# Using multimodal attention to design sensory substitution devices:
# Basic research and application

**Ivan Makarov**

Thesis for the degree of Philosophiae Doctor

November 2024

**School of Health Sciences**

**FACULTY OF PSYCHOLOGY**

**UNIVERSITY OF ICELAND**

# Using multimodal attention to design sensory substitution devices: Basic research and application

**Ivan Makarov**

Thesis for the degree of Philosophiae Doctor

**Supervisor(s)**

Rúnar Unnþórsson

Árni Kristjánsson

**Doctoral committee**

Rúnar Unnþórsson

Árni Kristjánsson

Ian M. Thornton

November 2024

# Skynskipti með fjölþættri athygli - grunnrannsóknir og hagnýting

**Ivan Makarov**

Ritgerð til doktorsgráðu

**Leiðbeinandi/leiðbeinendur**

Rúnar Unnþórsson

Árni Kristjánsson

**Doktorsnefnd**

Rúnar Unnþórsson

Árni Kristjánsson

Ian M. Thornton

Nóvember 2024

**Heilbrigðisvísindasvið**

**SÁLFRÆÐIDEILD**

**HÁSKÓLI ÍSLANDS**

ISBN 978-9935-9764-3-7

ORCID: 0000-0002-1942-1759

Reykjavik, Iceland 2024

# Ágrip

Skynjun á heiminum í kringum okkur takmarkast sjaldnast við aðeins eitt skynfæri. Skynjun heimsins er því fjölskynjun þar sem öll skynfærin leika hlutverk. Skilningur á því hvernig menn skynja og samþætta upplýsingar frá mismunandi skynfærum er nauðsynlegur til að þróa skilvirkan skynskiptibúnað. Í þessari ritgerð er flókið samspil mismunandi skynfæra kannað ásamt hlutverki athygli í skynjun og hvaða áhrif það hefur á hönnun skynskiptibúnaðar. Einkum er athyglinni beinta að snertiskyni og lögð áhersla á mikilvægi þess að skilja skynvillur og túlkun. Í grein I var ný skynvilla rannsökuð. Skynvillan felst í að sterkari snertiáreiti eru skynjuð ofar en þau sem eru af lægri styrk (e. Intensity order illusion, IOI). Ef seinna titringsáreiti á sama stað á mjóbakinu hefur hærri styrk en fyrra áreitið hafa þátttakendur tilhneigingu til að skynja það sem hreyfingu upp á við (og öfugt). Tvær tilraunir voru gerðar til að skilja hvað veldur IOI og til að komast að því hvort það virkar í láréttri vídd eða aðeins í lóðréttri vídd. Niðurstöðurnar úr grein I veita innsýn í IOI og hlutverk tíðni og styrkleika í skynvillunni. Auk þess staðfestir rannsóknin virkni Lofelt 5 titrara fyrir snertiörvun, sem skiptir sköpum fyrir þróun skynskiptibúnaðar okkar. Í greinum II og III er fjallað um fjölþætta athygli og áhrif samstillingar áreita fyrir mismunandi skynfæri fyrir frammistöðu í söfnunarverkefnum (e. foraging). Í grein II voru áhrif sjónar og heyrnar á söfnunarverkefni rannsökuð. Niðurstöður úr þremur tilraunum sýna að þegar markáreiti breytast í samstillingu sjónar og heyrnar, leiðir þetta athyglina að mörgum markáreitum. Engar vísbendingar voru um að leitin gengi betur þegar hljóð sem ekki er staðbundið fylgdi með, jafnvel þegar söfnunarverkefnið var flóknara (tilraun 3). Rannsóknirnar í grein III voru gerðar til að skilja hvort vísbendingar frá mismunandi skynfærum getu bætt frammistöðu í sjónrænu söfnunarverkefni. Vísbendi frá þremur mismunandi skynfærum (sjónræn, haptic og hljóð) voru borin saman við það þegar engar vísbendingar komu fram. Niðurstöðurnar sýndu greinilega að vísbendi frá öðrum skynfærum (heyrn og snerting) bættu frammistöðuna meira heldur en vísbendi innan sömu skynjunarvíddar. Vonast er til þess að niðurstöðurnar geti gagnast við að þróa skynskiptibunað og að þær veiti innsýn í hvernig upplýsingum sé best að miðla, og með hvað hætti, þegar þróa á skynskiptibúnað. Á heildina litið undirstrikar rannsóknirnar mikilvægi þess að huga að samþættingu mismunandi skynfæra og virkni athyglinnar við þróun skynskiptibúnaðar, en einnig almennt, til að skilja skynjun. Með því að nýta innsýn frá snertiskynjun og fjölþættri athygli, getum við aukið notagildi og skilvirkni skynskiptibúnaðar fyrir notendur sem glíma við truflanir á skynjun.

**Lykilorð:**

Skynskiptibúnaður; Þver-skilningarvit; Snertiathygli; Sjónræn Athygli; Fjölþætt athygl.

# Abstract

Perception of the world around us can never be limited to only one stream. The perception of the world is multimodal. Understanding how humans perceive and integrate information from multiple sensory modalities is essential for developing effective sensory substitution devices (SSDs). This thesis explores the complex interplay of sensory modalities and attention in perception and how it relates to the design of SSDs. The research investigates tactile perception and its implications for SSDs, highlighting the importance of understanding tactile illusions and anisotropies. In paper I the intensity order illusion (IOI) was studied. This illusion can be described as the mislocalization of the second stimuli following the first one. If the second stimulus on the lower back has higher intensity than the first stimulus then participants tend to perceive it as an upward movement. This also works vice versa. Two experiments were conducted to understand what causes the IOI and to find out whether it works in horizontal deimension or only in vertical dimension. The findings from Paper I reveal insights into the IOI and the role of amplitude for the IOI. Additionally, the study validates the effectiveness of Lofelt 5 actuators for tactile stimulation, crucial for SSD development involving tactile stimulation.

Examining multimodal attention, Papers II and III delve into the effects of synchrony and cross-modal cueing on foraging performance. In Paper II the focus was on studying the influence of visual and auditory synchrony on foraging. Three experiments were run to study the the influence of synchrony. The results show that when targets move in visual synchrony, this guides attention during multiple target visual search. These results can possibly be explained by the grouping of the targets. There was no evidence that participans have any search benefit from non-spatial sound even when the task was more complicated (Experiment 3). Paper III was made as an extension of the second paper. The main goal was to understand whether cues from different modalities can improve visual search. Three cues from different modalities (visual, haptic and sound) were compared to the condition where no cues were presented. The results clearly showed that cross-modal cueing (haptic cue and sound cue) is more effective compared to uni-modal cueing (visual cue). These findings can benefit in developing SSD and provide some insights about how information should be conveyed using SSD's.

Overall, the research underscores the importance of considering multimodal integration and attentional mechanisms in the development of SSDs, but also more generlly, for understanding perception. By leveraging insights from tactile perception and multimodal attention, future SSD designs can optimize sensory feedback delivery, enhancing usability and effectiveness for users with sensory impairments.

**Keywords:**

Sensory Substitution Devices; Cross-Modal Perception; Tactile Attention; Visual Attention; Multimodal Attention.

