a lack of effects for the audiotactile pairings might stem from the fact that audition and touch are less prone to spatial relations, and have overall a worse spatial resolution than vision (Zampini et al., 2005). However, the evidence as to whether spatial alignment influences, facilitates or alters perception for audiotactile stimuli, is still inconclusive, as can be also seen in the description of the studies included in this section.

To the best of my knowledge, it is still unknown whether the proximity or distance between the spatial locations of audiotactile stimuli can affect or enhance duration discrimination. The following proposed study attempted to explore the role that the spatial position of target and distractor (i.e. same, different) could play in the temporal perception of duration. The motivation behind this study was to understand if the same or different spatial locations represent a factor that could either enhance the integration between audition and touch or instead create an interference between them in duration discrimination. The research questions lied as to find out whether a different spatial location of target and distractor of audiotactile stimuli would reduce the interaction between the modalities, reflected in improved performance when the target and distractor modality provided different duration information of two events (bimodal incongruent condition). Also, the aim was to learn whether the same location of the stimuli creates a better duration discrimination when target and distractor present the same duration information in a pair of events (bimodal congruent condition).

4.2 Method

4.2.1 Participants

A total of 12 participants from the University of Trento participated in this experiment, in return for course credits or reimbursement (6 F; age range from 22- 26 years old; mean age 24). They participated in two different sessions, with one week separation from each session. They were not informed as to the purpose of the experiment, and gave prior written informed consent. The experiment was conducted in accordance to the ethical guidelines from the 1964 Declaration of Helsinki, as well as the guidelines from the University of Trento. All reported normal hearing and tactile sensitivity.

4.2.2 Apparatus and Stimuli

The apparatus used to deliver the stimuli were the same as in the previous experiment. The tactile stimuli were given by an Oticon BC 461-1 (100 Ohm, Oticon, UK) bone conductor vibrator (1.4 cm X 2.4 cm). The auditory stimuli were delivered via an external loudspeaker (Dell A215, USA) that was changed of location depending on the experimental conditions (i.e. to left or right of the experimental table). External white noise was played through 2 external loudspeakers (Dell A215, USA) located at the corners of the experimental table, in order to cover the sound produced by the tactile stimulator. It was controlled through an mp3 player (ipod Nano, Apple, USA), and had an intensity of 85.6 dB (SPL). Participants used headphones to cover the loudness and disturbance produced by the white noise, but we made sure that it was still possible to hear the other auditory stimuli.

The experiment was delivered through a PC laptop (Dell Windows 7) using PsychoPhysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) for MATLAB R2012b.

Tactile stimuli consisted on 200 Hz vibrations (0.085 N) delivered to the left or right hand’s index finger (depending on the experimental assigned group). The tactile stimulator was placed on top of a foam rectangle located at a distance of 26 cm from the fixation point on the monitor, and 38 cm away from the participant. Auditory stimuli consisted of pure tones with a frequency of 200 Hz and a constant intensity of 86 dB (SPL), and were located behind the Oticon, at a distance of 40 cm away from the participant. Responses were given through response keyboard located at a distance of 18 cm away from the participant.

The loudspeaker and the tactile stimulator (along with the index finger involved in the task) were covered by a black cloth and hidden from view, in order to avoid guessing the purpose of the experiment. For a clearer view of the experimental setup see Figure 15.

As in the experiments reported in Chapter 2, both tactile and auditory stimuli were presented in a fixed constant duration, that is, a standard duration, as well as in different variable durations, which were termed probe durations. The standard duration was of 240 ms, whereas the probe durations consisted of 200, 210, 230, 250, 270 and 280 ms each, so they could be either shorter or longer than the standard duration.

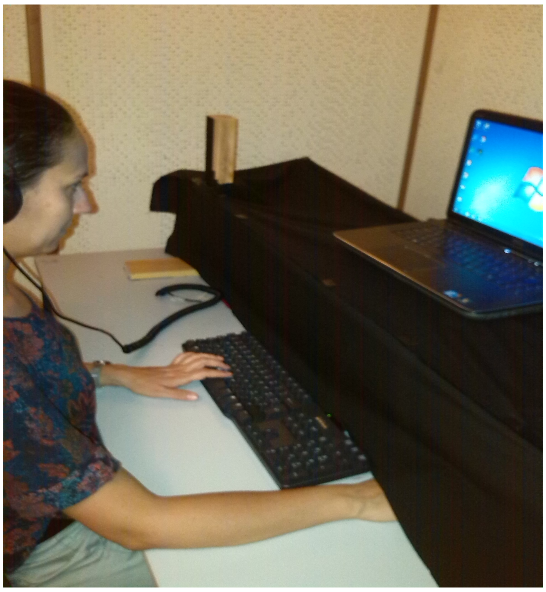


Figure 16.A view of the physical experimental setup

4.2.3 Design and Procedure

A within participants design was adopted with the following factors: Target (Auditory, Tactile), Position of the stimuli (Same, Different), Stimulus Presentation (Bimodal Congruent, Bimodal Incongruent and Unimodal) and Duration (200, 210, 230, 250, 270 and 280 ms). A 2AFC task was adopted for each of the possible modality targets, and consisted on attending to a target (either tactile or auditory) whilst ignoring duration stimuli from the distractor modality (i.e. the opposite modality, correspondingly).

Each trial was composed of a pair of events. Each event had a simultaneous presentation of a stimulus in the target modality, and a stimulus in the distractor modality.

For bimodal trials, both the target and distractor modality presented a fixed standard duration and a variable probe duration (detailed in the previous section). However, the manner or order in which the standard and probe duration were presented varied according to the experimental conditions. For the *bimodal congruent condition*, the target and distractor modality presented the same durations. For the *bimodal incongruent condition*, the events had the opposite duration information (i.e., when one of the target modality stimulus was presented in the standard duration, the distractor was presented in the probe duration, and viceversa for the other stimulus). The *unimodal condition* for each target modality was included to serve as a baseline for comparing the other conditions. Unimodal trials kept the same structure of events from the previous conditions (i.e., a standard and a probe duration stimulus). The order of presentation of the target standard and probe durations was counterbalanced across the two events of a trial, for all the conditions (e.g., standard – probe, probe – standard). The factor Stimulus presentation was presented randomly. For a clearer view of the conditions, see Fig. 2 included in Chapter 2.

The participants' task consisted on deciding which of a pair of stimuli from the target modality was the longest by pressing 1, if the first stimulus of the pair was the longest, or 2 if the second stimulus was the longest.

There was a total of two experimental sessions done in separated days. One session consisted of attending to only one target, either tactile or auditory. Each session was composed of two experimental blocks: **same** and **different** position, and a pause was presented between them to allow the experimenter to manually change the position of the stimuli, and give the participant a short break. The completion order of 'same/different' position blocks was counterbalanced, so half of the participants started with the 'same' position condition, and the other half with 'different' position for both targets. Each session took approximately 50 minutes to be completed.

After each block was completed, the experimenter entered the room and asked the participant to briefly leave it, allowing the experimenter to manually change the spatial location of the distractor, without having the participant to see this procedure.

Target task completion was also counterbalanced across participants, so half of them started with the auditory target task in the first session, then the tactile target for the second session, and viceversa for the other half of participants.

Additionally, the actual location (left or right) of the stimuli, both for target and distractors was counterbalanced: for the **same** position condition, half of the participants had the stimuli presented on the left side of the experimental table (the loudspeaker on the left, and their left hand index finger tactually stimulated); the other half of participants were presented with the stimuli on the right side of the table (the loudspeaker at the right, and their right hand index finger tactually stimulated). When presenting the **different** position condition, the distractor was moved to the opposite side (e.g. if the "same" position was on the left, then for the "different" position, the distractor would be on the right), nevertheless, the position for each type of distractor (for the auditory and the tactile task) was the same for each participant (e.g. if the participant started with the left for the auditory, then continued with the left for the tactile distractor).

The locations of the two oticons were inverted between participants to compensate for any difference in intensities between them.

Participants sat in front of the computer with a fixed location. We made sure that all participants always seated in the same position, with a fixed distance from the experimental table, trying to make their body midline coincide with the centre of the table. Instructions were provided verbally and on the computer screen.

A total of 180 trials were presented for each block. Each block was composed of 10 cycles, which involved the repetition of each duration (6) on each of the stimulus presentation conditions (3), so a total of 18 trials composed each cycle, and the trials were presented randomly. Once a cycle was completed, the next cycle followed within a block. There were 10 trials for each duration along the entire block.

4.3 Results and Discussion

The proportion of correct responses were obtained for each individual in each of the experimental conditions, in order to compute a repeated measures ANOVA with the factors TARGET (Auditory, Tactile), POSITION OF THE STIMULI (Same, Different), STIMULUS PRESENTATION (Bimodal Congruent, Bimodal Incongruent and Unimodal) and DURATION (200, 210, 230, 250, 270 and 280 ms).

Since the number of factors and interactions involved in the analysis is considerable, they are summarized in Table 3.

As the table indicates, there was a significant main effect of Stimulus presentation. Posthoc pairwise comparisons showed that the bimodal incongruent condition had the lowest performance (M = .538, SE. = .011,) in relation to the Bimodal Congruent (M = .696, SE =.02, p <.001) and Unimodal condition (M =

.673, SE = .018, p <.001). The latter two were not significantly different from one another. Figure 17 shows this tendency more clearly by target modalities.

A main effect of Duration was also found. Posthoc pairwise comparisons revealed that the duration of 200 ms (M = .715, SE = .025) provided greater overall performance than the durations of 230 (M = .578, SE = .013, *p* = .003), 250 (M = .542, SE = .018, *p* < .001), and 270 ms (M = .639, SE = .022, *p* = .031). Also, the duration of 210 ms (M = .676, SE = .023) had a significant greater performance than the durations of 230 (*p* = .002) and 240 ms (*p* < .001). The latter two durations were not different from one another, however the duration of 230 had a significantly lower performance in contrast to the duration of 280 ms (M = .663, SE = .026, *p* = .031). The duration of 250 ms had a worse performance when compared to the durations of 270 (*p* = .002) and 280 ms (*p* < .001). In sum, due to their proximity to the standard referent (240 ms), the durations closer to this standard duration (230 and 250 ms) provided the worst performance, because duration discrimination was harder to accomplish. See Figure 18 for a clearer view of the data.

Table 3. Summary of the data from the repeated measures ANOVA.

Factor	gL	F	*p*	Notes
Target	1, 11	3.39	.09	NS
Stimulus Presentation	1.20, 13.20	71.77	.000	** GGC(X^2 = 10.95, *p* = .004)
Duration	5, 55	23.66	.000	**
Position	1, 11	1.92	.193	NS
Target * Stimulus Presentation	2, 22	.14	.869	NS
Target * Duration	5, 55	.48	.783	NS
Stimulus presentation * Duration	10, 110	5.75	.000	**
Target * Position	1, 11	.56	.469	NS
Stimulus Presentation * Position	2, 22	.24	.788	NS
Duration * Position	5, 55	.69	.632	NS
Target * Stimulus Presentation * Duration	10, 110	1.10	.367	NS
Target * Stimulus Presentation * Position	2, 22	.24	.782	NS
Target * Duration * Position	2.78, 30.63	.35	.875	NS GGC(X^2 = 35.17, p = .002)
Stimulus Presentation * Duration * Position	4.78, 52.61	1.33	.221	NS GGC(X^2 = 91.50, p = .004)
Target * Stimulus Presentation * Duration * Position	4.64, 51.05	1.80	.06	NS GGC(X^2 = 82.80, p= .02)

** Indicate significance at p <.001, GGC indicates Greenhouse Geisser corrected

Since we did not have a specific hypothesis for the Duration factor, and also for data simplification purposes, the directions of the interactions for the Stimulus presentation x Duration are not included here.

All posthoc comparisons were made with Bonferroni corrections.

There was not a significant effect of spatial location, however for a clearer view of the trends, see Figure 19 and Figure 20 break down by modality targets.

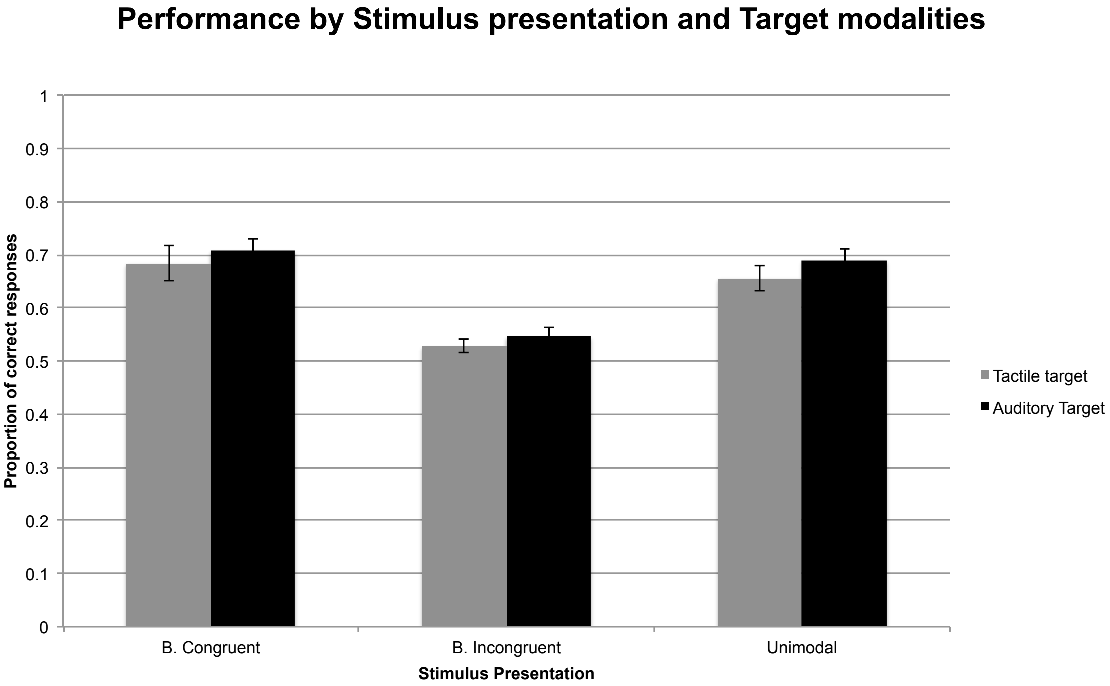


Figure 17. Performance by Stimulus presentation, according to the tactile and auditory targets. Gray bars indicate ‘tactile target’, whereas black bars ‘auditory target’. Error bars indicate standard errors.

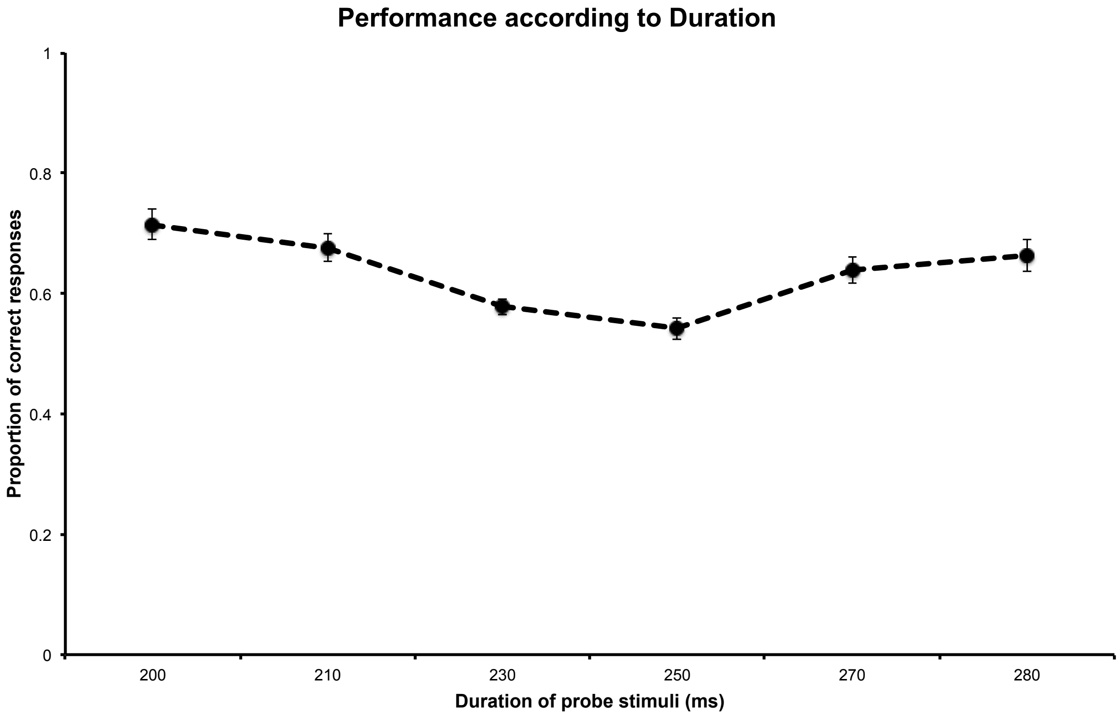


Figure 18. Overall Performance by Duration. Error bars indicate standard errors.

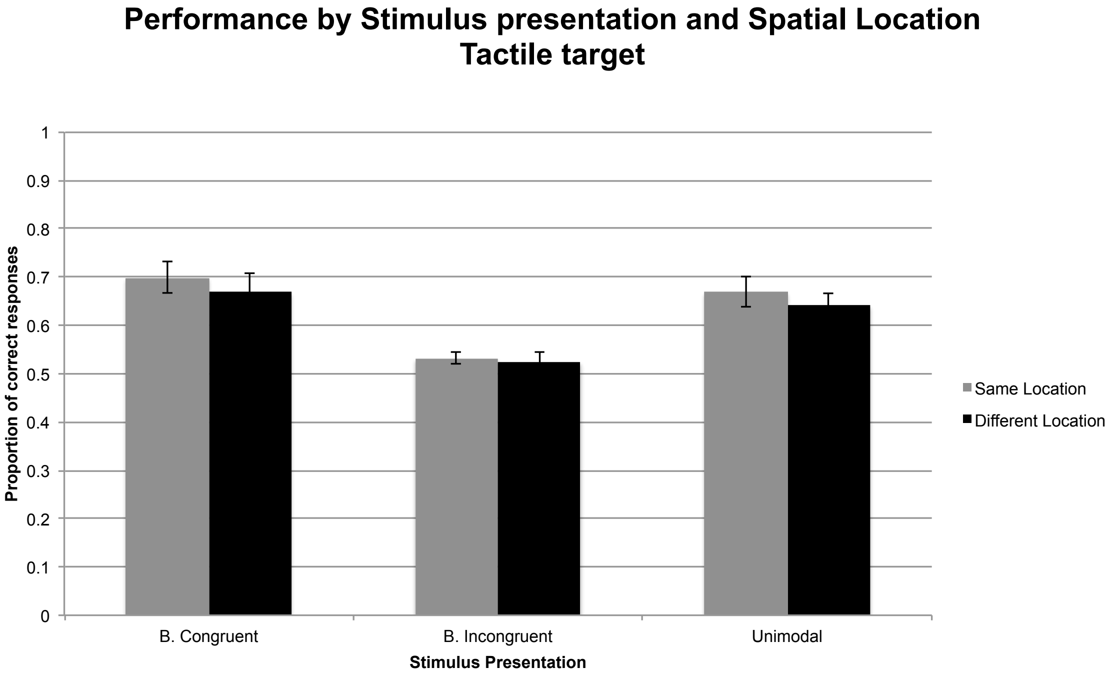


Figure 19. Performance by Stimulus presentation, according to a same or different spatial location position, for the Tactile modality as target. Gray bars indicate ‘same location’, whereas black bars ‘different location’. Error bars indicate standard errors.