# Acknowledgements

I would like to thank, first and foremost, Prof. Árni Kristjánsson and Prof. Rúnar Unnþórsson as my primary supervisors. Thanks to the University of Iceland for hosting. Finally, thank The Icelandic Technical Research Fund and The Icelandic Centre of Research (Rannís) for funding this work. I am deeply grateful to Árni and Rúnar for the life-changing opportunity they gave me three and a half years ago to pursue my PhD and to experience the beauty of Iceland, which has now become my second home.

I would like to thank Prof. Ian M. Thornton t the University of Malta, who is on my doctoral committee, for developing the tasks used in Papers II & III, for spending much of his time on our papers, for his patience and understanding, for his support, and for his strictness when needed.

I thank Dr. Sabrina Hansmann-Roth, Dr. Inga-Maria Ólafsdóttir, Prof. Heida Maria Sigurdardottir, and Prof. Árni Gunnar Ásgeirsson for productive discussions, interesting thoughts and questions, and wild parties in Iceland and outside it. I would like to thank all current and past Icelandic Vision Lab members for our productive working environment, interesting discussions, and tasty parties.

I would like to thank Hafliði Ásgeirsson for our friendship, productive work, and great fun together. I also want to address my Russian-speaking friends here in Iceland, Vlad and Anton, for the great time we spent together.

I would like to thank Elena Sergeevna Gorbunova for showing me the magic world of cognitive psychology and Stephan van der Stigchel for bringing me back to science.

I am incredibly blessed to have such a supportive family. My father inspired me to choose cognitive psychology as a career. My mother's unwavering support has been my strength during emigration and studying. My sister, who is launching her own career journey. I wish you all success and know that I am always here for you. Thank you for the short trips, long calls, and constant support. I love and miss you all!

I thank all my friends, Danya, Egor, and Ksyusha. We are spread throughout Europe, but our meetings are fun, eventful, and unforgettable! Dima, you are my besty, thank you for everything! I would like to thank my friends, who are far away in Russia and have supported me for decades – Kostya and Masha.

Last, I would like to give my thanks to Marta Þormóðsdottir Grönvold for everything that was and will be between us from the meeting in café Rosenberg to today and hopefully a bright future. Thank you for supporting me during these times, whether we were together or long-distance.

# Contents

# List of Abbreviations

3D – Three Dimensional

CI – Cochlear Implants

IOI – Intensity Order Illusion

ITT – Inter-target interval

L5 – Lofelt L5 actuators

SOV – Sound of Vision

SSD – Sensory Substitutional Devices

Px – Pixels

RT – Response time

TCT – Total Completion time

# List of Figures

# List of Original Papers

This thesis is based on the following original publications, which are referred to in the text by their Roman numerals (I, II, III, ... [as needed]):

I. Makarov, I., Stefánsson Thors, S. S., Ævarsson, E. A., Jörgensson, F. K. P., Yeganeh, N., Kristjánsson, Á., & Unnthorsson, R. (2024). The Haptic Intensity Order Illusion Is Caused by Amplitude Changes. *ACM Transactions on Applied Perception, 21*(1), 1–18. https://doi.org/10.1145/3626237

II. Makarov, I., Unnthorsson, R., Kristjánsson, Á., & Thornton, I. M. (2024). The effects of visual and auditory synchrony on human foraging. *Attention, Perception, & Psychophysics, 86*, 909–930. https://doi.org/10.3758/s13414-023-02840-z

III. Makarov, I., Unnthorsson, R., Kristjansson, A., & Thornton, I. M. (2024, January 11). Cross-modal cues improve the detection of synchronized targets during human foraging. https://doi.org/10.31234/osf.io/xrty3. Accepted in Multisensory Research.

All papers are reprinted by kind permission of the publishers.

# Declaration of Contribution

Ivan Makarov made contributions for designing and programming experiments, developing equipment, planning and performing data collection, data analysis, and writing manuscripts. The thesis was written by Ivan Makarov.

Árni Kristjánsson, Runar Unnthorsson provided supervision and contributed to experiment design, data analysis, writing paper manuscripts and thesis.

Ian Thornton provided supervision and contributed to experiment design (developing the tasks in papers II & III), data analysis, programming, writing paper manuscripts (papers II and III) and thesis.

Snorri Steinn Stefánsson Thors made contributions to programming experiment, performing data collection and data analysis (paper I).

Elvar Atli Aevarsson made contributions for designing experiments, developing experimental equipment, and performing data collection (paper I).

Finnur Kári Pind Jörgensson made contributions for designing experiments, developing experimental equipment (paper I).

Nashmin Yeganeh made contributions for writing manuscript (paper I).

# 1 Introduction

How do we perceive the world around us and ourselves in this world? Many people would say that most of the information about the world around us is perceived by our eyes. While this statement is correct, this sentence says nothing about other ways of perceiving or gathering information. Humans have seven sensory modalities: visual (sight), auditory (hearing), tactile (touch), olfactory (smell), gustatory (taste), proprioceptive (sense of body position and movement), and vestibular (sense of balance and spatial orientation). It is difficult to isolate modalities and say that we rely on only one modality at some particular point. Even if a person focuses their attention on a specific modality, information from other modalities continues to enter our brain unconsciously (Sinnett et al., 2006; Treisman, 1960). If we did not receive and constantly analyse information from our vestibular system, we would end up lying on the ground because of the force of gravity. Human beings constantly and simultaneously perceive a large amount of information from different modalities. Our representation of the perceptual world in our brain is, in other words, multimodal.

Analysing information from the different senses allows humans to reach conclusions about objects and describe and categorize them. There is no need for grown-up humans to start touching every object that they see in front of them. They usually have enough information to categorize and label stimuli without touching or smelling them. Nevertheless, it is undoubted that our brain perceives objects or events in real life through different modalities. Such integration of different modalities occurs, for example, when we look outside the window at the road. The cars will make a noise with their tires and engines, and we can understand whether the car is approaching us or driving away. Usually, we would also see cars moving towards or away from us. Combining auditory and visual information could save our lives if we cross the road and our visual attention is distracted away by something. This example shows why people should remove their headphones before crossing the road!

Synaesthesia, an extreme and rare condition, is a fascinating phenomenon where information from different modalities intertwines. This condition triggers a unique experience unrelated to the initial stimulus (Banissy et al., 2014). For instance, some individuals associate musical notes with specific colours or shapes, while others perceive a connection between letters and sounds. But a more common case is when multi-modal perception leads to non-synesthetic combinations such as high pitch tone and high position of an object or low pitch tone and large-sized objects (e.g., Evans & Treisman, 2011). Another example of integrating different sensory modalities is combining the information from the movement of the lips (visual modality) and voice (auditory modality). Throughout their life, most people unconsciously learn the position of the human's lips when pronouncing different words and sounds. This knowledge helps them to understand what their friend is saying within loud spaces.  However, this case also highlights a common multisensory illusion of our brain, the so-called ventriloquist effect (Bruns, 2019). This effect can be observed when people tend to mislocalize the source of the audio signal (the moving lips of the puppet). The ventriloquist illusion demonstrates the brain's remarkable ability to integrate sensory information from multiple modalities to construct a coherent perceptual experience.

### *The effects of sensory impairments*

How do humans perceive information from the world around them if one modality is missing? For example, suppose people are blind from birth or early in life. In that case, there is evidence that they can learn to substitute their visual modality with information from other modalities. For example, a blind person can have more developed tactile and audio modalities. Studies show that the later a person loses the ability to see, the worse their brain exchanges the visual modality with other modalitiess. This lack in adaptivity is caused by the reduced plasticity of the brain with age (Merabet & Pascual-Leone, 2010).

There are cases when people with normally functioning modalities cannot rely on them. For example, firefighters frequently face situations when the smoke blocks their visual modality. In such conditions of low visibility, instruments could be developed to help them perceive the world around them through another modality. Unfortunately, the development of these instruments is closer to the early beginning of such design than having arrived at the well-developed and "comfortable to use" stage. Several questions need to be studied to increase the quality of these instruments and make them more effective and easier to use. The results described in this work address some of the issues connected to the usage of these devices.

### *The outline of the thesis*

This thesis consists of six main chapters: Introduction, Aims, Materials and Methods, Results, Discussion and Conclusion. Additionally, a reference list and a list of original papers are provided in the end.

Chapter 1 is the "Introduction" and includes literature review about modalities addressed in this study: visual, tactile, and audio. Additionally, it provides an overview of multimodal integration mechanisms and more about the joint work of the respective modalities. Then it covers the sensory substitution mechanisms that naturally appear in our body and sensory substitution devices. Finally, it provides information about the current thesis, including motivation and scope.

Chapter 2 is called "Aims" and dedicated to state the research questions of the thesis and describes the aims of it. It is also describing aims and goals for each research separately.

Chapter 3 is called “Methods and Materials” and describes the methods, experimental design, and equipment that was used to address the research questions.

Chapter 4 is called “Results” and includes the results of the three papers described in the current thesis.

Chapter 5 is called "Discussion” where the results of the studies included in current thesis are reviewed.

Chapter 6 is called “Conclusion” where a summary of what was done in the thesis is provided and future directions discussed.

## 1.1 Multimodal sensory processing

Multimodal sensory processing refers to the ability of our brain to process information from the different senses simultaneously. It is well known that the different modalities have different speeds of delivering information to the brain and different processing times (Harrar & Harris, 2005, 2008). Our brain, therefore, has to solve many important tasks: temporal and spatial binding of the signals, dealing with sensory dominance, and assigning weights for the information coming from different modalities. However, let us focus on the separate modalities before going deeper into multimodal attention.

### 1.1.1 Visual modality

The visual modality is responsible for most of the information coming to our brain. Note that this applies to humans – other animals may prioritize other senses over the visual one. In what follows, I will briefly explain how the visual system works.

The visual system has the highest resolution of the human perceptual systems. Approximately 1 million fibres in the optic nerve convey information from the retina to the brain (Hall, 1999). The visual perception process can be described roughly in the following way. The first part is the sensation of the information that happens in our eyes, in the retina at the back of the eye. At this stage, physical stimuli (photons) are transformed into neural signals (electrical and chemical). Then, these signals go through the visual nerve toward the thalamus and then to the visual cortex. From the visual cortex, there are two possible pathways: ventral stream (from the visual cortex to the temporal lobe, analysing “what” the object is) and dorsal stream (from the visual cortex to the parietal cortex, analysing “where” the object is). At the end of this process, our brain (usually) recognises the object that was perceived (Mishkin et al., 1983).

The neural and biological sides of the question play a minor role in the current work. The major focus is on the cognitive processes connected to visual modality and their relation to other modalities. Visual attention is a broad cognitive process that helps humans filter irrelevant objects from the world around them while spotting and amplifying relevant objects (McMains & Kastner, 2009). Visual attention has two main stages: preattentive and attentive (see for review, Kristjánsson & Egeth, 2020). Anne Treisman's Feature Integration theory is one of the most well-known theories of visual attention. In the theory, Treisman stated that certain feature dimensions are processed separately and in parallel at the so-called preattentive stage. In the attentive stage, these dimensions are bound together. In their studies, A. Treisman and her colleagues found out that a target defined by a single feature pop out from other stimuli within a visual search array, while a target defined by a conjunction of the features requires more search time. For example, a red circle among green circles is found faster and is easier to find than a red circle among red squares and green circles. Finding targets defined by a conjunction of two features required more time and effort.

Moreover, the number of distractors (or the set size) did not affect the search time in the feature condition. In contrast, search time increased linearly in the conjunction condition as the number of targets increased. This pattern shows the attention-demanding feature binding (Treisman, 1977; Treisman & Gelade, 1980).

Jeremy Wolfe's guided search model involves a modification of the feature integration model and has practical implications in real-world scenarios. This model proposes that visual attention operates through a combination of bottom-up (attention is guided by the salience of stimuli) and top-down processes (attention is guided by the observer's goals) (Wolfe, 1994; Wolfe et al., 1989). It also suggests that unique object features are processed in parallel in the pre-attentive stage by feature maps. For example, if our goal is to find a red vertically oriented line, the colour feature map will highlight all the areas where the colour red appears, and the spatial orientation map will highlight all the areas with vertically oriented objects. These activations will be bound together, and attention will be directed to the area with the highest activation on colour and spatial maps.

The components of the guided search model involve bottom-up processing that refers to the automatic and involuntary capture of attention by salient features in the environment, such as colour, motion, or orientation. These features "pop out" and attract attention without conscious effort. In contrast, top-down processing involves the use of prior knowledge, expectations, and goals to guide attention in a more controlled and deliberate manner. Top-down processes influence the selection of relevant information based on task demands and internal goals (Wolfe, 2021).

Visual search tasks are instruments that allow us to study cognitive processes. In this paradigm, participants perform a variety of different tasks. They can search for one target (often asked to determine whether a target is present or not) or multiple targets. The search task could be based on the features of the targets, conjunction, or space (Á. Kristjánsson, 2015).

In the current work, I will use the so-called foraging task, which involves visual search for multiple targets. The term "foraging" originates from the behaviour of animals searching for and collecting food resources in their natural habitats. In experimental settings, foraging tasks simulate this behaviour by presenting participants with a series of choices or opportunities to gather resources or rewards (T. Kristjánsson et al., 2018). Usually, foraging tasks include visual search arrays where participants need to search for a number of targets that belong to a predefined category among distractors. Through the current paradigm, it is possible to study different cognitive mechanisms in humans (Bella-Fernández et al., 2023; Gil-Gómez De Liaño & Wolfe, 2022; Á. Kristjánsson et al., 2019; Ólafsdóttir et al., 2020, 2021).

A crucial question for the current work is connected to temporal synchrony and its effect on visual attention. This synchrony can appear in the stimulus array and at the neural level. It plays a large role in establishing and maintaining stable object representation (Blake & Lee, 2005; Treisman, 1999). How can participants possibly perceive visual targets with a synchronously changing colour? Moreover, will such synchronization improve visual search or foraging? One possible option is that targets may undergo global segregation from asynchronous distractors, potentially attracting attention in a bottom-up manner similarly to other unique features such as colour or orientation singletons. Evidence from the texture segmentation literature supports this notion, suggesting that rapid detection of temporally defined object boundaries can occur (Blake & Lee, 2005; Fahle, 1993; Farid, 2002; Leonards et al., 1996). Alternatively, mid-level visual grouping principles, such as synchrony, may contribute to perceptual organization (Alais et al., 1998; Lee & Blake, 1999; Wagemans et al., 2012). If global segregation occurs, target items can be expected to be quickly and easily located, leading to consistent response times and irregular spatial search patterns akin to feature-based foraging.