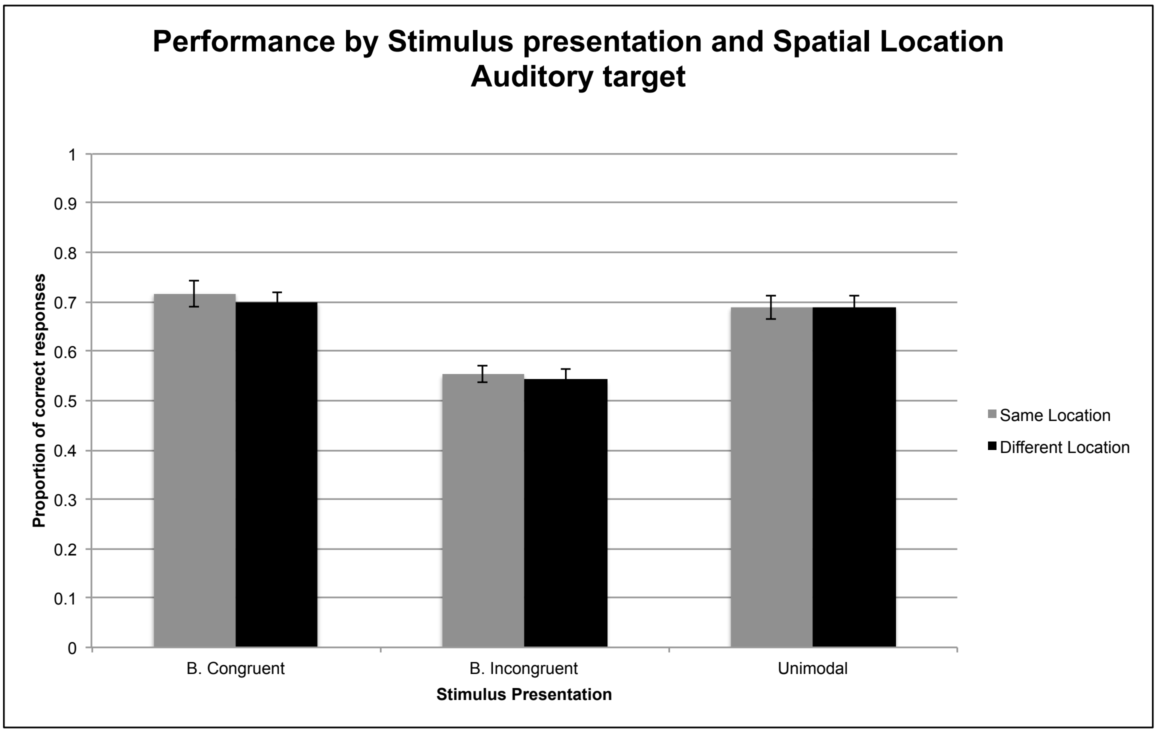


Figure 20. Performance by Stimulus presentation, according to a same or different spatial location position, for the Auditory modality as target. Gray bars indicate 'same location', whereas black bars 'different location'. Error bars indicate standard errors.

In this experiment, it is possible to notice a replication of the findings reported in chapter 1, in terms of the effects of one modality over the other one in the performance of an audiotactile duration task. Here again, the bimodal incongruent condition produced an overall poorer performance, indicating that audition and touch interfere with one another in a balanced manner, since the data does not suggest a difference in terms of modulation from one modality to the other (i.e., one modality does not 'impair' more strongly the performance in the other modality, etc.).

This modulatory effect can also be seen despite changes in the durations employed. In this experiment, we used longer durations on a range of 200 – 280 ms, to make the durations more salient or noticeable. It is also worth noticing that the found interference effect also holds despite the use of external speakers instead of headphones.

Nevertheless, there were no differences in performance according to the factor spatial position of the distractor stimuli; for example, there was no more integration between the modalities (as a facilitation effect would suggest) when being closer to one another, for the congruent condition, neither more/less interference from one modality to the other when being close enough, or far from each other. These findings are in line with Zampini et al. (2005) work on audiotactile temporal order judgments, where it is reported that redundant spatial cues did not play a role in the precision of these judgments. For example, presenting stimuli from different locations did not help temporal order judgments as compared to presenting stimuli from the same spatial location, i.e. there were no differences in the precision of temporal order judgments (Zampini et al., 2005).

In the presently reported experiment, it was expected to find no effects of the distractor modality in the bimodal incongruent condition, when the location of the target and distractor stimuli, so perhaps no decrement in performance. A spatial misalignment of both target and distractor would perhaps generate a more discernible situation, making it easier to disentangle the duration information from the target and distractor modalities. A possible reason for this lack of effect, could lie perhaps in the fact that the stimuli were not separated enough (52 cm distance separation), so possibly still created a sensation that they were coming from the same spatial location, despite that in the different spatial location condition, both the target and the distractor were presented in the opposite side of the participants' body midline (right - left). However it is rather hard to tell what distance is optimal for a multisensory optimal integration or interference.

Also, there was no facilitation or enhancement in duration discrimination when the target and distractor were presented in the same location, in the bimodal congruent condition. Again, this is in line with some of the studies discussed in this introduction, where there are no effects or differences in performance according to the spatial location of audiotactile stimuli, when presented in the front of the participant (Murray et al., 2004; Zampini et al., 2005) thus suggesting that interactions between audition and touch are not so spatial sensible, due to its overall lower spatial resolution (Zampini et al., 2005).

Nevertheless, evidence is inconclusive as to whether different spatial location helps or not.

Further work could be developed to test whether audiotactile interactions in the back of the head are more susceptible to changes in the spatial locations and benefit from this variable, in an audiotactile duration discrimination task, since previous work has shown that these interactions are stronger or benefit in relation to the same or different spatial location of bimodal stimuli, appearing on the rear part of the body (Kitagawa et al., 2005; Zampini et al., 2007).

However it is still not very clear why would a rear position for audiotactile stimuli would further help or show benefits from a same or different spatial location. Vision is mostly spatial and seems to benefit from the spatial alignment/misalignment, at least in temporal order judgment tasks, (Zampini 2003 a,b), so why would putting audiotactile stimuli in a position different from the usual visual (in front, with no visual feedback), i.e., making it even more different than vision, would generate a more powerful effect of spatial alignment/misalignment between the stimuli? It seems to be that probably audition and touch become more relevant when they are positioned in the back of the head, and therefore, the spatial cues available (since vision is obstructed), can only be taken from the information provided by audition and touch (Zampini et al., 2007).

In a different task, Tajadura-Jiménez et al. (2009) found that the integration of auditory-somatosensory stimuli (i.e. bimodal stimuli) benefits from the spatial alignment (i.e. same side presentation) between these stimuli and also depends on the body-part being stimulated. More specifically, the authors found that when performing a speeded detection task, participants were faster when there was spatial proximity to the sound source (20 cm), when the stimulated body part was closer to the participant's body (earlobe rather than hand), when the stimuli were aligned (i.e. both sound and electrocutaneous stimuli delivered to the same side of the body midline) and when the presented sound was at a high frequency (13-17 kHz) (Tajadura-Jiménez et al., 2009). These findings employed a task different to time perception paradigms, nevertheless they could provide hints for future studies. For example, a further study could be implemented on which the body part to stimulate could be the earlobe, rather than the hand as our study has done. Perhaps stimulating this body part, along with a closer distance to the sound source could more clearly show an effect of the spatial location on audiotactile stimuli.

References

Caclin, A., Soto-Faraco, S., Kingstone, A. & Spence, C. (2002). Tactile " capture " of audition. *Perception & Psychophysics, 64*, 616–630.

Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature, 387*, 808-810.

Harrington, L. K., & Peck, C. K. (1998). Spatial disparity affects visual-auditory interactions in human sensorimotor processing. *Experimental Brain Research*, *122*, 247–252.

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

Lloyd, D.M., Merat, N., McGlone, F. & Spence, C. (2003). Crossmodal links between audition and touch in covert endogenous spatial attention. *Perception & Psychophysics, 110*, 115–122.

Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., Schroeder, C.E. & Foxe, J. J. (2004). Grabbing your ear: Rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cerebral Cortex*, *15*, 963–974.

Spence, C. J, & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception & Performance, 22,* 1005-1030.

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

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

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

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

Zampini, M., Shore, D.I. & Spence, C. (2003a). Audiovisual temporal order judgments. *Experimental Brain Research, 152*, 198-210.

Zampini, M., Shore, D.I. & Spence, C. (2003b). Multisensory temporal order judgments: The role of hemispheric redundancy. *International Journal of Psychophysiology, 50*, 165-180.

Zampini, M., Brown, T., Shore, D.I., Maravita, A., Röder, B. & Spence, C. (2005). Audiotactile temporal order judgments. *Acta Psychologica, 118*, 277-291.

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

Chapter 5. The role of arms position in a tactile duration discrimination task

5.1 Introduction

The role of conflicting information from the different spatial frames of reference has been extensively studied in tactile perception. These spatial frames of reference can be the *anatomical* ‘coordinates’ (also known as somatotopic), which represent the median sagittal plane that defines the left and right coordinates of the body; as well as the *external* (also known as allocentric), which are the coordinates of the left or right side of a spatial configuration (Holmes, Sanabria, Calvert & Spence, 2005). Adopting a crossed arms posture has been employed as a way to understand how the brain deals with discrepant information provided by the anatomical and the external frames of reference when performing different tactile tasks that involve spatial judgments (such as tactile localization) as well as discrimination, for example (Holmes et al., 2005). This posture represents a puzzle to the brain, because it needs to update and solve the conflicting information brought up by both the anatomical and external referents, when localizing tactile stimuli for example (Azañón & Soto-Faraco, 2008).

One of the first reported studies exploring the effect of crossing the arms on temporal perception can be found in Drew’s (1896) pioneer work. The author designed a task which consisted on determining the temporal order of two tactile stimuli delivered to a finger from each hand in a non crossed and crossed arms position. The author showed that accuracy diminished when adopting a crossed arms position (Drew, 1896). This reduction in performance when crossing the hands is common and has been termed the ‘crossed hands effect’ (Holmes et al., 2005).

There is a body of work dedicated to study the crossed hands effect in the temporal perception of tactile events. Indeed, time is a very important variable when executing actions (mainly motor actions) which also involve the sense of touch (e.g. the action of tapping, movement, etc.; Studenka, Eliasz, Shore & Balasubramaniam, 2014).

Yamamoto and Kitazawa (2001) studied the role of arms posture on the perception of the temporal order of tactile stimuli, as way to know the moment when temporal order occurs, either earlier or after the stimuli are located in space (Yamamoto & Kitazawa, 2001). Since the crossed arms posture forces the brain to remap and update the new coordinates, the location of the stimuli in space has to occur, and is a process that takes extra time (in contrast to the uncrossed position). Therefore, this specific condition, needing more time overall, and needing this update in the localization of the new coordinates in space, is an opportunity that allows to observe whether temporal order judgments occur before or after the localization of stimuli has taken place in the crossed position, in contrast with the uncrossed arms position.

In their study, the authors had participants perform a task that consisted on deciding which stimuli (delivered sequentially to a finger in each hand) occurred earlier.

Their findings suggest that when participants had the crossed arms position, performance was worse when the interstimuli interval was inferior to 300 ms (i.e. 120, 200 ms). The authors propose that the remapping and update of the new spatial position of the stimuli in the brain begins after a 300 ms interval between the stimuli, to reach full re-establishment of the new spatial location after 1500 ms, which is reflected in a complete improvement of accuracy in performance (Yamamoto & Kitazawa, 2001). This meant that the errors were made in early interstimuli intervals because the location of the hands was not yet updated, and only happened after 300 ms. As the authors claim, the body has a default mode of assuming that the arms are not crossed, so with brief timing between stimuli, there is not enough time to update the system to the new crossed arms position (Yamamoto & Kitazawa, 2001). Therefore, some time is required to update the spatial location before the correct perception for temporal order perception can be attained. Therefore, these findings suggest that temporal order judgments can be done accurately only after the localization of the stimuli in the hands has been achieved in the crossed arms position (Yamamoto & Kitazawa, 2001).

Shore, Spry and Spence (2002) carried out a similar temporal order judgment task for tactile stimuli, and show the same pattern of findings, i.e., a decrement in performance when participants adopted the crossed arms posture (Shore et al., 2002). However, after an interval of 124 ms between the stimuli, participants' performance began improving for the crossed arms position (as shown in Experiment 1). This timing is dissimilar to the findings of Yamamoto and Kitazawa (2001), where improvement begins only at 300 ms. and later on. This difference in timing could probably reflect individual differences in the remapping process of localizing the stimuli, along with differences in the methods employed. For example, the various authors employed different ranges of stimuli onset asynchronies (SOA) along with very with different incremental steps (from 10 to 200 ms in Shore et al., 2002; from 10 ms to 1500 ms in Yamamoto & Kitazawa, 2001).

Another temporal aspect that has been studied in crossed arms paradigms is the study of temporal synchrony. Geffen, Rosa and Luciano (2000) addressed the effect of the crossed/uncrossed arms position in tactile simultaneity judgments, finding no differences in the simultaneity judgments according to the arms position. The question of why a crossed arms position affects temporal order judgment but not simultaneity judgments still remains unsolved and the possible explanations are still under debate. Some authors suggest that the temporal ordering and simultaneity judgements tasks are different in nature and require different processes (Hirsh & Sherrick, 1961; Shore et al., 2002; Yamamoto & Kitazawa, 2001). Hirsh and Sherrick (1961) propose that simultaneity judgments involve noticing whether two temporally separated events occurred, whilst temporal order judgments need further information, such as knowing which stimuli appeared first. Shore et al. (2002) even go further as to suggest that there might be different neural mechanisms involved in each kind of task.

Other studies have looked at the effect of interleaving the fingers of the two hands when adopting a crossed arms posture. For example, Zampini, Harris and Spence (2005) studied how this position affected the detection of the direction of trains of tactile stimuli delivered to adjacent fingers from the same hand or different hand, showing a decrement in performance when the hands were crossed and the fingers interleaved. The authors explained the crossed hands phenomenon as a result of a misalignment between the somatotopic and external frames of reference, when the direction of the stimulation (i.e. the sequence of vibrations, left-right, right-left) was opposite to the position of the hands (Zampini et al., 2005).

Other kinds of study have addressed the effects of crossing the arms in a cognitive advantage or ‘ability’ known as the “bimanual advantage”. This advantage is given when two fingers of the two different hands synchronously tap a surface, and the temporal variability in the execution of the tapping is reduced in contrast to tapping with a finger from one hand only (Studenka et al., 2014). Tapping with two hands rather than one supposes an advantage in terms of faster timing, since there might be an increment or addition of the sensory information obtained from the two hands to the timing system (Drewing & Aschersleben, 2003).

Findings suggested that crossing the arms increased the time variability when tapping with fingers from the two hands, as well as with only one, i.e. there were no statistical differences in terms of variability for the bimanual crossed hands and the unimanual condition, therefore the bimanual advantage is somehow lost when adopting a crossed arms position (Studenka et al., 2014).

Aglioti, Smania and Peru (1999) studied the effect of crossing the arms in patients with different degrees of tactile neglect or extinction, in an attempt to further understand if this type of lesions depends only on somatotopic areas or if it also relates to damage in higher-order areas that involve the integration of information coming from the spatiotopic frames of reference (i.e. somatotopic and extrapersonal space; Aglioti et al., 1999). In their tasks, participants needed to report whether they perceived a tactile stimulation in any of the hands or in both hands. Interestingly, participants with tactile extinction performed better when adopting the crossed arms posture in stimulation that was contralesional, than with the non-crossed arms posture (Aglioti et al., 1999). The authors conclude that both somatotopic information and spatiotopic is matched in the brain when perceiving somatic stimuli.

As previously described, there are studies that have addressed the role of arms posture in the perception of tactile temporal events, in a way to further understand how bodily information such as spatiotopic and somatotopic frames of reference are integrated and involved in the perception of tactile stimuli, among other motivations of study.

Specifically, the temporal aspects that have been addressed are the synchrony of tactile events (i.e. tactile stimulation; Geffen et al., 2000) as well as temporal order judgments (Drew, 1896; Shore et al., 2002; Yamamoto & Kitazawa, 2001). However, to the best of our knowledge, there are no reported studies exploring the effects of arms posture in the perception of tactile duration. The following experiment reported in this section, aimed to explore the role of adopting a non-crossed over a crossed arms position in a tactile duration discrimination task. The motivation behind this experiment was to test whether the crossing of the arms would impact the perception of the durations of tactile stimuli. Specifically, the intention was to check if the durations were perceived as longer in this condition, as a result of the updating and remapping process happening in the brain when adopting the crossed arms position. This process takes time, and it could affect the perceived duration of the stimuli processed when assuming this position.

5.2 Method

5.2.1 Participants

A total of 16 participants from the University of Trento took part in the experiment (9 F; mean age 26; age range from 21-38 years; 15 right-handed, 1 left-handed). All participants gave written informed consent prior to their inclusion in the study and took part on a voluntary basis, in exchange for course credit or reimbursement. They were not informed as to the purpose of the experiment, and all reported normal hearing and tactile sensitivity. The experiment was conducted in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki, as well as with the ethical guidelines laid down by the University of Trento.

5.2.2 Apparatus and stimuli

The tactile stimuli were delivered by means of two Oticons BC 461-1 (100 Ohm, Oticon, UK) bone conductor vibrators (1.4 cm X 2.4 cm), that were activated by two different soundcards (CREATIVE, USA). Auditory stimuli were delivered through headphones (Sennheiser HD 25-C11, Germany). White background external noise was sent to external loudspeakers (Dell A215, USA) located in a lower position from the monitor, parallel to the external corners of the computer used to deliver the task, and was controlled through an mp3 player (ipod Nano, Apple, USA). The experiment was delivered through a PC laptop (Dell Windows 7) using PsychoPhysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) for MATLAB R2012b. Tactile stimuli consisted on 200 Hz vibrations (0.085 N) that could either be delivered to the left or the right hand's index finger. The tactile stimulators (one for each hand) were placed on top of a foam rectangle located at a distance of 40 cm from the fixation point of the screen. The auditory stimulus consisted of a pure tone with a frequency of 200 Hz and a constant intensity of 50 dB (SPL). The auditory stimulus had a fixed duration of 240 ms. Durations of the tactile stimuli were variable and consisted of 200, 210, 230, 250, 270 and 280 ms. The screen appeared at a fixed distance of 85 cm from the participant

External white noise was presented to cover the sound made by the tactile stimulator with an intensity of 91 dB (SPL).