Another possible option is for synchronous targets to benefit from local grouping rather than global segregation. As observers select a target, they may become aware of synchronized items nearby, guiding their search towards these items. This local grouping could rely on either bottom-up or top-down mechanisms, resulting in sporadic movement between local patches of synchrony. While search speed would still be faster than the asynchronous condition, the serial nature of local grouping would lead to variations in inter-target times (ITTs) as the trial progresses.

In our studies, visual synchrony was sometimes accomplished with an auditory stimulus that appeared in synchrony with visual ones. However, how does the audio modality work? How do we perceive audio information? The attempt to answer these questions will be provided in the next section.

### 1.1.2 Auditory modality

The auditory modality processes information faster than all other perceptual modalities (Bizley & Cohen, 2013). This interesting fact can be explained by the distance that the signal needs to go from the basilar membrane, where physical pressure from the sound transduces into a neural signal that that is then conveyed to the auditory cortex located in the superior temporal lobe.

Even though the auditory modality is the fastest in terms of delivery of information, the number of neural fibres in the cochlear nerve is considerably smaller than in the visual system. There are roughly 30.000 fibres in the cochlear nerve compared to the roughly 1 million fibres in the optic nerve.

The loss of hearing dramatically reduces the quality of life. Cochlear implants (CI) can help patients restore their hearing, at least partially. A typical CI produces a sound sensation by stimulating the auditory nerve directly with electrical impulses from 22 micro-electrodes implanted in the inner ear. Remarkably, CIs can give a near-perfect understanding of speech in listening environments that do not involve external noise and have helped more than 600,000 deaf patients worldwide. However, CIs cannot restore normal hearing. CI users have great difficulties understanding speech when there is background noise and are unable to follow musical melodies as well as normally hearing listeners. CI users typically report perceiving music as unpleasant and out of tune.

Increasing music enjoyment for people with CI through stimulation through the tactile modality is a developing field nowadays (Nimmons et al., 2008; Prevoteau et al., 2018; Wright & Uchanski, 2012). In the next chapter the main focus will be on the tactile modality.

### 1.1.3 Tactile modality

The tactile modality is the least studied and, therefore, least well understood perceptual modality. There are many ways of studying the visual modality. Such studies can be performed on monitors, virtual reality glasses or even with physical displays. The analogous situation occurs regarding the audition and the ears. Ears are complicated organs, but they have specific localization. They can be reached and studied. Meanwhile, the skin is the largest organ of the human body. At the same time, skin has the lowest density of the fibres responsible for information (Gardner & Martin, 2013). The resolution of our tactile sensory system greatly varies by body part. Stimulating our fingertips – the most sensitive part of our body is not the same as stimulating the upper part of our back, with a much smaller density of receptors (Jarocka et al., 2021). The region responsible for the haptic information analysis in our brain is quite large. This region is divided into smaller parts, and the amount of neural hardware differs between these smaller parts. An experiment conducted with the fingertips, will therefore, lead to different results if tested on the lower back. All these facts complicate the investigation of tactile attention and tactile perception.

Another important topic to describe here is the so-called mechanoreceptors. These are the sensory receptors that perceive mechanical stimuli such as pressure, touch, vibration, and stretch. They are found throughout the body. There are several types of mechanoreceptors (Lumpkin et al., 2010):

1) Merkel Cells (Merkel Discs). These mechanoreceptors are found in the basal layer of the epidermis (outer layer of the skin), particularly in the fingertips and lips. They are responsible for pressure perception and participate in texture and fine tactile details perception.

2) Meissner's corpuscles. They are mostly located in the dermal papillae (the thin top layer of the dermis) of hairless skin: fingertips, palms, lips, and feet. These receptors are responsible for light touch and low-frequency vibrations, enabling the detection of tactile stimuli with high sensitivity and precision.

3) Pacinian Corpuscles. They are located deeper than the Merkel cells and Meissner corpuscles in the dermis (the inner layer of the skin) and hypodermis (the bottom layer of the skin, furthest from the surface — dermis). They are also present in joint capsules and other connective tissues. These mechanoreceptors are responsible for deep pressure and high-frequency vibrations. Pacinian corpuscles play a crucial role in proprioception.

4) Ruffini Endings (Ruffini Corpuscles). Ruffini endings are receptors located in the dermis and hypodermis, and they can also be found in joints. Ruffini Endings slowly adapt to skin stretching and sustained pressure. They give information about skin deformation and joint position.

Each mechanoreceptor type possesses unique traits and functions, enabling the body to perceive and differentiate diverse mechanical stimuli in its surroundings. Together, these receptors contribute to the intricate sensory perception of touch and perform vital functions in tactile sensation, proprioception, and detecting mechanical occurrences within the body.

Our research group has conducted studies of tactile attention and tactile perception for a long time. One of our findings is that the mechanoreceptors in the skin are more sensitive to vibrations than pressure (Hoffmann, Brinkhuis, et al., 2018). In this research Hoffman et al. (2019) also compared types of the tactile stimuli and their effectiveness for delivering information. Another important finding was the description of the optimal frequency values that humans perceive most accurately (Ævarsson et al., 2022). Participants were wearing wristbands with the L5 actuator on the inner and outer parts of the wrist. Results showed that the sensitivity pick is 200 Hz for the inner part of the wrist (100-275 Hz optimal range). The wrist pick's outer part was 125 Hz, with an optimal range between 75 Hz and 200 Hz. Hairy skin parts tend to be more sensitive than glabrous skin. The results provided us with parameters which were used in follow up studies.

Later, our group used the peak values from these measurements of the sensitivity range to find the optimal distance between actuators on the forearm (Yeganeh et al., 2022). To answer this research question, we created five different sleeves containing vibrotactile stimulators that observers wore on their forearms (Figure 1). These sleeves differ in the distance between the actuators (as close as possible, 5 mm, 10 mm, 15 mm, 20 mm). All sleeves had 3 actuators placed on a fabric, that allowed the user to wear them easily and adjust the sleeve for different sized forearms. We also compared the effectiveness of stimuli perception on the inner and outer parts of the forearm. Our results suggested that the highest accuracy was achieved when the actuators were placed with at least 2 cm distance between them. Another interesting finding tells us that there is no difference in providing sensory information between the inner and outer sides of the forearm. Note that these findings can be also explained with the high answer accuracy in general. One of the explanations for this, however, may be that the task was too easy to complete, so effectiveness of the stimuli perception was the same. Further testing of this finding is therefore needed for a conclusive answer.

**Figure 1.** The apparatus used in the study of our research group. Sleeves with different distance between actuators.

Another study performed by our group was conducted to find out the locations on the forearm where human perception is the most accurate (Yeganeh et al., 2023). Participants were stimulated with two sequential vibrations using the "best" sleeve from our previous study, where there was a 2cm distance between the tactors. Participants' task was to answer whether the second vibration was at the same tactor, higher or lower, than the first vibration. Another independent variable was the frequency of the stimulations (100 Hz, 150 Hz, 200 Hz, 250 Hz). These frequencies were chosen based on the previous studies of our research group (Ævarsson et al., 2022). The results showed that the most effective regions for perception of tactile information were near anatomical reference points, close to the the body's joints. In this study they were near the wrists and near the elbows. In contrast, the region closer to the middle of the forearm between the elbow and the wrist showed lower accuracy in information perception. Surprisingly, the accuracy did not depend on the frequency manipulations. These results provide insights into the best positions for the tactile stimulation of vibrotactile sensory substitution devices for the forearms.