5.2.3 Design and Procedure

The design of the experiment consisted of a within participants factor of Arms Position (crossed/non crossed) and Duration (200, 210, 230, 250, 270 and 280 ms).

The task consisted on attending to a pair of stimuli presented on each trial. Each trial had an auditory constant duration stimulus, and a tactile stimulus that could have any of the durations mentioned previously. The stimuli within the trial were presented sequentially. Participants needed to decide which of the stimuli from the pair was the longest by pressing the corresponding key: 'S' for auditory or 'V' for vibration.

The order of stimulus presentation (i.e. either auditory or tactile) within a trial was counterbalanced across the blocks and presented randomly. The instructions were delivered through the computer screen and were also explained by the experimenters.

Participants sat in front of the computer and had to wear the headphones and place their left index finger over a tactile stimulator located at their left side, and the index finger of the right hand over the tactile stimulator located at their right side. Both tactile stimulators were attached to the fingers by a velcro tape. The finger pressure was not controlled mechanically but participants were asked not to make any pressure over the tactile stimulator. Each participant was tested individually in a dimly illuminated soundproof booth and was instructed to keep her/his gaze at the fixation point in the center of the screen. When crossing the arms, participants were requested to always put the right arm in top of the left one (for a clearer view of the arms postures and the location of the tactile stimulators, see Figure 21).

A total of 240 trials were introduced with 10 experimental blocks: 5 for the crossed and 5 for the non-crossed arms position, and were presented sequentially (e.g. crossed – non crossed – crossed – non crossed, etc.), however the order of the experimental blocks was counterbalanced across participants, so half of them started with the non-crossed arms position, and the other half with the crossed arms position. 120 trials were presented for the crossed arms position, and 120 trials for the non-crossed. A total of 40 trials for each of the tactile durations were presented along the experiment (each duration being repeated 4 times in each of the experimental blocks). Three different interstimulus intervals of 600, 650 and 700 ms were randomly introduced across trials as a way to avoid delays that the use of two soundcards could generate. Intertrial intervals had a duration of 1 second.

A brief pause was introduced at the beginning of each block, to allow the participant to change the posture of the arms, according to the type of block, and to allow the experimenter to change the position of the response keyboard and check the participant's arms position. The experiment took around 40 minutes to be completed. 8 practice trials were introduced to familiarize the participant with the experiment.

The activation of the tactile stimulators used to deliver the tactile stimulation (either the right hand one, or the left hand one) was counterbalanced across trials, but presented randomly within each block. The hand to respond with (either right or left) was also counterbalanced across participants, so half of the participants were asked to respond with the left one, and the other half with the right one, along the entire task. Participants responded with the thumb finger of the assigned hand. The response keyboard had to be changed in location for each block (crossed or uncrossed) in order to be always in proximity to the thumb of the corresponding hand that participants responded with.

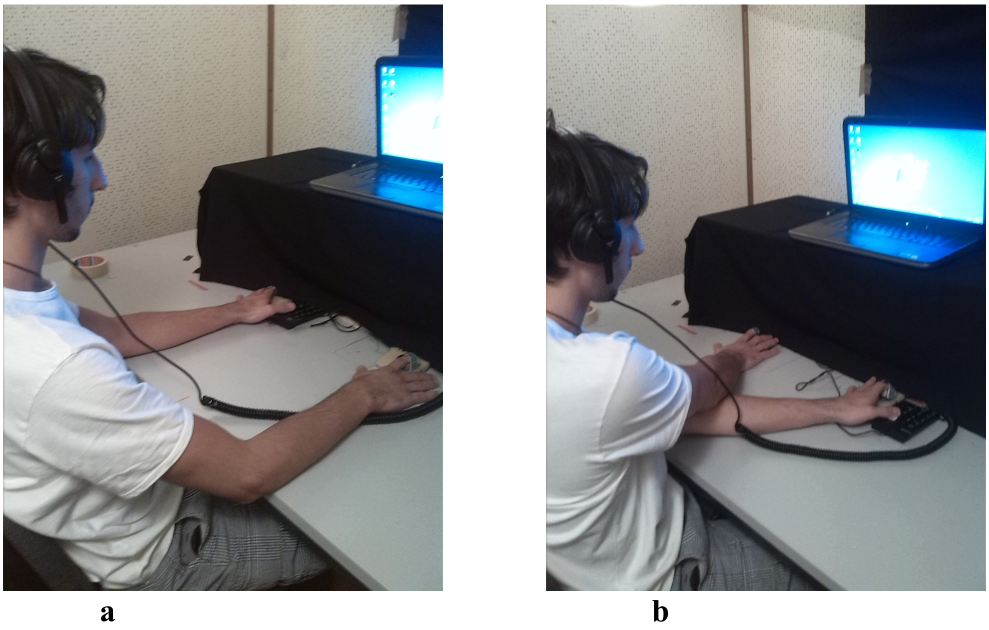


Figure 21. Image **a** illustrates the non-crossed arms position with the corresponding location of the tactile stimulator and response keyboard, whereas image **b** illustrates the crossed arms position also with the locations of the tactile stimulator and response keyboard.

5.3 Results and Discussion

5.3.1 Accuracy analysis

The proportion of correct responses was obtained for each individual in relation to the experimental conditions, and was used to calculate a repeated measures ANOVA with the Arms Position (Crossed/Non-crossed) and Duration (200, 210, 230, 250, 270 and 280 ms) as the within participants factors. The results show that the Arms Position factor did not show a significant main effect $F(1, 15) = .84$, NS, $p = .373$. However, the Duration factor provided a significant main effect $F(1.81, 27.16) = 7.34$, $p = .004$; Greenhouse Geisser corrected values are reported for this factor, since sphericity assumption was violated (Greenhouse

Geisser corrected, $X^2 = 60.15$, $p < .001$). Posthoc pairwise comparisons for the factor Duration, revealed the following directions: the duration of 200 ms (M = .742, SE = .035) provided better performance than the duration of 210 ms (M = .639, SE = .036, $p = .004$) and the duration of 230 ms (M = .506, SE = .035, $p < .001$); the duration of 210 ms had better performance than the duration of 230 ms ($p = .007$). Another significant comparison resulted from contrasting the duration of 250 (M = .592, SE = .036) with 280 ms, the latter showing better performance (M = .722, SE = .040, $p = .027$). Durations closer to the auditory constant referent of 240 ms, i.e. 230, 250 (M = .592, SE = .036) and 270 (M = .688, SE = .033) ms did not show significant differences between them in terms of performance. The rest of the comparisons did not show significant differences as well.

All post hoc pairwise comparisons were made using Bonferroni adjustments.

The interaction Arms Position x Duration did not show significance either $F(5, 75) = .48$, NS, $p = .785$.

For a clearer view of the results, see Figure 22 and Figure 23 below.

5.3.2 Point of subjective equality analysis

In order to test whether the crossed arms posture could generate an overall longer duration perception of the tactile stimuli, it was considered necessary to obtain the individuals' point of subjective equality (PSE) to check whether a shift towards longer durations was produced in the crossed hands condition, when comparing it to the non-crossed posture. In the psychometric function, the PSE represents the stimulus magnitude at which a stimulus appears to be equal to the magnitude of a standard or another stimulus (Kingdom & Prins, 2009). In other words, the PSE would represent the duration were the responses are equally probable to be considered as long or short.

If there were any influence of the crossed arms position, the PSE would be shorter in this condition in contrast to the non-crossed arms position.

For this analysis, the proportion of tactile longer responses was computed to calculate a psychometric fitting for each participant, in the crossed and the non-crossed arms condition. A logistic function was used to calculate the fittings, using the following formula:

$$F_L(x;\alpha,\beta) = \frac{1}{1 + \exp(-\beta(x - \alpha))}$$

Where x represents the tactile stimulus duration, β establishes the value of the slope and α corresponds to the threshold: F L (x = α ; α , β). The data was computed through Palamedes Toolbox in MATLAB (Prins & Kingdom, 2009).

The psychometric function fitting analysis provided a PSE value for each individual, according to the crossed and non-crossed arms condition. Since the fitting was not successful for all participants (we discarded 7 participants), only the data for 9 participants was employed.

The PSE values were then submitted to a paired samples t test, to contrast any possible difference between the PSE values for the crossed and non-crossed arms position. The findings suggested no differences in terms of the PSE values between the crossed arms (M = 224, SD = .006) and the non-crossed arms position (M= 224, SD = .011, t (8) = -.013, *p* = .990). The mean values indicate the duration in milliseconds.

This suggests that the tactile durations were not perceived as longer, when adopting the crossed arms position, thus this posture did not seem to affect the perception of duration.

The current findings indicate that performance was overall successful despite the crossed arms condition. The only decrement in performance seems to be happening in the tactile durations that are closer to the standard auditory stimulus, (e.g. the case for the lowest performance was seen at 230 ms, see Figure 22). There was not a particular hypothesis in relation to the factor duration, but we only expected performance to be worse when the durations were closer to the auditory standard referent.

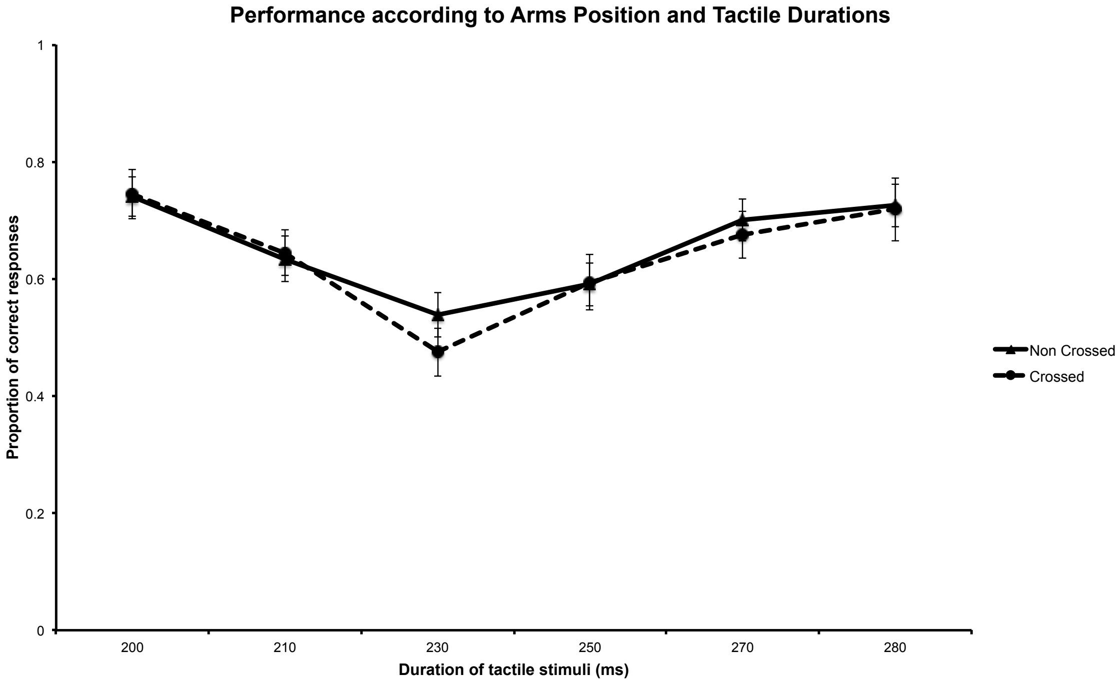


Figure 22 Performance according to the Arms Position factor and the different Tactile durations. The black straight line indicates performance for the Non-Crossed arms position. Dotted line indicates performance for the Crossed arms. Error bars indicate standard errors. As the pattern reveals, there are no differences in performance for the crossed and non-crossed arms posture.

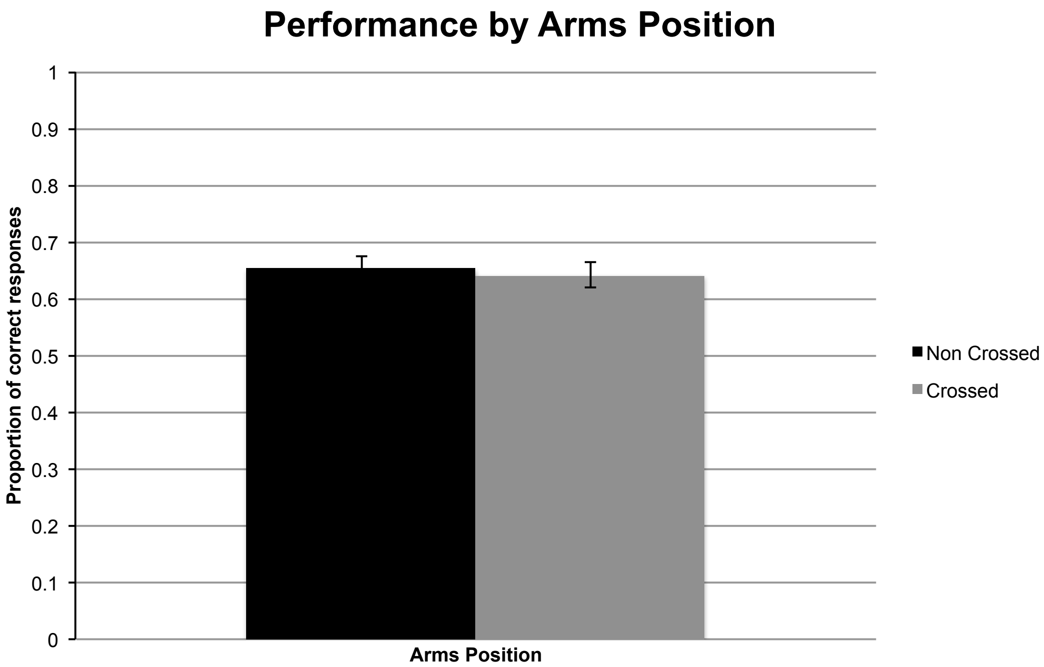


Figure 23. Mean performance for the Non-Crossed and Crossed arms position. The bars suggest no difference in terms of performance. Error bars indicate standard errors.

The hypothesis of having longer perception of tactile durations as a consequence of adopting a crossed arms posture could not be confirmed in this experiment. Despite the idea that the remapping process implies longer time for updating the spatial information, this did not seem to affect duration perception.

Evidence has shown deficits in performance for certain tactile temporal tasks such as temporal order judgments, when adopting a crossed arms posture (Drew, 1896; Shore et al., 2002; Yamamoto & Kitazawa, 2001). However, for the current set of data, this was not the case. It is true though that, for example, in Yamamoto and Kitazawa's (2001) findings, performance begins to improve after the intervals between the tactile stimuli grow from 300 ms onwards, in their temporal order judgment task; Shore et al. (2002) report this improvement, but in a shorter interval (124 ms). Even though the task carried out in this experiment is from a different nature, and involves duration perception discrimination, not temporal order judgments, it could be the case that the temporal sequence between the stimuli to discriminate duration was wide enough as to allow participants to update their position and their frames of reference as Yamamoto and Kitazawa

(2002) demonstrate; in fact this interval ranged form 600 -700 ms, therefore, the crossed arms posture did not impede participants to successfully discriminate between the different durations. However, it is also worthwhile considering that for practical purposes, participants changed the arms posture from block to block only. It could have been the case that they got used to the crossed arms posture quite fast, allowing successful discrimination in both arm postures (although we tried to keep the blocks as short as possible in order to avoid this kind of effect).

The pattern of results of the reported here experiment is more in line with the findings from synchrony perception tasks, where its judgments were not affected by the crossed arms posture (Shore et al., 2002). Indeed, duration discrimination is a different task that does not share much with synchrony or TOJ tasks, except for being temporal tasks. Therefore, the similarity in the pattern of results with the synchrony data is still unclear, and could be coincidental.

Another possibility could lie in the fact that perhaps the task was too easy or predictable. As a matter of fact, some participants reported noticing that the auditory stimulus was always the same, so this might have affected the accuracy, making duration discrimination more efficient.

To the best of our knowledge, this is the first task exploring the temporal aspect of duration in this kind of paradigm and helps further understand how duration discrimination is affected (or not) by adopting a posture different from the default mode. The arms posture seemed an interesting factor to explore, as a way to provide more evidence for the study of duration, present in this thesis.

References

Aglioti, S., Smania, N. & Peru, A. (1999). Frames of reference for mapping tactile stimuli in brain-damaged patients. *Journal of Cognitive Neuroscience, 11*, 67-79.

Azañón, E. & Soto-Faraco, S. (2008). Changing reference frames during the encoding of tactile events. *Current Biology, 18*, 1044-1049.

Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Drew, F. (1896). Attention: Experimental and Critical. *The American Journal of Psychology, 7*, 533-573.

Drewing, K., & Aschersleben, G. (2003). Reduced timing variability during bimanual coupling: A role for sensory information. *Quarterly Journal of Experimental Psychology, 56A*, 329–350.

Geffen, G., Rosa, V. & Luciano, M. (2000). Effects of preferred hand on the perception of tactile simultaneity. *Journal of Clinical and Experimental Neuropsychology, 22*, 219-231.

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

Holmes, N.P., Sanabria, D., Calvert, G. & Spence, C. (2006). Multisensory interactions follow the hands across the midline: Evidence from a non-spatial visual-tactile congruency task. *Brain Research, 1077*, 108-115.

Kingdom, F.A.A. & Prins, N. (2010). *Psychophysics: A practical introduction*. London, UK: Academic Press.

Kleiner, M., Brainard, D., & Pelli, D.(2007). "What's new in Psychtoolbox-3?" *Perception, 36*, 1–16.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Prins, N., & Kingdom, F.A.A. (2009). Palamedes: MATLAB routines for analyzing psychophysical data. http://www.palamedestoolbox.org .

Shore, D.I., Spry, E. & Spence, C. (2002). Confusing the mind by crossing the hands. *Cognitive Brain Research, 14*, 153-163.

Studenka, B.E., Eliasz, K.L., Shore, D.I. & Balasubramaniam, R. (2014) Crossing the arms confuses the clocks: Sensory feedback and the bimanual advantage. *Psychonomic Bulletin Review, 21*, 390-397.

Yamamoto, S. & Kitazawa, S. (2001). Reversal of subjective temporal order due to arm crossing. *Nature Neuroscience, 4*, 759-765.

Zampini, M., Harris, C. & Spence, C. (2005). Effect of posture change on tactile perception: impaired direction discrimination performance with interleaved fingers. *Experimental Brain Research, 166*, 498-508.

Chapter 6. Hierarchies between the senses of Audition, Vision and Touch in a duration perception task

6.1 Introduction

6.1.1 The senses of audition, vision and touch in the perception of duration

The perception of time can be accomplished through the information that comes from our different senses, as has been described in previous chapters. A body of work in the field of multisensory integration has dedicated the study of temporal perception through multisensory interactions.

In particular, duration perception is a time feature that has been explored in audiovisual combinations (Chen & Yeh, 2009; Klink, Montijn & van Wezel, 2011; Ortega et al, 2014; van Wassenhove, Buonomano, Shimojo & Shams, 2008; Walker & Scott, 1981) and visuotactile pairings (Rach & Diederich, 2006; Tomassini et al., 2011).

The vast majority of studies on temporal duration usually report audiovisual combinations and comparisons. Some of these demonstrate the phenomenon of auditory elongation, that is, the fact that an auditory stimulus is always perceived as longer than a visual with the same physical duration (Penney et al., 2000; Walker & Scott, 1981). It has also been demonstrated that for duration discrimination tasks, audition is usually a better modality than vision in terms of accuracy (Grondin, Meilleur-Wells, Ouellete & Macar, 1998; Grondin, 2003).

Some studies demonstrate the dominance of audition over vision, but not the opposite (i.e. vision dominating audition; Chen & Yeh, 2009; Klink et al., 2011). For example, Chen and Yeh's study documents this unilateral auditory influence in an oddball paradigm where participants contrasted a stimulus with a fixed duration (standard duration) to the duration of an oddball stimulus that could be presented either auditorily or visually, and was simultaneously presented with a distractor stimulus in the other modality. The distractor stimulus had the same duration of the oddball. The authors calculated an expansion ratio, (a measure obtained from dividing the standard duration by the point of subjective equality (PSE), as an indicator of the degree of time expansion. This ratio was used to contrast the unimodal to the bimodal conditions, and found that it was higher for the visual modality, whereas the auditory modality presented no differences. In this way, the authors concluded that the auditory modality dominated the perception of the visual modality, however the visual modality did not alter the auditory perceived duration (Chen & Yeh, 2009). Instead, van Wassenove, Buonomano, Shimojo and Shams (2008) demonstrated the opposite effect, where vision dominated the auditory perception of duration stimuli. In their tasks, they presented a series or trains of 5 durations in a target modality (that could be visual or auditory) and the fourth duration consisted on the duration to attend (target) within a trial. The target was presented simultaneously with an oddball stimulus in a distractor modality (i.e. visual or auditory, correspondingly). Participants had to report if the target duration was shorter or longer than the other target stimuli, within the train of stimuli. There were unimodal conditions for both modalities, bimodal congruent (where the oddball coincided with the duration of the target and bimodal incongruent (when the duration of the oddball was different from the target).

Perhaps some of the reasons for the obtained differences in the works of Chen and Yeh (2009) and van Wassenhove et al. (2008) can lie in methodological differences employed in their paradigms, such as the task design, the type of stimuli, the ranges of durations and more importantly, the experimental conditions employed (for example, Chen and Yeh did not include an incongruent condition in their tasks).

Walker and Scott (1981) however, document a number of cases of auditory dominance over visual duration judgments, but also the opposite in only one condition, where the auditory stimulus had a lower intensity, resulting in a dominance of vision on the perceived auditory duration (Experiment 3).

There is a significant smaller number of studies addressing visuotactile combinations in the perception of duration. Rach and Diederich (2006) presented a visuotactile task where participants had to detect visual stimuli, whilst ignoring a series of tactile distractors (which had short and long durations). With analysis of reaction times, the authors concluded the presence of facilitation effects for short tactile distractors. The findings were discussed in relation to an effect they termed ‘the inverse effectiveness of duration’, where the short durations might have generated an enhancement for the bimodal presentations (Rach & Diederich, 2006).

The study of Tomassini et al. (2011) also documents visuotactile interactions in a duration reproduction task. When comparing performance for the tactile and visual modalities, the authors found that duration reproduction was better and more precise when using the tactile modality than the visual modality. The tactile modality results showed duration reproductions closer to the actual physical duration than for the visual reproductions. The authors also found that there were no advantages of bimodal combinations (visuotactile pairings) in contrast to unimodal comparisons (Tomassini et al., 2011).

6.1.2 Studies of temporal perception involving the three modalities: audiovisual, audiotactile, and visuotactile

6.1.2.1 Synchrony perception, temporal order judgments and simultaneity

Fujisaki and Nishida (2009) conducted a series of studies in order to compare the temporal resolution or accuracy of audiovisual, visuotactile and audiotactile pairings in the perception of synchrony.

In their experiment, pairs of bimodal stimuli (with the combinations mentioned above along with a tactile only condition) were presented synchronously o asynchronously. The stimuli could be in the manner of repetitive pulse trains or single pulses (Experiment 1). Participants needed to respond whether the presented stimuli were in synchrony or in asynchrony. The findings indicated the following hierarchy in temporal resolution: audiotactile combinations showed the best temporal resolution, followed by visuotactile, and lastly, performance was worst for the audiovisual combination. Nevertheless, tactile bimodal stimuli (i.e. two tactile stimuli) and audiotactile pairings showed similar accuracy in performance (Fujisaki & Nishida, 2009). Experiment 4 tested the same modality pairings in a further simultaneity judgment task, along with temporal order judgments of the above-mentioned pairings.

The findings replicated the same hierarchy for the simultaneity judgments: the audiotactile stimuli showed a better temporal resolution than visuotactile and audiovisual stimuli, overall. In the case of temporal order judgments however, further comparisons revealed no significant differences in between audiotactile and audiovisual pairings, in terms of just noticeable difference values (Fujisaki & Nishida, 2009).

6.1.2.2 Duration perception of empty intervals

In an interest to test contextual effects on temporal perception, Burr, Della Rocca and Morrone (2013) conducted a study to see the effects of distractor duration stimuli in the discrimination of duration. The authors employed the auditory, tactile and visual modalities in short and long empty intervals. Through a two alternative forced choice task, in most of the cases, both the distractor intervals and the test duration intervals (to be discriminated) were presented in the same modality. Also, under some conditions, the distractor intervals were introduced in asynchrony with the test trials. Participants needed to compare the durations of the stimuli from the task, whilst ignoring the distractor.

There was, however, a task relevant to the experiment proposed in this chapter. The task involved a synchronous bimodal presentation of a distractor duration in a modality different from the test task, along with a duration of the task modality. The task involved only auditory and tactile stimuli that acted both as distractor or test durations, and vice versa. The findings indicated that audition as distractor affected the visual judgments in a greater extent than the visual, the latter influencing only the auditory duration discrimination judgments on half of the times (Burr et al., 2013). However, their findings are only preliminary, since the obtained data comes from only one participant.

This provides additional evidence as to the persistent influence of the auditory modality over different modalities in temporal perception.

To the best of my knowledge, there are no reported studies exploring how the modalities of audition, vision and touch influence one another, altogether in the in the perception of duration. The following experiment reported here, documents the influence of audition, vision and touch in a duration discrimination task, when these modalities served both as targets and as distractors, correspondingly. This was accomplished by means of 2AFC tasks, where a modality served as target, whilst the other two modalities served as distractors. Therefore, there was a total of three tasks, with the auditory as target (with vision and touch as distractors), vision as target (with audition and touch as distractors), and tactile as target (with audition and touch as distractors), in a within participants design.

Two experimental questions motivated this experiment. One was to establish the hierarchies of the interactions between the three senses in a duration perception paradigm, with an interest to find (based on previous literature) whether auditory unimodal performance was better/equal to the tactile performance (as suggested by the results found in Chapter 3, experiment 2), and higher than the visual in a duration perception task. In sum, to test whether:

$$\text{Unimodal Auditory} \geq \text{Unimodal Tactile} > \text{Unimodal Visual}$$

The second question consisted on finding out the hierarchy of influence from the different distractor modalities over the given targets. For example, taking into account the previous documented close interactions between audition and touch, it was our interest to find out whether touch would be more disruptive (incongruent condition) or facilitatory (congruent condition) as a distractor, than the visual modality when audition was the target modality, along the different experimental conditions.

6.2 Method

6.2.1 Participants

A total of 12 participants, from the University of Trento took part in this experiment (10 F, mean age 21.4 years, range from 19 to 26 years), in return for course credit exchange or reimbursement. They participated in three separate sessions, with one-week distance separation from each of the sessions. They were not informed as to the purpose of the experiment, and gave written informed consent prior to the beginning of the experiment. The experiment was conducted in accordance to the standards from the 1964 Declaration of Helsinki, as well as the ethical standards from the University of Trento. All participants reported normal hearing, seeing and tactile sensitivity.

6.2.2 Apparatus and stimuli

The equipment used to deliver the tactile and auditory stimuli, was the same as in the previous reported experiments. Auditory stimuli consisted of pure tones with a frequency of 200 Hz and a constant intensity of 50 dB (SPL). Tactile stimuli consisted on 200 Hz vibrations (0.085 N) delivered to the index finger of the left hand (the auditory and tactile stimuli were the same adopted in previous experiments). In this experiment, the auditory stimuli were delivered via headphones. The visual stimuli were delivered through a small LED lamp with a red light emitting diode.

White background noise was presented to cover the sound coming from the tactile stimulators and the loudspeaker that generated the light (this loudspeaker was hidden from view). The white noise was controlled through a generic mp3 player with an intensity of 85.6 dB (SPL) and was played through 2 external loudspeakers (Dell A215, USA), located at the corners of the experimental table.

The LED lamp was attached to the center of the monitor used to deliver the instructions and the overall experiment, and was located at a distance of 40 cm away from the participant. The lower part of the monitor was concealed with a black cloth covered cardboard and served as a structure to support the LED lamp. A small response keyboard was placed on the right side of the participant, close to the monitor. The tactile stimulator was at a fixed position (located on the left side of the participant), at a distance of 26 cm from the fixation point on the monitor and within a 45 cm distance from the participant. For a clearer view of the experimental setup, see Figure 24.

Figure 24. A view of the experimental setup

The experiment was delivered through a PC laptop (Dell Windows 7) using PsychoPhysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) for MATLAB R2012b.

As in previously reported experiments, the tactile, auditory and visual stimuli were presented in a fixed constant duration, that is, a standard duration, as well as in different probe durations. The standard duration was of 240 ms, whereas the probe durations consisted of 200, 210, 230, 250, 270 and 280 ms each, so they could be either shorter or longer than the standard duration.

6.2.3 Design and Procedure

The experiment consisted of a within participants design with the factors of:

a) TargetDistractor combination: AudioVisual (AV), AudioTactile (AT), VisuoTactile (VT), VisuoAuditory (VA), TactileAuditory (TA), TactileVisual (TV)
b) Stimulus Presentation (Bimodal Congruent, Bimodal Incongruent, Unimodal)
c) Duration (200, 210, 230, 250, 270 and 280 ms).

A 2AFC task was employed for each of the three possible modality targets, therefore a total of three tasks. As in the within participants task from chapter two, the design and procedure of the task was identical with two exceptions:

Instead of having only one modality distractor, two modality distractors were employed, for each of the target modalities. For example, when the auditory modality was the target, the task was composed of the visual and tactile modalities as distractors.

The distractor information in the bimodal incongruent condition presented the standard durations only, in both events of the pair of stimuli from each trial. This is different from the previous procedure, where the distractor information had a standard and a probe duration, in a trial. The motivation behind this change, was to see whether there were elongation/shortening effects of the perceived target duration, according to the duration of the distractor modality.

The procedure used for this experiment was similar to the one used in Experiment 2, (Chapter 2). A bimodal trial consisted on a simultaneous presentation of a pair of events. Each event, consisted of a simultaneous presentation of a target stimulus (in a given modality) and a distractor stimuli (in another modality). The durations of the target and distractor modalities were the same in the congruent condition (i.e. a probe and a standard duration). For the incongruent condition, the target had a probe and a standard duration, while the distractor only kept the standard durations (as explained above). Unimodal trials consisted on a presentation of a pair of events (probe and standard durations) within a single modality. Order of probe-standard duration was counterbalanced along each block.

Participants had to come to three separated sessions. Each session consisted on attending to only one target (either auditory, tactile, or visual) whilst ignoring the corresponding modality distractors. The orders of target completion along with each of the Target-Distractor combinations were counterbalanced across all participants.

In each session, participants completed one experimental block with a target and one type of modality distractor. When finished, the block with the other modality distractor followed. Each block with a Target-Distractor combination consisted on a total of 180 trials, so a total of 360 trials were completed for each session (and a total of 1080 trials for the whole experiment). Each duration under each of the stimulus presentation conditions, was repeated a total of 10 times for each of the blocks. A practice session was introduced for each of the Target-Distractor combinations, with a total of 12 trials each.

A fixed pause of one minute was introduced in the middle of each block (TargetDistractor combination). After each block was finished, participants were asked to briefly leave the experimental room, to allow the experimenter to manually change the apparatus connections, in order to introduce the new block with the new TargetDistractor combination.

The three sessions were conducted in a sound attenuated booth with no illumination, to allow participants to see the emitted light from the LED lamp.

6.3 Results and Discussion

The proportions of correct responses were obtained for each of the participants, according to the different experimental conditions.

A series of different analyses were performed to explore the data.

However, before describing the data and analyses, a note to data interpretation is added here. First, it is important to clarify that the factor Duration, and the corresponding interactions involving this factor will not be addressed in the current discussion of the data. The reason behind this lack of explanation is that we did not have a particular hypothesis for the different levels of the duration factor, therefore we do not consider necessary to address this issue, added to the fact that its omission from the discussion allows data simplification.

Also, the interactions of Distractor x Stimulus presentation are described in detail in the third analysis, since the findings are similar to the first analysis. This is to allow data simplification and repetition of the effects.

6.3.1 First analysis. Individual ANOVAS for each of the target modalities

In order to have an initial inspection of the data, individual repeated measures ANOVAS were performed for each of the modality targets. The ANOVAS with their corresponding factors are listed below:

Auditory target: Distractor (2) x Stimulus presentation (3) x Duration (6)
Visual Target: Distractor (2) x Stimulus presentation (3) x Duration (6)
Tactile Target: Distractor (2) x Stimulus presentation (3) x Duration (6)

6.3.1.1 Auditory Target results

Despite the fact that the factor ‘Stimulus presentation’ resulted significant, pairwise comparisons of this main effect showed no significant differences for the performance in the bimodal congruent (M = .730, SE = .030, NS.), bimodal incongruent (M = .691, SE = .034, NS.), and unimodal conditions (M = .732, SE = .031, NS.) as seen from Table 4 .For a clearer view of the data, see Figure 25.

Table 4. Summary of the data from the repeated measures ANOVA for Audition as target.

Factor	gL	F	*p*	Notes
Distractor	1, 11	1.022	.33	
Stimulus presentation	2,22	3.616	.044	*
Duration	5,55	28.737	.000	**
Distractor * Stimulus presentation	2, 22	3.058	.067	NS
Distractor * Duration	5, 55	.651	.662	NS
Stimulus presentation * Duration	4.84, 53.28	2.058	.034	* GGC (X^2 = 80.37, *p* =.03)
Distractor x Stimulus presentation * Duration	10, 110	1.770	.074	NS

** Indicate significance at p <.001. GGC indicates Greenhouse Geisser corrected values

Despite the fact that the interaction Stimulus presentation x Distractor was marginally significant (*p* = .067), it was considered relevant to continue with the corresponding posthoc comparisons for explorative reasons. The posthoc comparisons are detailed next:

For the At combination (i.e. auditory as target, tactile as distractor), the bimodal incongruent condition had the lowest performance (M = .654, SE = .033) in comparison to the bimodal congruent condition (M = .732, SE = .039, *p* = .009). Also, contrasts revealed a marginal significant difference between the bimodal incongruent and the unimodal condition, being the incongruent the lowest accurate, as opposed to the unimodal condition (M =.729, SE = .033, *p* = .069).

Regarding the Av combination (i.e. auditory as target, vision as distractor), results showed no significant differences between the congruent (M = .238, SE = .035) incongruent (M = .728, SE = .037, NS *p* = 1) and unimodal condition (M = .735, SE = .035, NS, *p* = 1) in terms of performance.

Also, performance was significantly better (however, marginally significant) in the incongruent condition when the distractor was the visual (M = .728, SE = .037), than when the distractor was the tactile (M = .654, SE = .039, *p* = .056). This finding could perhaps suggest that touch as a distractor, interfered more with audition, than vision as a distractor, thus suggesting that it was probably harder to disentangle the duration information between audition and touch, which is consistent with previous literature that suggests that audition and touch share close links (Occelli et al., 2011; Soto-Faraco & Deco, 2009; Von Békésy, 1959)

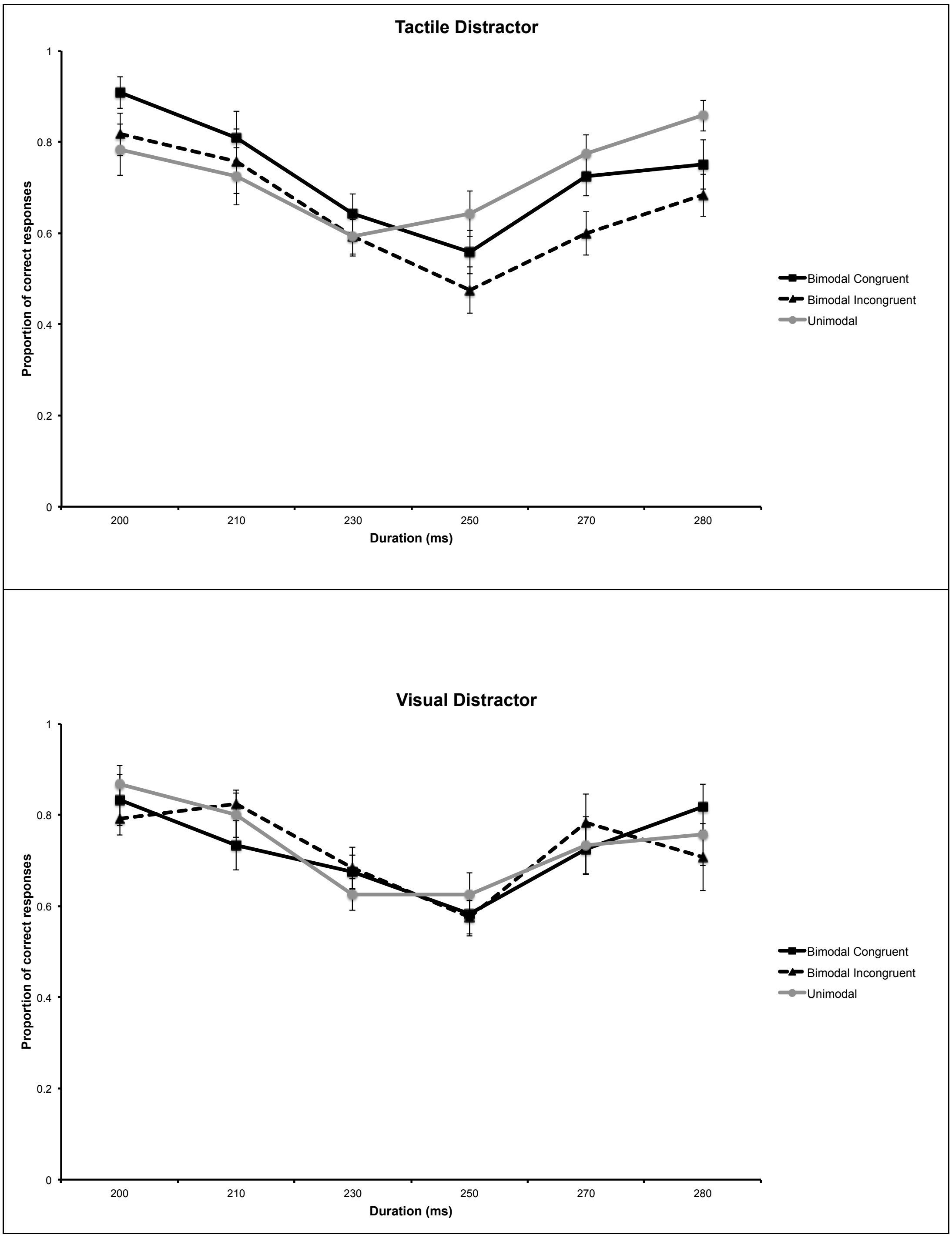


Figure 25. Graphs of Performance according to Distractor, Stimulus Presentation and Duration for the Auditory Target

6.3.1.2 Visual target results

Table 5 summarizes the findings from the visual as target modality. Posthoc pairwise comparisons for the main effect of Stimulus presentation revealed that performance was better in the bimodal congruent condition (M =

.682, SE = .028, and significantly higher than the unimodal (M = .583, SE = .021, p = .001) and the incongruent condition (M = .562, SE = .018, p = .008). However, unimodal and incongruent were not significantly different from one another, which is different from the findings of the experiments reported in previous chapters. For a clearer view of the data, see Figure 26.