Our latest study compared the sequential versus simultaneous presentation of tactile information (Yeganeh et al., 2024). For this study, our group created a sleeve with six actuators. The current sleeve was made to send signals that signify letters from the Braille alphabet. In this study my colleagues also compared short patterns (2 and 3 vibrations) to long patterns (4 and 5 vibrations). The participant's task was to recreate the presented pattern. It turned out that simultaneous presentation takes more work to recreate. Another finding showed that participants are significantly less effective in recreating long patterns. This finding means we should search for a more effective way of presenting complicated and extensive information. Overall, these results provide insights into the limits of tactile perception.

Another important part of tactile attention studies is connected to tactile illusions. Our group recently discovered the intensity order illusion (IOI; Hoffmann et al., 2019). The IOI involves a systematic mislocalization of the second stimulation of two stimulation. If the second stimulation with higher intensity follows the first stimulation with lower intensity, the second stimulation tends to be perceived above the first one. The IOI illusion also involves the opposite effect: when the second stimulation has a *lower* intensity than the first one, people will tend to report that the second stimulation was lower than the first one (Hoffmann et al., 2019). More information about this illusion can be found in sections dedicated to Paper I and in the Appendix.

Illusions provide us with insights about the expectations of the brain from the world around us. They also provide information about basic principles of the perceptual system and about the operating principles of perceptual interpretation. There are some well-studied illusions in the tactile modality. An illustrative instance of a perceptual phenomenon within tactile perception is the sensory saltation illusion, also referred to as the cutaneous rabbit illusion (Geldard & Sherrick, 1972). This phenomenon involves rapid, repeated stimulation at adjacent locations on the skin. For instance, a sequence of three taps applied to two skin sites can elicit the sensation of an object hopping along the skin (Flach & Haggard, 2006), resulting in an illusory perception of stimulation occurring between the actual stimulated locations. Similarly, another spatiotemporal haptic illusion known as the funneling illusion (Gardner & Spencer, 1972) involves the perception of two adjacent simultaneous vibratory stimuli originating from the space between the two actuators. These illusions underscore the extensive spatiotemporal interactions (Flach & Haggard, 2006) involved in processing tactile stimulation, potentially reflecting modulations in the activity of the somatosensory cortex (Blankenburg et al., 2006).

Moreover, evidence suggests that well-known illusions observed in different sensory modalities also occur in the tactile modality. For instance, visual illusions, including the Bourdon (Day et al., 1990), Müller-Lyer (Gentaz & Hatwell, 2004; Millar & Al-Attar, 2002), Ebbinghaus illusion (Ziat et al., 2014), have tactile counterparts. Additionally, illusions associated with underestimating the distance between stimuli and overestimating time have been observed in the tactile modality. These include the tau effect (Lechelt & Borchert, 1977), the kappa effect (Suto, Y., 1952), and the apparent haptic movement illusion (Carter et al., 2008; Sherrick & Rogers, 1966), which may be related to the cutaneous rabbit illusion discussed earlier. Cross-modal illusions can be explained by similar operational and perceptual organisational principles that apply for different modalities.

As mentioned before, on very rare occasions in real life, people rely on separate modalities while perceiving the world around them. The perception process is usually cross-modal. The next chapter will intoduce cross-modal perception and multisensory integration of stimuli coming from different modalities.

### 1.1.4 Multisensory Integration

In real-world situations, humans receive information from multiple modalities simultaneously. The auditory and visual modalities can be used simultaneously to estimate an object's shape and distance to it. Sometimes, when we do not know the depth of a river and cannot see the riverbed, we measure it using our feet. Searching for a light switch in a dark room will also require our tactile modality. Each modality, each way of perceiving the world around us, has its strengths and weaknesses. In humans, several classical studies have shown that the visual modality dominates in spatial tasks (McGurk & McDonald, 1976), while the auditory modality dominates in temporal tasks (Welch et al., 1986).

Despite dominating one of the modalities dependently on the task, there is research showing that modalities can work together to increase the effectiveness of visual orienting, for example when we are searching for a particular target. The so-called "Pip-and-pop" (Van der Burg et al., 2008) and "Poke and pop" (Van der Burg et al., 2009) phenomena, described by Van der Burg and his colleagues, are great examples of multisensory integration. In this research, the participants searched for a vertically or horizontally oriented target stimulus among differently oriented distractors. All the stimuli on the screen were changing colour back and forth at different rates. In a crucial manipulation, Van der Burg and his colleagues changed the colour of the target stimulus synchronously with non-spatial sound (the sounds did not come from any specific location of one of the items, Van der Burg et al., 2008). In another study, the sound was replaced with a non-spatial vibration that was presented on the back of the left hand (Van der Burg et al., 2009). In both studies, search time was significantly faster in the conditions where the colour change of the target was synchronized with stimuli from another modality. For audio-visual integration, the authors argued that the interaction appears in the early processing stages without any specific effort. Moreover, they proposed that this interaction can appear in the early visual cortex. They reached the same conclusion for tactile perception in the "Poke-and-pop" effect (Van der Burg et al., 2009). In their conclusions, Van der Burg and his colleagues refer to earlier work conducted by Violentyev et al., which shows that participants' perception of a visual stimulus (flash) was affected by two touches (Violentyev et al., 2005). So, participants saw two flashes even though only one was presented to them if this occurred simultaneously with two haptic stimulations. All these authors concluded that their research provides strong evidence that visual information processing is not unimodal and can be affected by other modalities.

Karla Evans and Anne Treisman have also studied multisensory integration in a series of Experiments (Evans & Treisman, 2011). In their study Evans and Treisman presented several experiments showing the connection between high tone pitch and bright surface (Martino & Marks, 1999), small sizes (Gallace & Spence, 2006; Mondloch & Maurer, 2004), angular shapes (Marks, 1974), arrows pointing up for six-months-old kids (Wagner et al., 1981). They gathered evidence showing that a multisensory association appears between a high-pitch tone and several features, such as the small size of the visual stimuli and the high position of stimuli. Participants were asked to judge different features of the visual object (size, position, density of the lines inside a circle, contrast of the circle) and the type of sound (high vs low pitch tones played either on a violin or a piano). Evans and Treismsn investigated the level of the possible integration if there is any. Two types of experiments were conducted to address this question involving either "direct" or "indirect" tasks. Participants were asked whether the pitch was high or low in the "direct" task. In the "Indirect" task, the question was about the type of instrument playing the sound (piano or violin). The instruments' sound could be as high or low as the pitch in the "direct" task. The association between the "high" pitch and small size and high position appeared in both types of tasks. These results mean that these associations between visual and audio modalities appear automatically and affect the task even if the sound task is to define whether violin or piano, which is an indirect task in this context.

Nevertheless, the results of Evans & Treisman also showed that, at the same time, this integration was more effective when the word for the type of tone was the same as the position of the visual stimulus (high-high). This work shows that multisensory integration can have influences on different aspects of information processing and decision-making. The current findings may also be part of the explanation of the intensity order illusion (Hoffmann et al., 2019; Makarov, Stefánsson Thors, et al., 2024), where high amplitude stimulation following a low stimulation at the same actuator will most likely be perceived as upward movement.

Despite the wealth of research in recent decades, the question of how the brain combines information from different modalities remains unclear. Spence discussed this question in his review paper (Spence, 2011). Spence empathised the importance of cross-modal correspondences in shaping perception, cognition, and behaviour, and emphasized the need for further research to explore their mechanisms and practical applications in various domains. It is clear that cross-modal correspondence can appear on different levels. The following are the variants outlined in Spence (2011):

1) Structural Correspondences. These correspondences are based on similarities in the physical attributes of stimuli across different sensory modalities. For example, higher-pitched sounds are often matched with brighter colours, while lower-pitched sounds are associated with darker colours.

2) Statistical Correspondences. This level of interaction is based on the co-occurrence of sensory features in the environment. For instance, in English, we often refer to smaller objects using higher-pitched sounds (e.g., "small" vs. "big"), and this pattern is found across many languages.

3) Semantic Correspondences. These correspondences are based on the meaning of the stimuli. For example, sweetness in taste is often associated with round shapes and high-pitched sounds, while bitterness is associated with angular shapes and low-pitched sounds.

4) Emotional Correspondences. These correspondences are related to the emotional content of the stimuli. For example, bright colours and major chords in music are generally associated with positive emotions, while dark colours and minor chords are associated with negative emotions.

5) Learned Correspondences. These correspondences are based on individual or cultural learning. For instance, in some cultures, red is associated with danger and green with safety, influencing, for example, how people perceive traffic lights.

6) Developmental Correspondences. These are correspondences that are based on developmental stages. For example, children often associate high-pitched sounds with smaller or more delicate objects and low-pitched sounds with larger or more robust objects.

7) Synesthetic Correspondences. These correspondences involve a blending of sensory experiences, as seen in synesthesia. For example, a synesthete may perceive specific colours when they hear certain musical notes.

By understanding these levels of interaction, researchers can better understand how our senses work together and how they can influence each other. Nevertheless, this knowledge is already widely used in sensory substitution technologies.

## 1.2 Sensory substitution

The section on *sensory substitution* in this thesis involves two main parts: discussion of sensory substitution mechanisms and of sensory substitution devices. The first subchapter will discuss the physiological and psychological processes underlying sensory substitution mechanisms. The second part will cover physical sensory substitution devices. These devices differ from, for example, classical cochlear implants. Sensory substitution devices use information from one modality to exchange lacking information from a deficient modality. In addition, the application of these devices can be wider than used within medicine. For example, the Sound of Vision device, recently designed in our laboratory (Brooks et al., 2023; Hoffmann, Brinkhuis, et al., 2018) is intended for use for visually impaired individuals where information from head mounted cameras is conveyed to the user via haptics and audition. All this information will be covered in the second part (1.2.2).

### 1.2.1 Sensory substitution mechanisms

What enables our body and brain to substitute another modality for a compromised one? The compensation depends on the time when the ability to perceive information through this modality was lost. Compensatory sensory substitution mechanisms typically do not work as well for grown-up people as for children beacause of how neuroplasticity becomed much worse as people age (Collignon et al., 2006). Children's brains are generally more plastic or adaptable than adult brains, meaning they may have an easier time learning to use new sensory substitution devices or adapting to changes in sensory input. Nevertheless, adults can also benefit significantly from sensory substitution, particularly if they receive appropriate training and support (see e.g. Hoffmann et al., 2018).

The main principle underlying the compensatory mechanism is that working sensory modalities take on some functions of the dysfunctional modality and in addition become more proficient in processing information. This process occurs at different levels of the nervous system. Árni Kristjánsson and his colleagues made a broad review of the sensory substitution mechanisms that provide different insights about the nature of these processes (Á. Kristjánsson et al., 2016).

Neuroimaging and transcranial magnetic stimulation research provide evidence of cross-modal plasticity, indicating that brain regions conventionally associated with specific sensory modalities, such as vision or audition, can flexibly engage in processing information from other modalities (Merabet & Pascual-Leone, 2010; Pasqualotto & Proulx, 2012).

Brain regions traditionally labelled as sensory-specific, like the visual cortex, may possess supermodel functions, capable of responding to inputs from diverse sensory modalities, suggesting an integrated processing system (Ricciardi et al., 2014). For example, tactile stimulation can activate regions associated with vision, revealing functional overlap and integration among sensory systems (Poirier et al., 2007; Renier et al., 2005).

Studies demonstrate that sensory inputs from one modality can activate brain areas typically associated with another, suggesting cross-modal interactions in perception (Ortiz et al., 2011; Sadato et al., 1996). In blind individuals, tactile tasks (such as Braille reading) can trigger activation in the visual cortex, indicating adaptive processing of non-visual sensory information (Striem-Amit et al., 2015). Additionally, the so-called fusiform face area in the temporal lobe, that is selective for processing visually presented faces in those with intact vision, becomes activated when visually impaired people touch faces (Ratan Murty et al., 2020).

Research into sensory substitution techniques, such as Braille reading and auditory sensory substitution, illustrates the brain's capacity to adapt to alternative sensory inputs (Amedi, 2002; Amedi et al., 2001; Bedny et al., 2015). This adaptive plasticity can, for example, involve recruiting visual cortex for processing tactile or auditory information, enabling individuals with visual impairments to perform tasks typically associated with sight (Ptito et al., 2009; Reich et al., 2011). These discoveries are very important. They help us to understand of sensory processing. They also give us ideas for new ways to help people who have trouble with their senses. For example, we can use what we know to create new methods to help them, and design advanced sensory substitution devices (Calvert et al., 1999; Kupers et al., 2010). By understanding how the brain handles it when a sense doesn't work, we can make treatments to help people improve their senses.

While the human body can demonstrate remarkable adaptability, it is not always able to fully compensate for a compromised modality. Therefore, the development of sensory substitution devices is crucial. These devices can significantly improve the quality of life for blind or deaf individuals, offering them a more comfortable and independent lifestyle.

### 1.2.2 Sensory substitution devices

Our research group has worked on the development of sensory substitution devices (SSD) for several years. The main principle behind these devices involves converting information from one compromised modality to another. What are the most common variants of SSD for visually impaired people?

The most common variants of SDDs are connected to the audio modality. Efforts to assist the visually impaired using auditory cues have explored echolocation, which can be hindered by noise concerns, and camera-based auditory feedback systems, offering a more feasible alternative (Ifukube et al., 1991; Kolarik et al., 2014; Loomis et al., 1998, 2005; Marston et al., 2006, 2007; Van Erp, 2005). Virtual sounds, rather than speech, have shown better effectiveness in aiding navigation for the blind (Loomis et al., 2005). The PSVA device, linking captured image pixels to sound frequencies, activates visual cortex regions in early blind individuals during pattern recognition tasks (Arno et al., 2001). Similarly, the EyeMusic device translates visual data into musical notes, enhancing spatial perception and movement accuracy in blindfolded subjects (Abboud et al., 2014; Levy-Tzedek et al., 2012). Moreover, the EyeCane, featuring infrared sensors for distance and obstacle detection, significantly improves navigation skills for the visually impaired (Chebat et al., 2015; Maidenbaum et al., 2014).