Table 5. Summary of the data from the repeated measures ANOVA for Vision as target.

Factor	gL	F	*p*	Notes
Distractor	1, 11	.111	.745	NS
Stimulus presentation	1.25, 13.76	15.228	.000	GGC (X^2 = 9.12, p = .01)
Duration	5, 55	4.974	.001	**
Distractor * Stimulus presentation	2, 22	5.201	.014	*
Distractor * Duration	5,55	1.505	.203	NS
Stimulus presentation * Duration	10, 110	3.395	.001	**
Distractor * Stimulus presentation * Duration	10, 110	.742	.684	NS

** Indicate significance at p <.001. GGC indicates Greenhouse Geisser corrected values

The interaction Stimulus presentation x Distractor was significant and relevant for our experiments, therefore poshtoc comparisons were made, revealing that for the Va combination (i.e. vision as target, audition as distractor), performance was overall better for the congruent condition (M = .712, SE = .034) in contrast to the incongruent (M = .554, SE = .022, p = .007) and the unimodal (M = .554, SE = .022, p = .000). However, the latter two were not significantly different from one another (NS, *p* = 1).

Comparisons for the Vt combination (with touch as distractor) showed a marginally significant difference in performance according to the stimulus presentation factor between the congruent performance (M = .651, SE = .026, NS) and the incongruent performance (M = .569, SE = .021, *p* = .056). However, the unimodal condition was not significantly different from the incongruent or the congruent condition (M = .597, SE = .023, NS).

Interestingly, this significant interaction (i.e. Stimulus presentation x Distractor) could be also driven by the fact that performance in the congruent condition was significantly better when the distractor was the auditory (M = .712, SE = .034), than when the distractor was the tactile (M = .651, SE = .026, *p* = .035). With this pattern of results, one might wonder why was audition a better enhancer for the visual modality, in contrast to the tactile modality and whether this suggests that audition is better overall for temporal perception, and whether this leadership is driving a better performance for the visual modality as target.

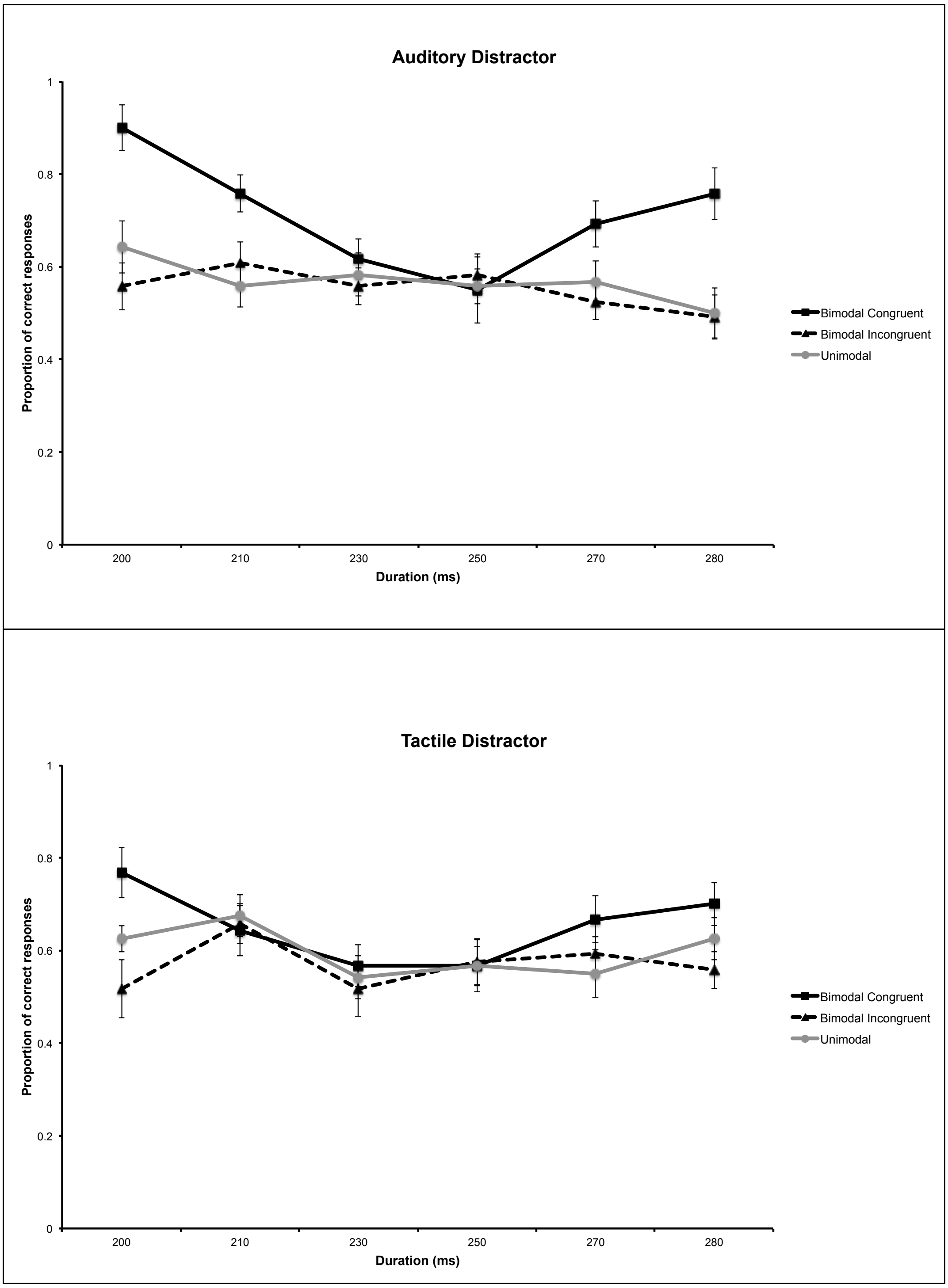


Figure 26. Graphs of Performance according to Distractor, Stimulus Presentation and Duration for the Visual Target

6.3.1.3 Tactile target results

In Table 6, it is possible to find the summary of the results. Even though the fact that the factor ‘Stimulus presentation’ showed significance, pairwise comparisons of this main effect revealed no significant differences for the performance in the bimodal congruent (M = .684, SE = .021), bimodal incongruent (M = .647, SE = .017), and unimodal conditions (M = .687, SE = .017). For a clearer view of the data, see Figure 27.

Table 6. Summary of the data from the repeated measures ANOVA for Touch as target.

Factor	gL	F	*p*	Notes
Distractor	1, 11	.180	.679	NS
Stimulus presentation	2, 22	4.251	.027	*
Duration	5, 55	27.256	.000	**
Distractor * Stimulus presentation	2, 22	4.443	.024	*
Distractor * Duration	2.45, 26.96	.466	.800	NS GGC (X^2 = 25.19, p = .03)
Stimulus presentation * Duration	10, 110	2.740	.005	*
Distractor * Stimulus presentation * Duration	10, 110	1.376	.201	NS

** Indicate significance at p <.001. GGC indicates Greenhouse Geisser corrected values

The posthoc pairwise comparisons of the Stimulus Presentation x Distractor interaction showed that for the Ta combination (i.e. tactile target, auditory distractor), performance was overall worse for the incongruent condition (M = .611, SE = .023) in contrast to the congruent (M = .704, SE = .024, *p* = .016) and the unimodal (M = .692, SE = .024, although marginally different, p = .063). The congruent and unimodal conditions did not show significance (NS, p = 1).

Despite that some of the interactions were marginally significant in this target, the pattern of results is in line with findings from the second chapter (Experiment 2).

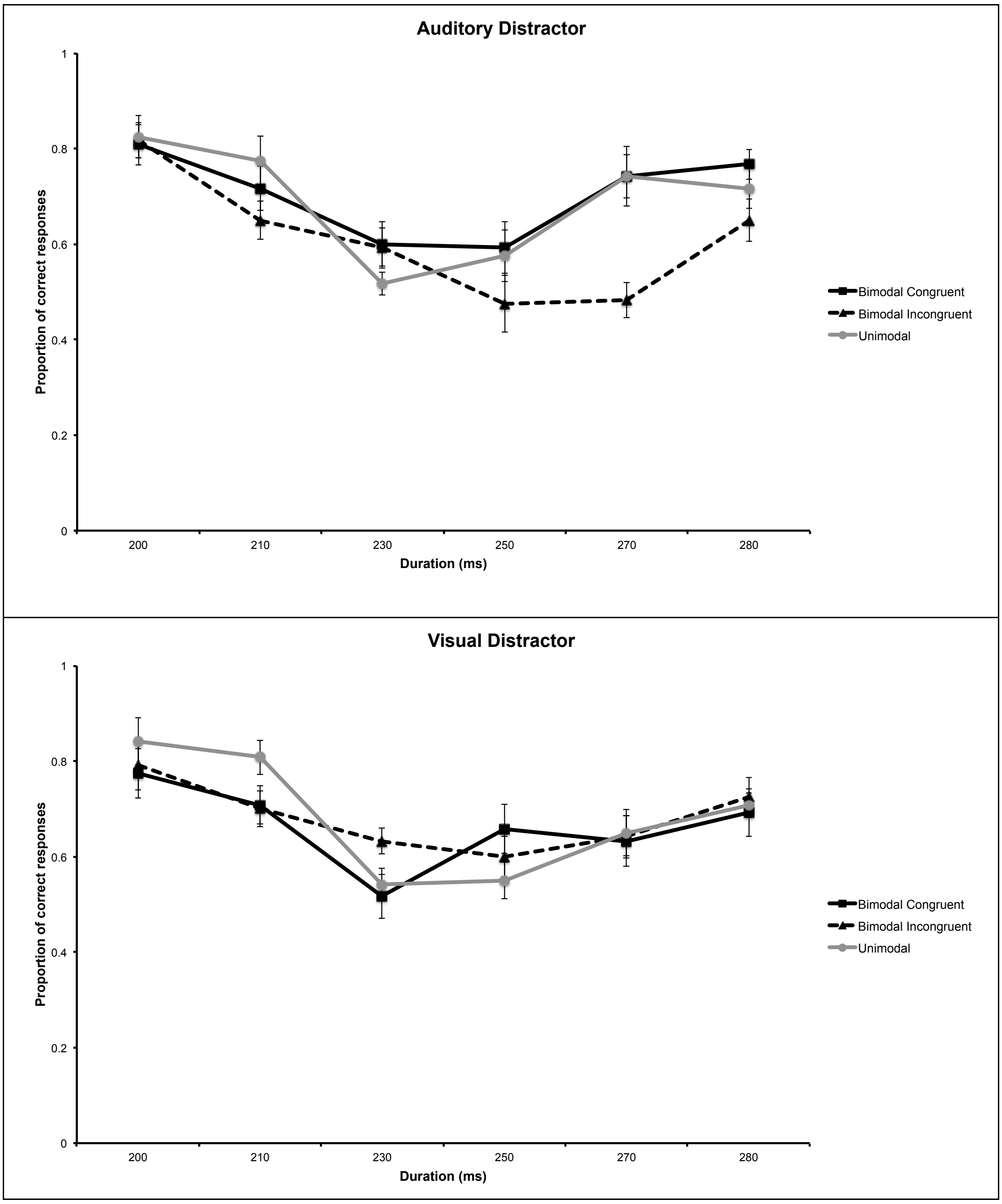


Figure 27. Graphs of Performance according to Distractor, Stimulus Presentation and Duration for the Tactile Target

In the Tv combination (with vision as distractor), there were no differences between the congruent (M = .664, SE = .023), incongruent (M = .682, SE = .021) and unimodal conditions (M = .683, SE = .024, NS, p = 1.).

The interaction can be also partly driven by the fact that performance in the incongruent condition was significantly better when the distractor was the visual (M = .682, SE = .021), than when the distractor was the auditory (M = .611, SE =

.023, *p* = .029). This could suggest that the auditory modality as a distractor interfered more greatly with the tactile target, than the visual distractor, suggesting again that it becomes harder to separate the duration information from both audition and touch. This finding is similar to the pattern of results reported for the auditory target, and could be additional evidence as to the close interactions and connections that these two modalities share with one another.

6.3.2 Second analysis. Comparison of unimodal performance across the target modalities

This analysis was considered pertinent as a way to explore each individual modality performance in duration discrimination. The analysis aimed to know which modality was better or more accurate when having to discriminate between stimuli with different duration. For this purpose, only the unimodal responses were employed, along with the different durations employed in the task. A within participants ANOVA with the factors of Target modality (Auditory, Visual and Tactile) and Duration (200, 210, 230, 250 270 and 280 ms) was run to check for this hierarchy of the senses when perceiving duration.

Table 7. Summary of the data from the repeated measures ANOVA Target modality and Duration.

Factor	gL	F	p	Notes
Target	2, 22	22.459	.000	**
Duration	5, 55	17.792	.000	**
Target * Duration	10, 110	3.826	.000	**

** Indicate significance at p <.001

Table 7 summarizes the data found for the unimodal comparisons. Posthoc pairwise comparisons of the factor Target modality revealed that performance in the unimodal condition was overall better for the Auditory (M = .732, SE = .031) and Tactile modalities (M = .687, SE = .017) in contrast to the Visual modality, which had a significantly lower performance (M = .583, SE = .021, p <.001). There were no significant differences between the Auditory and Tactile modalities, thus suggesting that duration discrimination is easier or more accurate when any of these two modalities present duration information. For a clearer view of the performance, see Figure 28.

All posthoc comparisons were made with Bonferroni corrections.

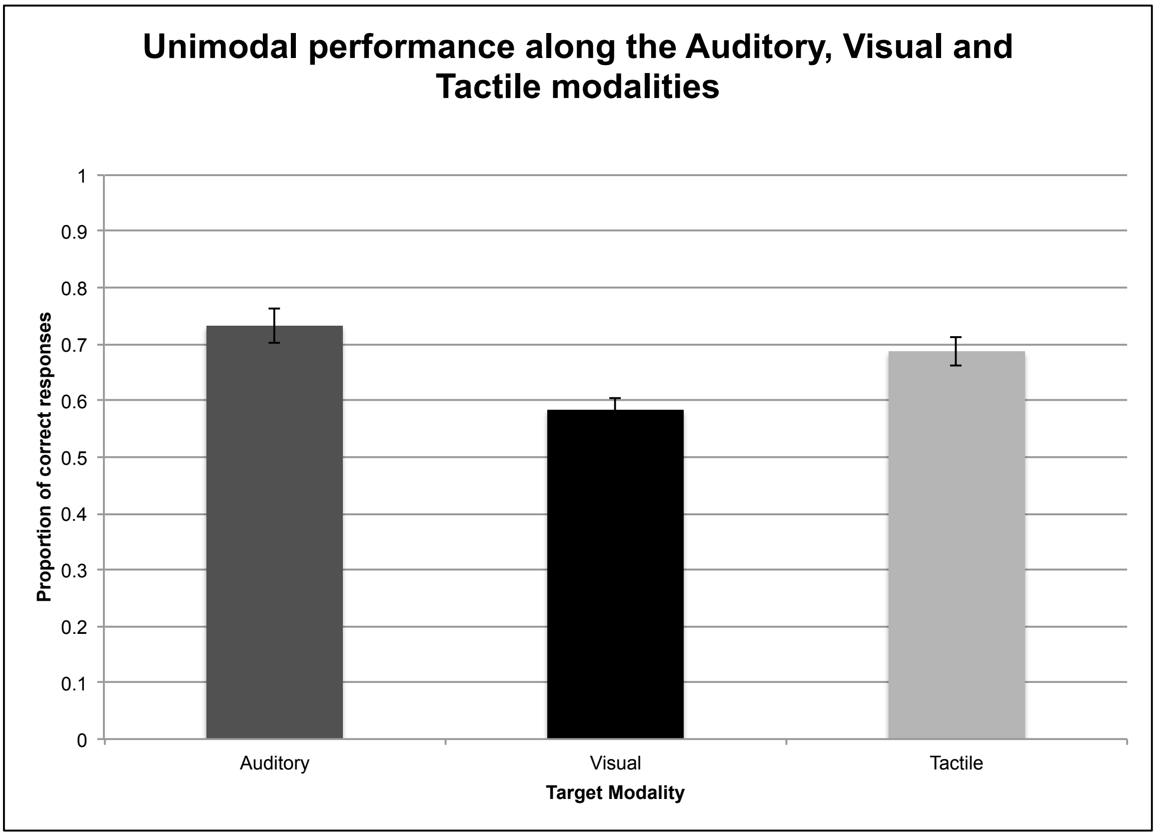


Figure 28. Unimodal performance along Auditory, Visual and Tactile modalities.

In sum, some of the most interesting findings from this experiment showed that in unimodal performance, audition and touch are significantly better than the visual modality, in this particular duration task. This finding is in line with previous claims that the auditory modality has a better temporal resolution than vision (Grondin, 1993), and that touch is more accurate than vision in duration discrimination (Tomassini et al., 2011). Also, the particular combination of audition and touch (audiotactile combinations) seems to have better temporal resolution than other modality pairings (Fujisaki & Nishida, 2009). This might be in relation to the fact that audition and touch are very much related, in several ways, for example, both being sensitive to the same type of stimulation (vibrations; Von Békésy, 1959; Soto-Faraco & Deco, 2009). There is however, a possibility that the estimation of the duration of the visual stimuli was very hard to accomplish (as reported by some of the participants); it was hard to differentiate between the durations of the stimuli coming from the LED lamp perhaps by the nature of the visual stimuli themselves, or the small differences in durations. It could be the case that vision does not have a very good resolution for duration perception.