Another variant of SSD involves haptic devices and revolves around utilizing haptic feedback to provide environmental information to individuals with visual impairments. These devices face challenges such as touch intricacies and device practicality. Early endeavours, exemplified by the work of Bach-Y-Rita et al., displayed potential but encountered limitations (Bach-y-Rita et al., 1969). Research indicates that while the hands offer high spatial acuity for haptic feedback, usability remains a concern. Promising approaches include employing haptic belts for directional cues and tongue-based devices for tactile stimulation. Studies like those conducted by Kupers et al. (2010) demonstrate the potential of training blind individuals in navigation tasks.

**Figure 2.** The apparatus used in the SOV device of our research group. Left top photo: The tactile belt prototype version 4 from our Sound of Vision project. Left down photo: closer view on the prototype version 4. Right photo: A person wearing the prototype version 4.

Our research group has also conducted research connected to navigation tasks. In the Sound of Vision (SOV) project, the task was to create a device that would scan the visual information around the user, encode this information into tactile and audio signals, and send it to the skin and ears. One of the problems that were faced during this project was connected to sensory input limits. It was very important to find the right way not to overload the user's sensory system (Hoffmann, Spagnol, et al., 2018). The SOV system consisted of headgear with three-dimensional cameras and sound presented throughthe headphones, a tactile belt with 60 vibrating actuators at the abdomen, and a computer in a backpack (meant for the prototype only). The header with cameras works as a perception organ, substituting eyes, while the laptop processes the data and transforms it into audio and tactile signals. A learning procedure that lasted 8 hours was run for six visually impaired people. Figure 2 shows the sensory substitution device.

The participant's task was to avoid physical obstacles in the space relying on the SOV system. Results showed that the learning process was successful (see e.g., Hoffmann, Spagnol, et al., 2018). Participants were also asked to avoid collisions in the changing environment. Importantly, the results showed that participants were learning how to use the auditory and tactile information from the SOV device, instead of simply learning to navigate the space with memory.

A major issue with combining haptic and audio signals involves the danger of overwhelming the attentional and perceptual systems regarding the amounts of information coming to the brain simultaneously. Humans have a finite attentional capacity, and important information may, therefore, be missed because other information is selectively attended (Broadbent, 1958; Á. Kristjánsson, 2006; Most et al., 2005; Treisman, 1960). This phenomenon transcends sensory modalities, impacting vision (Grimes, 1996; McConkie & Zola, 1979), haptic perception (Gallace & Spence, 2006), and audition (Vitevitch, 2003). Attentional strain resulting from sensory substitution devices may divert attention from other environmental cues (Kärcher et al., 2012), emphasizing the need for designs with minimal attentional demands (Liu, 2001). Effective sensory substitution devices must avoid overwhelming users with information and should not disrupt the natural sensory environment (Härmä et al., 2004). Neglecting attentional constraints can limit the practicality of such devices, underscoring the importance of understanding sensory and attentional mechanisms in their development (Loomis et al., 2018).

Many questions are connected to the most comfortable and efficient usage of the SSDs. For example, what is the best place to use a tactile device? The most common wearables for SSDs involving tactile stimulation are bracelets. However, these devices can provide only limited information with a small resolution. The most sensitive parts of our body are the fingertips (Jarocka et al., 2021). The problem is that if you create a device, that is sending the information to the fingertips, then you will lose the ability to operate with your hands properly if used concurrently with the device. That is why one of the very important questions in developing tactile devices is searching for the optimal balance between the comfort of the user and the most effective way to deliver information.

## 1.3 The scope of this thesis

This thesis investigates the mechanisms underlying tactile perception and the effects of multimodal attention on human performance. It aims to enhance our understanding of how different sensory modalities interact and how this knowledge can be applied to improve sensory substitution devices (SSDs) and multimodal interfaces.

The motivation for this research is driven by the need to bridge gaps in our understanding of sensory interactions and leverage this knowledge for practical applications. Enhancing SSDs and tactile communication devices is crucial for improving accessibility and quality of life for individuals with sensory impairments. Additionally, optimizing multimodal interfaces can significantly benefit various professional and everyday applications, ensuring information is presented in the most accessible and efficient manner.

The research begins by exploring the Intensity Order Illusion (IOI) in tactile perception, specifically examining whether the illusion is caused by changes in amplitude or frequency of sequential stimulations. Utilizing advanced tactile actuators that allow independent manipulation of these parameters, the study seeks to isolate the factors contributing to the IOI. Additionally, the thesis assesses the generalizability of this phenomenon across different directional stimulations and various types of actuators.

In addition to the unimodal study of tactile perception, the thesis investigates the effects of visual and auditory synchrony on attention and perceptual performance in visual foraging tasks. Inspired by the "pip & pop" phenomenon, the research examines how synchrony in visual and auditory stimuli can guide attention, improve search accuracy, and reduce search time. The interaction between these modalities is further explored to determine the most effective methods of presenting information through various sensory channels. Building on these findings, the thesis also examines the role of cross-modal cues—visual, auditory, and haptic—in enhancing the detection of synchronized targets during complex multitarget foraging tasks. The research aims to identify the optimal ways of presenting information through different modalities, considering the complexity of the task and the effectiveness of unimodal versus multimodal cueing.

The motivation for this research is driven by the need to bridge gaps in our understanding of sensory interactions and leverage this knowledge for practical applications. Enhancing SSDs and tactile communication devices is crucial for improving accessibility and quality of life for individuals with sensory impairments. Additionally, optimizing multimodal interfaces can significantly benefit various professional and everyday applications, ensuring information is presented in the most accessible and efficient manner.

The work presented in this PhD thesis was split into two parts: basic scientific research into perceptual mechanisms and applied research of SSDs. In the so-called scientific part, the focus was on studying cross-modal attention. I also studied how we perceive information through our skin and how the interaction between different modalities can affect visual search. The second part of the work in this thesis was dedicated to solving applied research questions. My research group was searching for the best way to deliver information, developing prototypes, and evaluating the effectiveness of information delivery. Further scientific research conducted by my colleagues and me will be described.

# 2 Aims

The title of this thesis was chosen to indicate that it is divided into two main parts. The first part focuses on fundamental research of multimodal attention. Detailed knowledge of the perceptual mechanisms is crucial for developing effective and usable sensory substitution devices. The second part of this thesis involved developing and testing sensory substitution devices. My research group tested different actuators for our devices. Our group designed and assembled the equipment for our experiments. This equipment can likely be used for various purposes in the future, such as navigation for visually impaired people or helping people with prosthetics to understand the positions of their prosthetic leg or arm in space.

The current thesis will address the following research questions:

1) What are the underlying mechanisms of the Intensity Order Illusion (IOI) in tactile perception?
2) How do visual and auditory synchrony influence attention and perceptual performance in foraging tasks?
3) What are the effects of cross-modal cues on the detection of synchronized targets in complex visual foraging tasks?

The current thesis is mainly focused on my research as a first author, describing research that I led, so further in the text, only the aims of these papers will be described. A description of the research conducted by my colleagues and where I am the second author can be found in the Introduction.