Furthermore, when the auditory modality was the target, the tactile modality had a greater interference when it played as a distractor, than the visual distractor, and is in line with the idea that audition and touch are strongly linked, as mentioned above. In relation to the stimulus presentation condition, the Av pairing showed that congruent performance was not different to the incongruent performance, or to the unimodal performance. This finding is surprising, and very different from the findings reported in previous chapters of this work. It is true however, that the incongruent condition is different here, but still provided conflicting information in relation to the duration information from the target modality, and had indeed effects such as decrement in performance in other modality combinations. The question as to why there were no effects of the stimulus presentation factor is still hard to solve, and there is not a specific hypothesis as to the possible reasons involved in the difference of stimulus presentation effects.

In contrast with Av, the At combination showed that performance in both the congruent and unimodal conditions was better (and equally good) than the incongruent condition.

When the visual modality played as target, the auditory distractors helped vision in a greater fashion than touch helped vision in the congruent condition. This could be seen as an enhancement effect. One might wonder whether this finding could be explained by the possibility that audition shares more properties with vision (such as spatial relations) than touch with vision, and therefore, vision becomes more susceptible to auditory information. Or is it because audition is always better in temporal resolution, overall? This does not seem to be the case at least in this duration task, according to the unimodal analyses previously discussed.

Furthermore, the pattern of results from the main findings of the factor stimulus presentation revealed that for the Va combinations, the incongruent and unimodal conditions presented no differences in terms of performance. A suggestion for explaining this similarity could be that (as mentioned before) it very hard to discriminate between the visual unimodal durations, and this adversity somehow paralleled the difficulty set by the incongruent condition.

The tactile target seemed to be more affected by audition as distractor, than the visual modality. This is reflected in the fact that, in the bimodal incongruent condition (which represents the hardest condition of the stimulus presentation factor), vision provided better performance. It seems relevant to ask why was vision less disrupting than touch. Again, this might be related to the fact that audition and touch seem to be more intertwined (Occelli et al., 2011; Von Békésy, 1959), therefore the separation of duration information becomes more difficult.

Regarding the stimulus presentation performance, in the Ta combination the incongruent modality had the worst performance, and the congruent and unimodal appeared to be equally good. However, for the Tv pairing, all the conditions appeared to have the same performance, so there were no enhancement effects or interference. This finding is different from the previous chapters but similar to the findings for the Av pairing.

Further statistical analysis (e.g. an overall ANOVA for the target modalities) should be performed to allow comparisons between the target modalities and the stimulus presentation levels, as a way to compare these data with the findings from the previously reported experiments. This analysis would allow for example, comparisons of the nature: Is the At (auditory target, tactile distractor) incongruent condition better than Ta (tactile target, auditory distractor) incongruent? as a way to check for balanced interferences between these particular combinations or see whether one modality interfered more with the other, etc.

References

Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Burr, D., Della Rocca, E. & Morrone, M.C. (2013). Contextual effects in interval-duration judgments in vision, audition and touch. *Experimental Brain Research, 230*, 87-98.

Chen, K.M. & Yeh, S.L. (2009). Asymmetric cross-modal effects in time perception. *Acta Psychologica, 130*, 225–234.

Grondin, S. (2003). Sensory modalities and the temporal processing. In H.Helfrich (Ed.), *Time and Mind II*, pp.61–77, Hogrefe & Huber, Göttingen, Germany.

Grondin, S., Meilleur-Wells, G., Ouellette, C., & Macar, F. (1998). Sensory effects on judgments of short time-intervals. *Psychological Research, 61*, 261–268.

Fujisaki, W. & Nishida, S. (2009). Audio-tactile superiority over visuo-tactile and audio-visual combinations in the temporal resolution of synchrony perception. *Experimental Brain Research, 198*, 245-259.

Kleiner, M., Brainard, D., & Pelli, D.(2007). "What's new in Psychtoolbox-3?" *Perception, 36*, 1–16.

Klink, P.C., Montijn, J.S. & van Wezel, R.J.A. (2011). Crossmodal duration perception involves perceptual grouping, temporal ventriloquism, and variable internal clock rates. *Attention, Perception & Psychophysics, 73*, 219–236.

Occelli, V., Spence, C., & Zampini, M. (2011). Audiotactile interactions in temporal perception. *Psychonomic Bulletin & Review,* 18, 429–554.

Ortega, L., Guzman-Martinez, E, Grabowecky, M & Suzuki, S. (2014). Audition dominates vision in duration perception irrespective of salience attention and temporal discriminability. *Attention, Perception & Psychophysics, 76,* 1485–1502.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Rach, S. & Diederich, A. (2006). Visual-tactile integration: does stimulus duration influence the relative amount of response enhancement? *Experimental Brain Research, 173*, 514–520.

Soto-Faraco, S. & Deco, G. (2009). Multisensory contributions to the perception of vibrotactile events. *Behavioural Brain Research, 196*, 145–154.

Tomassini, A., Gori, M., Burr, D., Sandini, G. & Morrone, M.C. (2011).Perceived duration of visual and tactile stimuli depends on perceived speed. *Frontiers* *in Integrative Neuroscience, 5*, 1–8.

van Wassenhove, V., Buonomano, D.V., Shimojo, S. & Shams, L. (2008). Distortions of subjective time perception within and across senses. *PLoS ONE, 3,* e143.

von Békésy, J. (1959). Similarities between hearing and skin sensations. *Psychological Review, 66*, 1-22.

Walker, J.T. & Scott, K.J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones and gaps. *Journal of Experimental Psychology. Human Perception and Performance, 7*, 132.

Chapter 7. A study exploring elongation effects of the auditory modality in the perception of tactile durations

7.1 Introduction

The present study intends to further explore the extent and nature of the influence of the auditory modality on the perception of tactile durations. A mutual reciprocal interference from the auditory and tactile modalities has been consistently demonstrated in the previous studies described in the present thesis. However, the previously employed methodology has not allowed deepening into how these modalities impact one another. For example, it is still unclear whether the perceived durations in the target modalities are really the result of an 'expanded or compressed' duration caused by the distractor modality at play. In other words, it is still unknown whether the final percept of the target modalities is actually truly perceived as longer or shorter as the distractor modality dictates, or whether this is a result of participants finding it hard to ignore the distractor modalities in the incongruent condition. Hence, the distractor modality might have interfered with the processing of the target modality, but not necessarily elongate or shorten the perceived duration in the target modality.

The following experiment wished to overcome the methodological issue raised above, by means of an aftereffect paradigm. Therefore, the experiment included a training session that involved the exposure to auditory-always-longer durations, to check for possible elongation effects on a subsequent tactile discrimination task. Control conditions were also introduced, and will be explained in the Design and Procedure section.

7.1.1 Audition expanding the perceived duration of a visual stimulus

The work of Romei, De Haas, Mook and Driver (2011) attempts to explore the extent of influence of the auditory modality on the perceived durations of the visual modality. The effects of the auditory modality were tested in a 2AFC task. Here, participants were presented with pairs of concurrent stimuli occurring both in the visual and auditory modality (Experiment 1a). A congruent condition was included, where auditory and visual duration information was the same, and an incongruent condition, where the pairs of bimodal stimuli had opposing duration information. For example, in the first event, the visual modality would have a fixed duration and the auditory could either have a longer or shorter duration; in the second event, the visual modality would acquire the longer or shorter duration (i.e., the same duration of the auditory first event), and the auditory would have the fixed duration. Participants needed to determine which from the pair of visual stimuli was the longest, whilst ignoring the auditory information. A visual 'only' condition was introduced to serve as baseline, as well as auditory shorter or auditory longer durations (than the visual target), as distractors. Sensitivity measures (d') were computed, and indicated enhancement effects for the congruent condition, and impairment in performance or sensitivity, for the incongruent condition (Romei et al., 2011). These findings are interesting, however the authors jump to conclude that the visual events were perceived as longer when they were simultaneously presented with a longer auditory stimulus. That is, that the visual duration was expanded by the longer auditory durations, reflecting true perceptual changes (Romei et al., 2011). However, it is not known from their work, whether participants were attending to the auditory stimuli only (for example, due to a noticeable saliency of the auditory stimuli) and responded according to the auditory durations only, rather than attending/responding according to the visual durations. Therefore, the question of whether the visual durations were actually enlarged still seems to be unsolved. This perceptual riddle seems to be present along these kinds of studies, when it is unknown if they are decisional biases or whether they obey to real perceived changes (Soto-Faraco & Deco, 2009). In a further study though, De Haas, Cecere, Cullen, Driver and Romei (2013) made a different attempt to explore the stretching effects of duration in a task that did not involve the direct judgment of duration stimuli, rather, it demanded other non-temporal visual characteristics. The task had two conditions: visual and audiovisual. For the visual only condition, a visual target Gabor patch stimulus (which could have different variable durations) was presented embedded in either one of two additional visual stimuli that had a constant duration. These additional visual stimuli were dynamic white noise rectangles. For the audiovisual condition, the Gabor patch had a fixed duration and was embedded in one of the two additional visual stimuli (dynamic white noise rectangles). Additionally, in the audiovisual condition, a pure tone was embedded in both of the dynamic white noise rectangles. The tone had the same variable durations of the Gabor patch from the visual only condition. Participants had to respond as to whether the Gabor patch appeared in the first or second visual dynamic noise rectangle. The findings indicated that there was an increase of visual discrimination sensitivity (in terms of d’ of the signal detection theory) as a function of the visual variable durations for the ‘visual only’ condition. In other words, the variable visual durations enhanced visual sensitivity and this was reflected in a better performance, in contrast to the audiovisual condition, in which there was a decreased visual sensitivity, i.e. lower performance for half of the durations employed (de Haas et al., 2013). Therefore, it was believed that, for this task, the sound had an unfavourable effect for a given set of durations, and this is quite different from previous studies (Romei et al., 2011). Additionally, the paradigm design did not allow to show stretching effects.

A study by Heron et al. (2013) also documents an apparent ‘expansion’ and ‘compression’ of the perceived duration of visual stimuli, after pairing bimodal stimuli with auditory incongruent durations and visual durations. Specifically, in their first experiment (multisensory experiment), the authors presented participants with a two interval forced choice duration discrimination task. The task was to focus on a given target modality deciding which of a pair of stimuli was the longest, whilst ignoring the duration information on the distractor modality. Here however, in the pair of events within a trial, the distractor modality duration was simultaneously presented with only one of the target events of the pair (different from Romei’s et al., 2011 experiment, which paired the two events). The distractor duration could either be shorter or longer than the target standard stimulus (i.e., in the target modality). Both the auditory and visual modalities served as targets, and distractors, accordingly. The results showed significant differences in the point of subjective equality (PSE), only when the auditory stimuli served as shorter and longer distractors. In other words, the authors claimed that the visual referent was compressed when paired with the short auditory distractor (reflecting a lower PSE), and thus indicating that the duration was perceived as shorter, and enlarged (higher PSE) when paired with the long auditory distractor (Heron et al., 2013). However, as explained in the beginning of the chapter, the problem with this kind of studies is that one cannot rule out the possibility of having participants being distracted by the (in this case) auditory distractors, which should have been ignored. Therefore, attending to the auditory distractor could have caused to respond in accordance to its own duration, and not really generating an expanded or compressed perception of the visual stimuli.

The next section is relevant for this chapter because it addresses some techniques employed to explore the effects of bimodal training in subsequent perception of unimodal events. This is important for our study, since the aim was to know whether the auditory modality actually enlarges the perceived duration of subsequent tactile stimuli, and the selected methodology to test this question was through the use training and aftereffects testing.

7.1.2 Training and Aftereffects techniques to study the influences between modalities

Some studies have shown that after training effects with multisensory stimuli are reflected in later perceptual processing (Shams, Wozny, Kim & Seitz, 2011). For example, it has been shown that being exposed to bimodal situations (audiovisual) in a training stage, subsequently aids later unimodal recognition, even when the other modality is absent (Seitz, Kim & Shams, 2006; Kim, Seitz & Shams, 2008). Wozny and Shams (2011) show similar results, however they presented conflicting audiovisual information. For example, in training sessions they presented auditory and visual bimodal stimuli that were separated spatially. They also presented these stimuli unimodally. The task was to later identify the location of the auditory stimulus alone; when it was preceded by a visual different location, the subsequent perceived location of the auditory stimulus was in agreement with the previous visual location, rather than the actual position (Wozny & Shams, 2011).

Levitan, Ban, Stiles and Shimojo (2015) employed an aftereffects paradigm to test the overall idea of whether there is a common mechanism for rate perception shared between the senses, in this case, audition and vision. In order to explore this notion, they tested whether rate could be transferred across the modalities by means of adaptation stages. Specifically, their experimental design provided an adaptation stage and a main test task. The adaptation stage presented trains of stimuli with different frequency rates and these stimuli could either be visual or auditory. These trains of stimuli had gaps between them and participants had to state the number of perceived gaps. This allowed participants to become aware of the rhythm of the train of stimuli. Pre-adaptation and post-adaptation test tasks were introduced. These test tasks consisted on the presentation of stimuli in the opposite modality from the adaptation task (this happened only for the cross-modal condition, which is the relevant one for the studies reported in this introduction). In this test task, different pulse rates were presented (and were distinct from the rates used in the adaptation task). Here, participants focused on whether the presented stimuli were fast or slow, and had to be judged individually (and in relation to the mean speed rate of the previous stimuli. For this, they were given feedback only on the pre-adaptation test trials. The findings indicated that there was a transfer effect between the modalities, where the perceived rate shifted in relation to the adaptor rate in a negative aftereffect; for example, if there was an exposure to fast stimuli in the adaptation stage, the stimuli in the post test tended to perceived as slower when contrasted with the perceived rate of the pre adaptation test (Levitan et al., 2015). The authors suggested that their findings provide hints as to the possibility of a shared mechanism between the senses for time perception (Levitan et al., 2015).

These works suggest that multisensory training can be effective and still impact the later processing of events. Here, we were interested to test whether audition can truly change the perceived duration of subsequent tactile stimuli, after trainings of longer auditory-than-tactile durations. Our rationale was if there is a true elongation effect from the auditory modality, the subsequent tactile perceived duration would be shifted towards the auditory duration, that is, longer than its real (tactile) physical duration.

To the best of my knowledge there are no studies searching the extent of the influence of auditory durations on the perception of tactile durations.

In order to test these effects, an aftereffects paradigm was designed. This consisted on a training session that included pairings with longer auditory durations than the tactile ones (incongruent training). A control training condition was introduced with pairings of auditory and tactile stimuli that had the same durations (congruent training). A tactile duration discrimination task was introduced before the trainings (pre-test stage) and after the trainings (post-test stage), to check for possible tactile stretching effects.

7.2 Methods

7.2.1 Participants

A total of 25 participants from the University of Trento took part in this experiment (17 F, mean age 22.7 years old, age range from 18 - 41), in return for course credits or reimbursement. Due to a technical error from the experimenter, a participant was discarded from the analysis; therefore only 24 participants were employed for the analysis. They were not informed as to the purpose of the experiment, and gave prior written informed consent. The experiment was conducted in accordance to the principles from Declaration of Helsinki (1964) as well as the ethical guidelines from the University of Trento.

All participants reported normal hearing, seeing and tactile sensitivity.

7.2.2 Apparatus and stimuli

The equipment used to deliver the auditory, visual and tactile stimuli was identical to the one reported in Chapter 6 ‘Hierarchies of the senses experiment’.

Also, the auditory stimuli were delivered through headphones. White background noise (85.6 dB) was also presented to cover the noise made by the tactile stimulator and the loudspeaker that produced the light. The white noise was played using the same previously reported loudspeakers.

The experimental setup was also identical to that one used in the previously reported experiment and the same key response was employed. This time there were four possible response keys, which were used interchangeably in accordance to the task. For a clearer view of the setup, see Figure 29.

The experiment was delivered through the same PC laptop using PsychoPhysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) for MATLAB R2012b.

Figure 29. A view of the experimental setup

The experiment involved two different tasks, a visuo tactile duration discrimination task, and a training task. Details of these are further explained in the section ‘Design and Procedure’. However, the durations of the stimuli from the visuo tactile duration discrimination task are the following:

The visual stimulus represented the standard stimulus, and consisted of a red light with a constant duration of 125 ms across the entire task; it served as a reference duration to allow comparison with the tactile stimuli. The durations of the tactile stimuli consisted of 100, 110, 140 and 150 ms.

The durations for the stimuli in the training task consisted in the following:

a) For the Congruent training: auditory and tactile stimuli had the same durations, which were 100, 110, 140 and 150 ms.
b) For the Incongruent training: auditory stimuli always had a longer duration than the tactile (the latter had the same durations as in the

congruent training). Auditory durations were of 140, 150, 180 and 190 ms.

c) Deviant stimuli: these consisted on a simultaneous presentation of both a sound and a vibration. There were three kinds of deviant stimuli:
    1) when both sound and vibration had a brief interruption in the mid part of the stimuli
    2) only the sound presented an interruption in the middle
    3) only the vibration presented an interruption in the middle

All the deviant stimuli had a constant duration of 180 ms. and their interruption always started at 60 ms and lasted 20 ms. The deviant stimuli were always longer than any of the stimuli in the congruent condition (tactile and auditory) and tactile stimuli in the incongruent condition.

7.2.3 Design and procedure

As previously mentioned, the experiment consisted of two different tasks, delivered in separate blocks. At the beginning of the experiment, a visuotactile duration discrimination task was implemented (‘pre’ test stage), followed by a training task, and subsequently followed by the same visuotactile task (‘post’ test stage). The experiment structure was designed this way, in order to test the effects of the training stage on the post stages of the visuotactile discrimination task, when comparing the pre and post test stages. There were a total of 9 experimental blocks. For a better understanding of the block sequence across the whole experiment, see Figure 30.

This sequence would allow to check for the presence of elongation effects on the perceived duration of the tactile stimuli at the post stage, when the training stage (incongruent version) consisted of auditory longer stimuli. Details are given below as to the nature of the tasks.

7.2.3.1 Visuotactile duration discrimination task

The initial task consisted on a visuotactile duration discrimination task ('pre' stage). Here, participants had to decide which from a pair of visual and tactile stimuli was the longest by pressing the corresponding key ('L' for the visual stimuli (as for 'luce' in Italian), "V' for the tactile stimuli (as for 'vibrazione', in Italian). The stimuli consisted on a light (emitted by the LED lamp) and a vibration, presented sequentially (the order of presentation of the light and the vibration was counterbalanced across trials). After the training phase of the experiment, the same visuotactile task was implemented ('post' stage) as shown in figures a and b.

Each duration was 16 presented times, therefore a total of 64 trials were included for each stage (pre and post). The aim of this task was only to contrast the proportion of tactile longer responses before ('pre' stage) and after ('post' stage) the training phase. 8 practice trials were always introduced in the first 'pre' stage of the experiment.

7.2.3.2 Training task

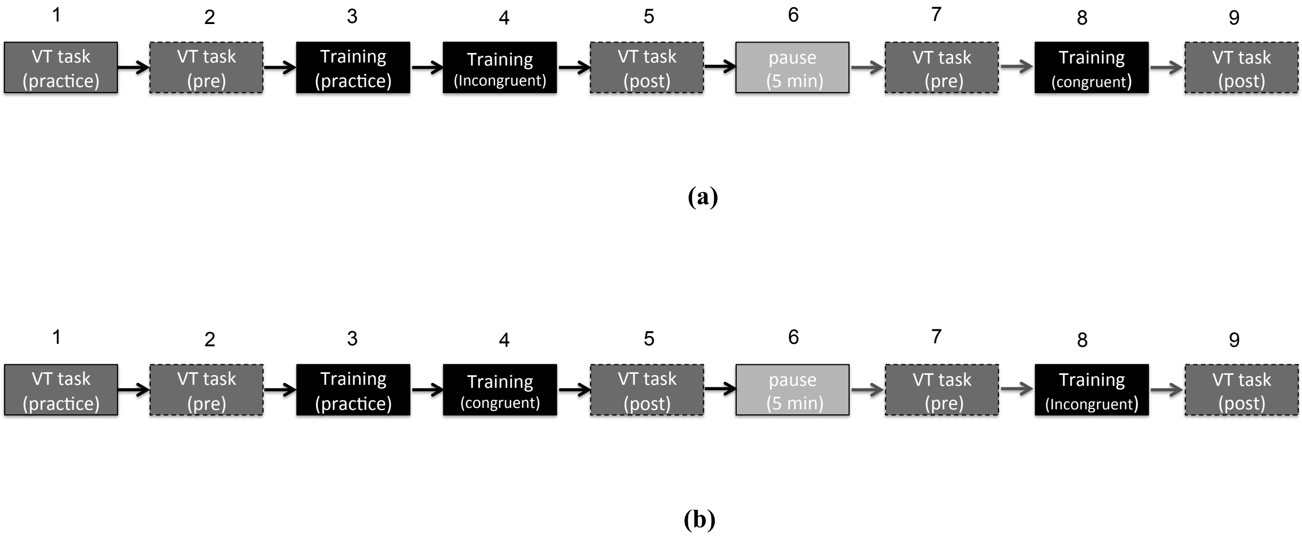
This task consisted on exposing the participants to audiotactile stimuli in two different conditions. In one condition, the auditory and tactile stimuli were presented simultaneously and had the same durations (congruent training). In the other condition, the auditory stimuli were always longer than the tactile ones by a constant 40 ms difference (incongruent training). Here, both auditory and tactile stimuli presented the same onset and clearly, different offsets.

Both conditions also presented deviant stimuli (which were detailed in the previous section), and the participant's task was to detect the presence of these deviant stimuli, i.e., detect an interruption in the stimuli. Participants pressed the 'Y' key if they detected an interruption, and the 'N' key if they did not.

This task allowed participants to be exposed to and actively attend every stimulus' duration.

Each training phase consisted on a total of 118 trials. 100 trials consisted on repetitions of each of the 4 different durations (each duration repeated 25 times). Each of the 3 possible deviant stimuli was repeated 6 times, with a total of 18 trials for the deviant stimuli. All the trials were randomly shuffled for each of the participants.

There were two training phases introduced in the experiment. Also, a total 7 practice trials were introduced at the beginning of the first training phase.



(a)

(b)

Figure 30. The figures indicate the sequence of blocks within the experiment. Dark gray squares correspond to the visuotactile duration discrimination task, which included practice, pre and post stages. Black rectangles indicate the training stages, which included practice, congruent and incongruent versions. Notice that two rounds of vt and training tasks were implemented. **a** and **b** symbol the counterbalance in the presentation of the congruent and incongruent training phases.

Participants sat in front of a screen and were asked to wear the headphones employed to deliver the auditory stimuli. The tactile vibrations were always delivered to the left hand index finger. Participants were asked to always fixate on the LED lamp located at the center of the screen. The beginning of each trial was marked by a change in the color of the screen (i.e. grey). In the rest of the trial, the screen was kept in a black color.

The instructions were showed in the screen and also repeated orally by the experimenter.

The experiment took place in a dark, sound attenuated booth, to allow participants to see the light emitted by the LED lamp.

A long pause of 5 minutes was introduced once the first round of ‘pre’, ‘training’ (either congruent or incongruent) and ‘post’ test stages were completed. Here, participants were asked to leave the sound proof booth to allow them to take a break from the task. After the pause, participants completed the second round of ‘pre’, ‘training’ (congruent or incongruent, accordingly) and ‘post’ blocks.

The experiment had a total duration of about 50 min and had a total of 507 trials.

Completion of a or b versions of the experiment were counterbalanced across participants.

7.3 Results and Discussion

7.3.1 Analysis 1. Contrasting tactile longer responses before and after training

An initial inspection of the data was made by the following analysis, where the aim was to compare whether the number of tactile longer responses was increased as a function of exposure to longer auditory durations (as was set in the incongruent training). For this purpose, the proportion of tactile longer responses was computed for each individual, and were employed to compute a within participants repeated measures ANOVA with the following factors: TRAINING STAGES (Congruent, Incongruent) x STAGES (pre training and post-training) x DURATION (100, 110, 140 and 150 ms). The results are summarized in Table 8.

Table 8. Summary of the data from the repeated measures ANOVA of Training type x Stages x Duration

Factor	gL	F	*p*	Notes
Training type	1, 23	.083	.776	NS
Stages	1, 23	.808	.378	NS
Duration	1.34, 30.90	100.57	.000	* GGC (X^2 = 49.65, *p* = .000)
Training type * Stages	1, 23	.516	.480	NS
Training type * Duration	3, 69	.920	.436	NS
Stages * Duration	3, 69	1.01	.392	NS
Training type * Stages * Duration	3, 69	.250	.861	NS

GGC indicates Greenhouse Geisser corrected values

As seen from the table, the only factor that showed significance was Duration, and it only reflected that the number of tactile longer responses increased for the longer durations, as expected. The mean number of tactile longer responses for each duration was the following: for 100 ms (M = .254, SE = .029, *p* =.000), 110 ms (M = .366, SE = .027, *p* =.000), 140 ms (M = .679, SE = .027, *p* =.000), and for 150 ms (M = .742, SE = .031, *p* =.000). These differences in duration can also reflect that participants were doing the task correctly.

Additionally in this analysis, it was expected to find an increased number of tactile responses in the post-training stage (M = .506, SE = .023) as compared to the pre-training stage of the incongruent training (M = .509, SE = .019, NS), however, as shown by the data, this was not significant. For a visual description of the data, see Figure 31 and Figure 32.

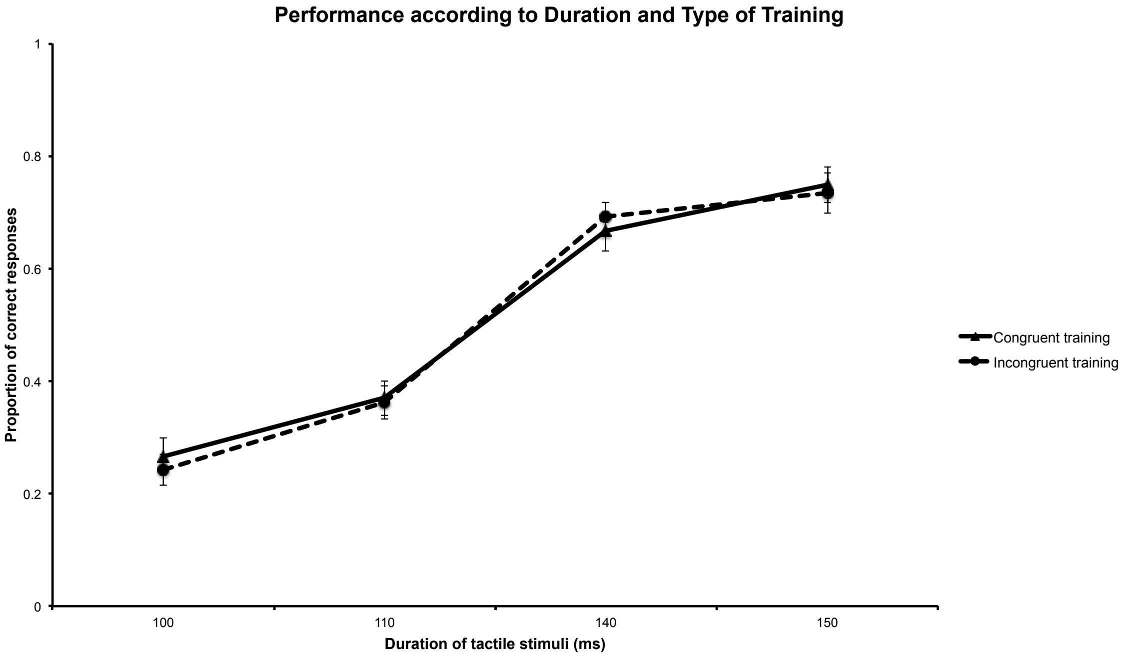


Figure 31. Performance according to Duration and Type of Training. First analysis

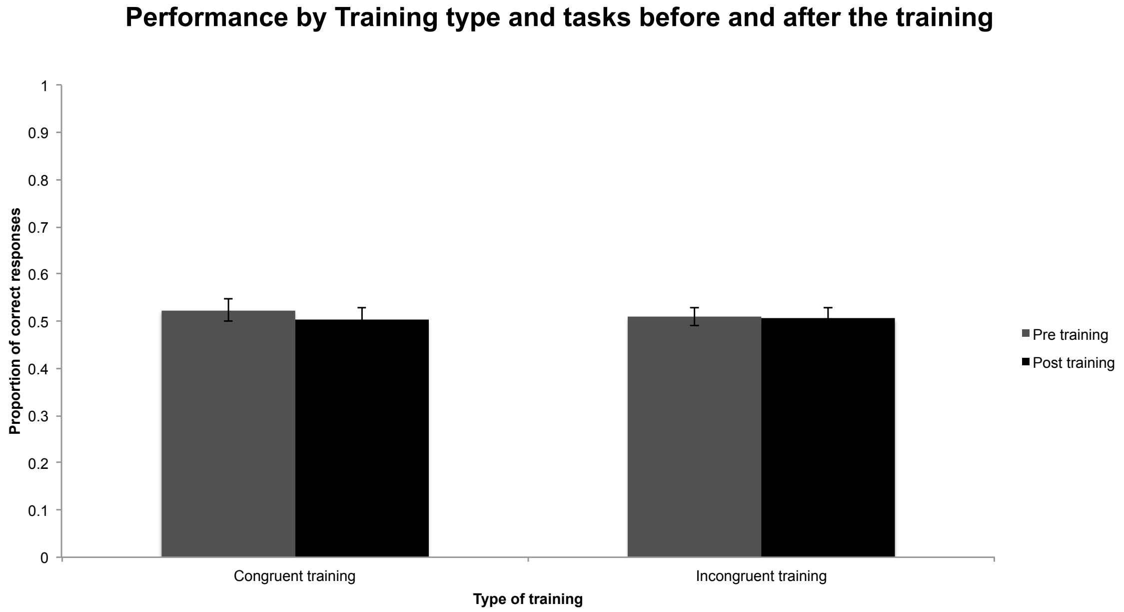


Figure 32. Performance by type of training before and after the training tasks. First analysis

7.3.2 Analysis 2. Contrasting the accuracy of visual longer and tactile longer responses

This analysis represented an alternative way of inspecting the data, in order to check for possible aftereffects. The rationale was that, if the exposure to longer auditory durations worked, it could be reflected on an increase in the accuracy of the tactile longer responses, as contrasted to the number of accurate visual longer responses. Thus, if the incongruent training was successful, performance should be better for the tactile longer duration stimuli (i.e. 140 and 150 ms), than for the visual longer duration stimuli (i.e. 100 and 110 ms), since the task involved participants responding as to which stimulus in a pair, either visual or tactile, was the longest).

In this analysis, the factor duration was compressed to allow a more precise view of the data, and to maximize the possibilities of observing an effect.

Therefore, the proportions of correct responses for the visual and tactile longer stimuli were first obtained to perform a within repeated measures ANOVA with the following factors: MODALITY (Visual/Tactile), TYPE OF TRAINING (Congruent, Incongruent) and STAGES (pre training and post-training). The results can be seen in Table 9.

Table 9. Summary of the data from the repeated measures ANOVA of Modality x Type of training x Stages

Factor	gL	F	*p*	Notes
Modality	1, 23	.196	.662	NS
Training type	1, 23	.739	.399	NS
Stages	1, 23	.436	.516	NS
Modality * Training type	1, 23	.077	.784	NS
Training type * Stages	1, 23	.495	.489	NS
Modality * Stages	1, 23	.830	.372	NS
Modality * Training type * Stages	1, 23	.500	.487	NS

As described in the table, neither the main effects nor the interactions between the factors were significant. The type of training did not exert an influence in performance according to the Congruent (M = .696, SE = .020) and Incongruent training (M = .707, SE = .019, NS).

Moreover, it was expected to find a better performance for the tactile modality, when the received training consisted of pairings with an auditory-always-longer stimulus. However, the proportion of correct responses from the tactile modality (M = .713, SE = .028) was not significantly different than performance in the visual modality (.701, SE = .027, NS.), in the incongruent training. More specifically, performance was not improved after the incongruent training in the post- training task, when comparing the visual (M = .702, SE = .035) and the tactile (M = .711, SE = .029, NS) performance.

For a visual inspection of the data on this analysis, see Figure 33 and Figure 34, which reveals no differences among the controlled variables in the experiment.

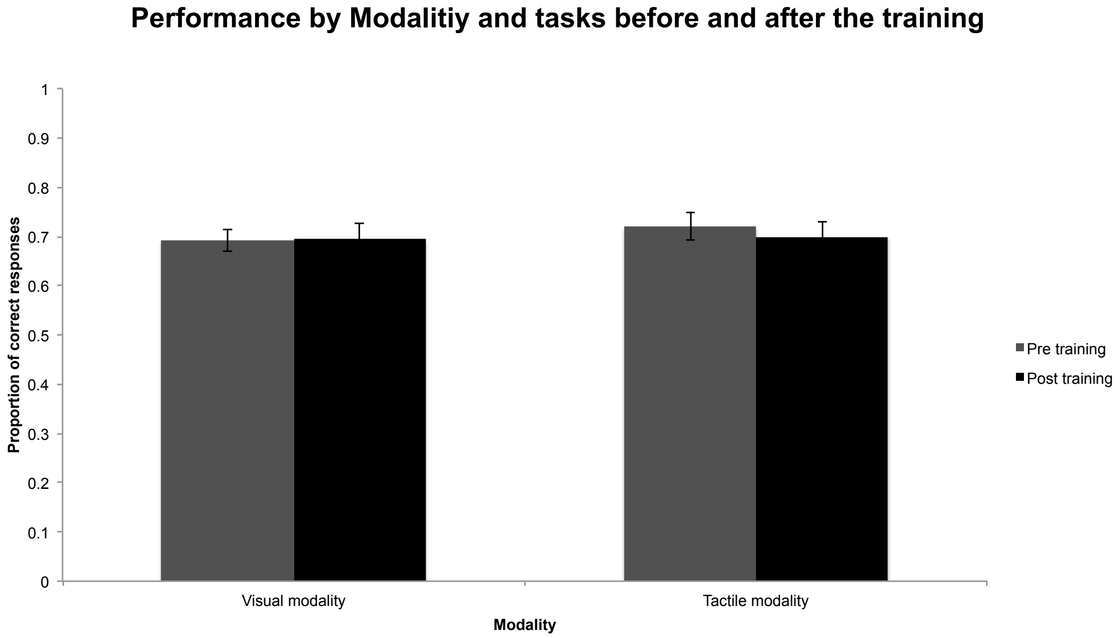


Figure 33. Performance of Visual and Tactile longer accuracy for the before and after training tasks.

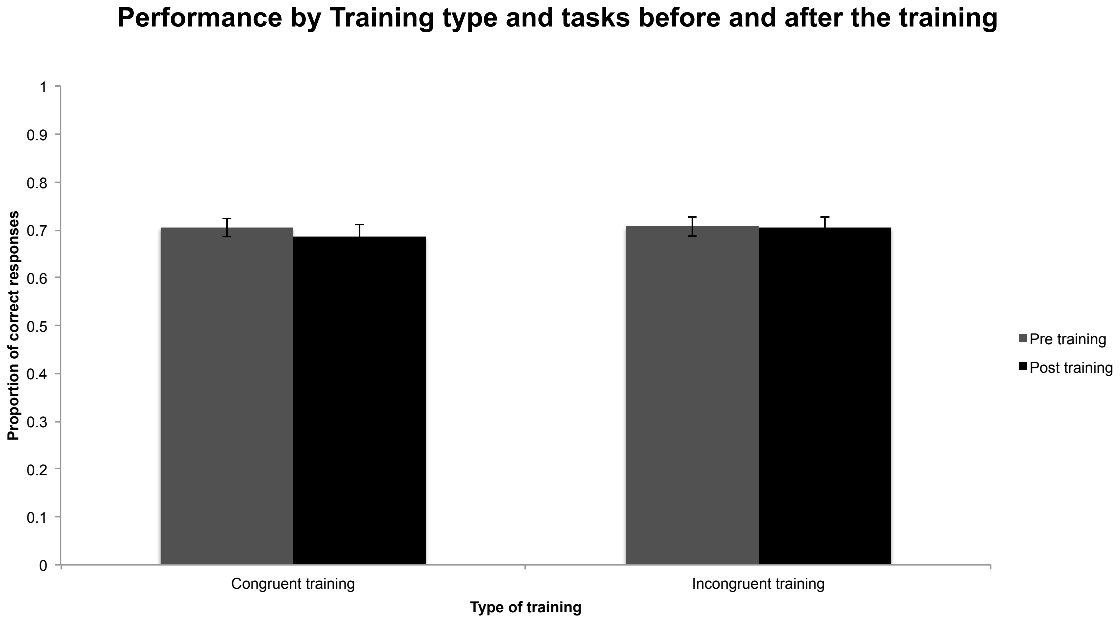


Figure 34. Performance by type of training before and after the training tasks. Second analysis

7.3.3 Analysis 3. Comparing JND and PSE values

The following analysis was made to allow the comparison of the point of subjective equality' (PSE) and the 'just noticeable difference' (JND) values, as a way to check for aftereffects of the auditory training. For example, a shift in the PSE would be expected if the incongruent training session worked, in direction of tactile longer perceptions. In other words, if the tactile stimuli were being perceived as longer, the PSE should be shorter on the post training stage, as compared to the pre-training stage.

For this purpose, the proportion of tactile longer responses was employed to calculate a psychometric fitting for each participant in the pre and post stages of the incongruent training. This was done in order to obtain the PSE and JND values.

A logistic function was employed to calculate the fitting, using the following formula:

$$F_L(x; \propto, \beta) = \frac{1}{1 + \exp\left(-\beta(x - \alpha)\right)}$$

Where x represents the tactile durations, β establishes the value of the slope and $\propto$ corresponds to the threshold: $F_L\ (x = \propto\ ; \propto, \beta)$. The data was computed through Palamedes Toolbox in MATLAB (Prins & Kingdom, 2009).

Once the JND and PSE values were obtained, paired t-tests were computed, in order to compare the pre-training and post-training performance. Since the fitting was not successful for all the participants, only the data of 14 participants was employed to calculate the t-tests.

The JND comparison revealed no significant differences, (t = .900, *p* = .384) in terms of the number of tactile longer responses on the ‘pre’ training condition (M = .030, SD = .054) and the ‘post’ training condition (M = .018, SE = .013), when the auditory information was longer than the tactile (incongruent training condition).

Regarding the PSE comparisons, the data showed no significant differences (t = -.872, *p* = .399) between the ‘pre’ (M = 112, SD = .021) and ‘post’ (M = 116, SE = .012) tactile longer responses. In this case, the mean values represent values in terms milliseconds.

If the incongruent training would have worked, then participants would probably perceive the tactile stimuli as longer after the incongruent training, however, this was not the case. Furthermore, since it was assumed that performance would improve with practice, the post- stage training should have rendered more tactile longer responses, at least after the incongruent training and this was not the case.

The results also show that the type of training did not have a different impact on the evaluation stages.

The analyses described above showed that the present experiment was not able to show any differences neither in the proportion of tactile longer responses nor on the proportion of correct tactile longer responses, or in the point of subjective equality from the pre-training and post-training test stages. This suggests that there were no effects of exposing tactile stimuli to auditory longer durations, and could perhaps imply that audition did not enlarge the perceived duration of tactile stimuli. The findings could also reflect that the impact of the auditory durations was not as relevant or influential to the subsequent tactile processing of the durations.

There could have also been methodological weaknesses in the design that could account for the lack of effects. Perhaps a larger number of expositions to the auditory-longer stimuli in the training stage are needed, since only 100 trials were employed in each training stage.

Another possibility could be that the training manipulation did not work because maybe the training task was not as engaging, so the participants might have not been aware of the auditory-always-longer duration stimuli, or might have not payed enough attention to them. A lot of variables could be influencing this lack of effect. For example, maybe participants forgot this auditory-longer relation because of the timing that passed between the training and the post-test stage, that is, a natural memory decay could have interfered. Probably, the overall paradigm was not suitable or sensitive to test elongation effects.

Also, the use of three modalities in the experimental design was perhaps not optimal, since there was too much different duration information from different sensory channels involved; therefore, a design with only two modalities would be purer to test the aftereffects, but could also bring however, methodological limitations.

Further work could be addressed to improve this paradigm and test the real nature of these multisensory interactions, as a way to learn whether they are real perceptual changes or only the effect of biases at responding, the so called ‘decisional biases’. As Soto-Faraco and Deco (2009) suggest, one way to solve this problem in these paradigms could be by using distractor stimuli set to an intensity so low, that is imperceptible to the subject. In this way, there would be no direct influence of the distractor at the decision or response stage, but however still be influential to perception (Soto-Faraco & Deco, 2009).

A different alternative could be in implementing a duration reproduction task, which could perhaps constitute a more direct assessment of an expansion effect in duration. In particular, it could allow seeing whether longer durations in touch are reproduced after a training period of longer durations in the auditory modality. Moreover, this could be done in a much simpler way, without combining modalities as in the study of Levitan et al. (2015), which managed to show a cross-modal transfer or aftereffect between audition and vision in rate perception, even when these modalities were not presented together, thus suggesting a common mechanism for encoding time (Levitan et al., 2015).

References

Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

De Haas, B., Cecere, R., Cullen, H., Driver, J. & Romei, V. (2013). The duration of a co-occurring sound modulates visual detection performance in humans. *PLoS ONE, 8*, e4789.

Heron, J. Hotchkiss, J., Aaen-Stockdale, C., Roach, N.W. & Whitaker, D. (2013). A neural hierarchy for illusions of time: Duration adaptation precedes multisensory integration. *Journal of Vision, 13*, 1-12.

Kim, R.S., Seitz, A.R. & Shams, L. (2008). Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS ONE*, 3, e1532.

Kleiner, M., Brainard, D., & Pelli, D.(2007). "What's new in Psychtoolbox-3?" *Perception, 36*, 1–16.

Levitan, C.A., Ban, Y.A., Stiles, N.R.B & Shimojo, S. (2015). Rate perception adapts across the senses: evidence for a unified timing mechanism. *Scientific Reports, 5*:8857.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Prins, N., & Kingdom, F.A.A. (2009). Palamedes: MATLAB routines for analyzing psychophysical data. http://www.palamedestoolbox.org .

Romei, V., De Haas, B., Mok, R.R. & Driver, J. (2011) Auditory stimulus timing influences perceived duration of co-occurring visual stimuli. *Frontiers in Psychology, 2,*1–8.

Seitz, A.R., Kim, R. & Shams, L. (2006). Sound facilitates visual learning. *Current Biology, 16*, 1422-1427.

Shams, L., Wozny, D.R., Kim, R. & Seitz, A. (2011). Influences of multisensory experience on subsequent unisensory processing. *Frontiers in Psychology, 2*, 1-9.

Soto-Faraco, S. & Deco, G. (2009). Multisensory contributions to the perception of vibrotactile events. *Behavioural Brain Research, 196*, 145–154.

Wozny, D.R. & Shams, L. (2011). Recalibration of auditory space following milliseconds of cross-modal.discrepancy. Journal of Neuroscience, 31, 4607-4612.

Chapter 8. General discussion

As has been previously exposed along this thesis work, the general purpose of this thesis work consisted on the exploration of audiotactile interactions in the perception of duration, and how different variables that have been previously found to create multisensory integration, work for this particular modality combination when perceiving temporal duration. These variables, as exposed in the following thesis, involved for example, the spatial proximity between stimuli from audition and touch, along with the intensities of the stimuli from audition and touch.

Furthermore, we were interested to know the hierarchies or patterns of dominance between the senses of audition, vision and touch in the perception of duration, in an interest to learn more about the interplay among these three modalities when perceiving this particular temporal feature. Moreover, we wanted to know which of these three modalities were more accurate for duration discrimination.

Additionally, the following work aimed to further understand the extent of the modulatory effects found between touch and audition, as a way to know whether they can truly generate perceptual changes on the perception of the other modality. Specifically, there was an interest to know whether the distractor modality (in this case only the auditory was tested) was able to expand or shorten the perceived duration of the target modality (in this case, the tactile). This meant an insight of the impact of the auditory modality that could be revealed in a final, modified perceived tactile duration.

It is worth mentioning that the temporal factor of duration served mainly as a tool to further explore the interactions between audition and touch along the entire experimental work reported in this thesis.

Prior to the discussion, a brief summary of the findings of the following thesis is presented next. The first experiments demonstrated that audition and touch interfere with one another when serving as distractor modalities, in a duration perception task. Specifically, they seemed to alter performance in a condition where the duration information between the two modalities was conflicting. The next experiments tested the impact of the intensity of the distractor stimuli in the same duration perception task and revealed different findings. One of the experiments testing intensity showed a balanced interaction between audition and touch, in terms of the same interference with conflicting duration information along with facilitation effects (in some cases), when the duration information was congruent. These findings happened despite changes in the intensities of the modalities. However, a further experiment revealed that audition altered touch in a greater manner when duration information was conflicting (regardless of the different intensity levels employed for audition), and that touch facilitated audition with congruent durations (only when the tactile intensity was set to a high level).

A further experiment explored whether the spatial location of target and distractor modalities would alter the interaction between audition and touch, in the same duration perception task. Specifically, to see whether performance would improve as a function of different or same spatial arrangements of the stimuli from the two modalities. The results showed that changes in the spatial location did not modify the interaction between the modalities, thus the same balanced interference effect was found as reported in the very first experiments.

An additional experiment included in this thesis addressed the role of adopting a crossed arms posture as a way to test the impact of conflicting frames of reference (anatomical and allocentric) on a tactile duration discrimination task. The results indicated that there was no effect of this posture as in contrast to non-crossed arms position on tactile duration discrimination.

One of the main findings revealed a balanced interference between the tactile and auditory modalities, mainly in a bimodal incongruent condition, where the duration information in a pair of events in the two modalities was the opposite. To the best of my knowledge, this is the first reported study addressing the interactions of audition and touch in the perception of duration.

This balanced interference proved to be a robust effect, since it was thoroughly replicated along this experimental work, despite changes in: 1) having different vs. same group of participants performing the task, 2) changes to the intensities or saliency of the distractor modalities (but however this only happened in one of the intensities' experiments), 3) changes to the spatial location of the stimuli from the distractor modalities (same or different location in relation to the target modality), and 4) in different ranges of durations (in ranges from 50 to 130 ms, and in ranges from 200- 280 ms).

It is not possible to draw a direct comparison of the latter findings with the experiment testing the hierarchies between audition, vision and touch in the perception of duration, due to methodological differences (for example, in the way of presenting the incongruent condition), as well as the inclusion of a new modality, (i.e. vision). Nevertheless, it is still possible to see a similar trend in the hierarchies experiment, in terms of the suggested strong links between audition and touch, and the difficulty of disentangling or separating the duration information of these modalities, despite the explicit request to ignore them in the task. For example, the findings showed that, for the incongruent condition, performance seemed to be more disrupted when the auditory played as a distractor in the tactile target (in contrast to the visual distractor), and when the tactile modality played as distractor for the auditory modality as target (again, in contrast with performance from the visual modality as distractor). This tactile effect however was marginally significant (*p* = .056) according to current conventions, but the effect appears strong enough to be discussed and relevant to the set of data reported in this work.

Additionally, when drawing direct comparisons between audition and touch in the 'hierarchies of the senses' experiment, audition seemed to provide better performance in the congruent condition than the tactile, when both acted as distractors for the visual modality as target. This difference in facilitation effects between the two modalities is interesting, because overall, according to the reported-here data, they have previously shown a balance in interference effects. One might wonder whether audition is still more accurate overall (than the other modalities) when perceiving duration, or that perhaps the audiovisual combination is more optimal than the visuotactile for improving duration perception. The first question though, is answered in the following finding from this experiment, which is that audition and touch showed to be equally accurate when discriminating between pairs of duration stimuli in unimodal contrasts, as compared to visual duration discrimination, which showed the worst performance in this overall comparison.

Another interesting remark of the experiment testing the hierarchies of the senses, is the fact that it was not able to replicate the same incongruency effects shown in the other experiments, where the incongruent condition provided the worst performance. This difference might stem from the differences in the incongruent condition. In this particular case, the durations of the distractor modality were the same, i.e., they adopted the standard duration for both events of a pair (but however could be shorter or longer than the probe durations). In most cases the performance was equal among the congruent, incongruent and unimodal conditions. Perhaps here the incongruent condition was less confusing, and more clear thus providing overall better performance.

This finding slightly resembles the findings from Romei et al (2011), in audiovisual conditions, where longer or shorter distractor durations in the incongruent condition did not affect performance and sensitivity measures.

Most of the above mentioned findings provide additional evidence that could suggest that audition and touch share close links. In this work, they proved to be equally accurate to discriminate duration and they affect one another in a balanced manner, as opposed to other sensory interactions, in which other directions in modulation have been seen; for example in audiovisual interactions, usually one dominates the other one depending on the examined variable; if it is spatial, usually vision dominates (Bertelson & Radeau, 1981; Howard & Templeton, 1966;), if it is temporal, usually audition dominates (Morein-Zamir et al., 2003; Repp & Penel, 2002).

In relation to time perception theories, there have been different proposed models that wish to explain the timing mechanisms in humans. Some of them, as explained in chapter one, propose the existence of a single or multiple clock mechanisms. The notion of a single vs. multiple timing clocks derives partly from findings regarding duration perception. For example, these findings indicate differences in the perception of the duration of a same event by different modalities: an auditory stimulus is usually perceived as longer (despite having the same physical duration) from a visual one (Walker & Scott, 1981). If there was a unique clock for all the modalities, then the phenomenon of *auditory-always-longer* stimuli could not be explained. Rather, multiple clocks theories would argue that there are different clocks for each modality, with an individual pacemaker that operates at different emission rates depending on the modality at play (Wearden, Edwards, Fakhiri & Percival, 1998). The existence of single or multiple modal timing mechanisms is still an open debate.

The findings from this thesis are difficult to be directly compared with the above-mentioned theories, because firstly, the overall paradigm employed here does not directly measure whether the auditory, tactile or visual stimuli employed in the thesis were perceived as equal, longer, etc. in duration terms. That is, there is no information here that compares whether an auditory stimulus is perceived as longer than a tactile one, or whether a visual stimulus is perceived as longer than a tactile stimulus, etc. Rather, the paradigm informs us about the degree of interference or facilitation between the modalities, when discriminating the duration of pairs of stimuli. However, the findings from this thesis could suggest that audition and touch could share a common mechanism for duration perception, or at least a component from the clock models due to the difficulty for separating these modalities when discriminating between durations. This sharing could allow a flow of information and constant cross talk from one modality to the other, that has been reflected as a disruption of performance when they are presented simultaneously, as seen from the findings. However, one should consider that this explanation is not completely consistent with the findings observed for vision. Therefore, it may also be the case that audition and touch, but not vision, share a timing mechanism. The existence of a shared clock (or component of the timing model) between audition and touch, as opposed to vision could lie perhaps in the nature of the stimuli that is perceived by these modalities: since both tactile vibrations and pure tones operate in forms of vibratory cues, audition and touch could be sharing the same pacemaker or internal timing component that could be sensitive to the same type of stimulation.

The main finding of a balanced interference between audition and touch has been highlighted along this discussion. However it could be claimed that this modulation is only partial, since there has not been a constant pattern of facilitation, as demonstrated by the results reported here. Some possible reasons could be that either our paradigms have failed to show this facilitation, or the audiotactile interactions when perceiving duration is only attained by interferences with one another, and not necessarily *helping* one another. As has been documented in the multisensory literature, one of the rules for attaining multisensory integration between the modalities establishes that the intensities of the stimuli are critical for the integration to occur, integration that at times can derive in enhancement. Perhaps the intensities involved along the series of experiments were not necessarily low enough to promote this integration (despite the explicit manipulation and efforts in some experiments to set them low).

A further exploration as to the nature of the modulation between audition and touch was made in this work, in an attempt to learn whether there is a real expansion of the perceived duration according to the duration distractor modality. The paradigm employed in this work provided no significant results that could hint for a real expansion of the tactile durations as a result of exposure to auditory longer durations. Previous evidence has found the existence of cross-modal transfer of information between the modalities (for example, in rate perception Levitan et al., 2015) with aftereffects paradigms. One could speculate that there might also be a strong cross-modal transfer of duration information between audition and touch, in particular for these two modalities that happen to show such a tight bond. Therefore, further attempts with new methodologies could be done to explore expansion effects of a distractor modality on the perceived duration in the other modality.

In sum, the present work with the above-mentioned findings suggests that touch and audition are related, extending previous evidence indicating that these modalities share close links. Furthermore, they proved to be equally accurate for the discrimination of duration. To the best of my knowledge, this is the first empirical exploration of this sensory interaction in the perception of temporal duration.

8.1 Further implications

The reported here findings help broaden the current account of audiotactile interactions, informing us about the perception of duration. In particular, they provide a clearer picture regarding the relation of audition and touch when perceiving duration, proving to be equally accurate for this feature, in contrast to vision. This can add to the understanding of the hierarchy of the senses regarding the perception of time, as well as documenting evidence for multisensory integration principles.

These findings can also give hints for current time perception theories and models, perhaps leaning towards a common timing mechanism that could allow a cross talk between these particular modalities.

Multisensory situations usually derive in human performance improvement such as faster responses (Ho, Reed & Spence, 2007) and efficient learning (Shams & Seitz, 2008), to name a few. The facilitatory effects of multisensory situations can have several applications, where the use of rapid responses is needed. However, in this thesis work, the facilitatory effects of audiotactile pairings in the perception of duration did not find a clear systematic pattern or repetition. As has been discussed previously, this lack of pattern could be due to methodological issues or that our paradigm was not sensible for demonstrating facilitation. Nevertheless, other studies have shown facilitatory effects in audiotactile interactions (e.g. Gillmeister & Eimer, 2007; Zampini, Torresan, Spence & Murray, 2007). If we take this into consideration, there might be some useful applications of this particular modality combination.

Previous evidence by Ho et al. (2007) has shown that audiotactile cues (in form of horn sounds and vibratory belts) facilitate the driver in avoiding a collision with other vehicles in simulated environments. These bimodal cues improved performance in contrast to unimodal cues, by providing effective warning alarms that were activated whenever a lead car decremented its speed. Similarly, findings reported in this thesis could be employed for generating informative alarms as to the closeness of another car or obstacle when parking, for example. In this case, congruent (in terms of duration) audiotactile alarms could be an indicator as to how far or close our car is to an obstacle/another car when parking. For example, a short duration of the audiotactile signals would indicate that the obstacle is very proximate, whilst longer durations could imply that the obstacle is further away in distance. This would mean making an analogy of duration to the distance between objects.

This type of audiotactile warning principle could also be used in medical surgery equipment or in a platform for ‘surgery learning’. For example, audiotactile-congruent-duration stimuli could provide information as to the location in space of an area to reach in surgery, etc. A short duration would mean that the area is proximate, a long duration that the location is far away, etc. This would be very helpful when there are no visual cues available.

Another application of audiotactile signals could be found in the field of music. For example, audiotactile cues could be useful for learning musicians who need feedback about the duration of a certain musical note or rhythm. Specifically, audiotactile pairings with congruent duration could provide information about the duration of a tone, indicating when they need to finish or change a note, etc. Overall, these cues could prove efficient for the learning musician.

References

Gillmeister, H. & Eimer, M. (2007). Tactile enhancement of auditory detection and perceived loudness. *Brain Research, 1160*, 58-68.

Ho, C., Reed, N. & Spence, C. (2007). Multisensory in-car warning signals for collision avoidance. *Human Factors, 49*, 1107-1114.

Levitan, C.A., Ban, Y.H.A., Stiles, N.R.B. & Shimojo, S. (2015). Rate perception adapts across the senses: evidence for a unified timing mechanism. *Scientific Reports, 5*, 1–5.

Shams, L. & Seitz, A.R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences, 12*, 411-417.

Walker, J.T & Scott. K.J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones and gaps. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 1327-1339.

Wearden, J.H., Edwards,H., Fakhiri, M. & Percival, A. (1998). Why "sounds are judged longer than lights': Application of a model of the internal clock in humans. *Quarterly Journal of Experimental Psychology, 51B*, 97-120.

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