## 2.1 Paper I: The Haptic Intensity Order Illusion Is Caused by Amplitude Changes

The first paper is part of the unimodal research about tactile attention. The main purpose of the paper was to study a phenomenon called the Intensity Order Illusion (IOI). This phenomenon was discovered by our group (Hoffmann et al., 2019). This haptic mislocalization effect appears in the lower back with two sequential stimulations in the same location. If the second stimulation has a higher intensity than the first one, it was perceived as an upward movement. When the second stimulation had a lower intensity then, it tended to be perceived as a downward movement. I tried to reach the following aims in the following research:

**The first aim** of the research was to study whether the IOI is caused by amplitude changes or frequency changes. The actuators that were used in the original report on the illusion provided eccentric rotating force, both normal (NERMS) and parallel (PERMS) to the skin surface, but notably, these actuators did not allow independent manipulation of frequency and amplitude. This meant that we could not address the contribution of each of these parameters (amplitude or frequencey) to the illusion. For this reason, in the current study, I used Lofelt L5 actuators, which allowed us to manipulate frequency and amplitude separately (*L5 Actuator Description*, 2019).

**The second aim** involved psychophysical tests of the L5 actuators. The plan is to use L5 actuators to convey information for cochlear implant users or provide tactile information through sensory substitution devices. There are several options of how they can be used: music enjoyment, tactile attention studies, helping visually impaired people to navigate, helping people with prosthetics, developing the wearables for professional needs, e.g., firefighters who often have to operate in low visibility conditions (such as in smoke-filled buildings).

**The third aim** was to investigate whether this illusion appears in the horizontal direction, as well as vertically, as tested originally. This knowledge will be used in developing SSD since it is important to search for the optimal way to present tactile information.

**The fourth aim** of our study was to replicate the IOI using a different set of tactile actuators. It cannot be emphasized enough how crucial it is to check and replicate different studies and test for their generalizability.

## 2.2 Paper II: The effects of visual and auditory synchrony on human foraging

The second study involved multimodal research. It was inspired by the so-called "pip & pop" phenomenon found for single-target search (Van der Burg et al., 2008). In this phenomenon, the participant's task was to search for vertically or horizontally oriented targets. All the stimuli changed colour independently from each other. A non-spatial sound that appeared in synchrony with the target colour change reduced the search time of the target (Van der Burg et al., 2008).

In this paper, we studied the synchrony effect on multiple-target visual search. This study was aimed at analysing whether synchrony in stimulation can guide attention and aid perceptual performance within a foraging task. The main question was whether visual synchrony can speed up the visual foraging and improve accuracy. Additionally, we studied if audio synchrony can have any additional effect on the visual search for multiple targets.

This research was conducted to broaden the knowledge about the multimodal attention and the interaction between different modalities. The results of this paper offer valuable insights into the most effective methods of presenting information through various modalities.

## 2.3 Paper III: Cross-modal cues improve the detection of synchronized targets during human foraging

Based on the results of "The effects of visual and auditory synchrony on human foraging," the third study of this thesis was designed. Findings from the second study indicated that there were no clear benefits to foraging from audio cues. This result contradicts what could be expected based on previous visual search studies with multimodal cues (Gao et al., 2021; Van der Burg et al., 2008; Zou et al., 2012).

One of the possible reasons for the lack of multimodal aid in the second study was that the task was too easy. Participants relied only on the visual modality and did not use multimodal cueing. Another potential reason is that participants were not searching for synchrony. Synchrony itself was irrelevant to the task. Based on this outcome, in the current study participant's task was to search for synchrony, considering synchronisation from different cues.

This study aimed to investigate what kind of cues will be the most effective in facilitating the multitarget visual foraging task, while the task itself will be too complicated to rely only on the grouping effect as it was in the previous study.

Another aim was to compare the effectiveness of unimodal and multimodal cueing. In the current study, visual cueing was compared to audio cueing and haptic cueing. Current knowledge provides important insights about the most effective way of presenting information through different modalities.

# 3 Materials and Methods

The experiments' designs were different in all three studies. The following provides brief information on the materials and methods used in all three papers. Expanded information can be found in the original publications section.

## 3.1 Paper I: The Haptic Intensity Order Illusion Is Caused by Amplitude Changes

Two Experiments were conducted in the current study. Both experiments were conducted in the laboratory within an anechoic chamber. The experiments were approved by the National Bioethical Committee of Iceland (VSN-15-107) and the methodology conformed to the Declaration of Helsinki.

### 3.1.1 Paper I. Participants

Seventeen people participated in Experiment 1. All participants were naive to experimental hypothesis. Seven participants reported themselves as women, ten participants reported themselves as men. The age range was between 21 and 41 years old (7 women, mean = 30.5 yr, SD = 6.8 yr).

Ten people participated in Experiment 2. All participants were naive to the experimental hypothesis. Four participants reported themselves as woman, and six participants reported themselves as man. The age range was between 20 and 34 years old (mean = 27 yr, SD = 4.3 yr)

### 3.1.2 Paper I. Apparatus

The same equipment was used in both Experiments. A custom-built 4x4 array of Lofelt L5 actuators was used to stimulate the lower back. Figure 3a shows a photo of the custom-built array; Figure 3b shows the apparatus placed on the participant's lower back.

Both experiments were programmed in a Python environment. Our lab-tactile research hardware system generated signals and transferred them to the equipment. This system consists of an RME MADIface XT audio interface, two Ferrofish A32 AD/DA converters, and a custom-made multi-channel 2 W amplifier. The vibration calibration measurements were performed using a Bruel and Kjaer 4507 accelerometer and a Bruel and Kjaer 2270 handheld sound analyser. Participants responded during the experiment with a standard keyboard.

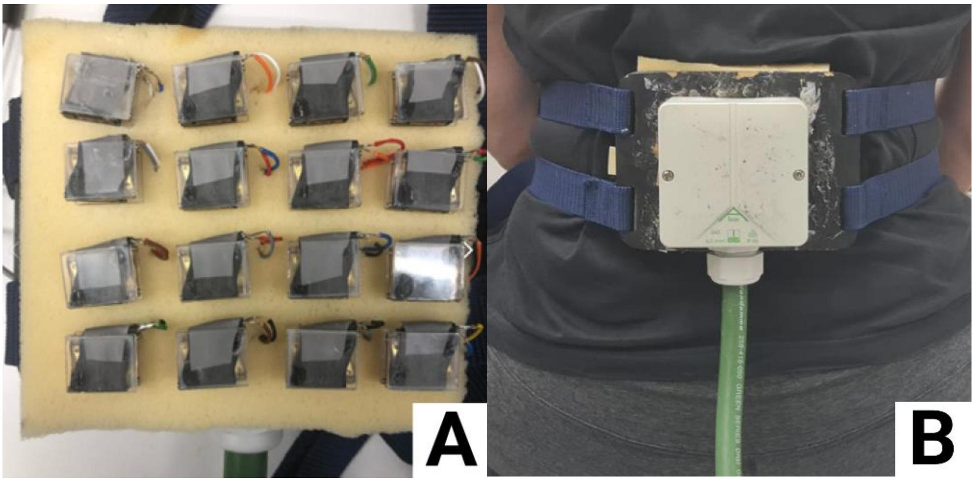

**Figure 3.** The apparatus used in Experiments 1 and 2 of Paper I. (A) The mounting base with a 4 ×4 array of L5 actuators placed at 10-mm centre-to-centre distance. (B) One mounting base at a time was strapped around the participant’s waist so that the tactor array stimulated the lower thoracic region.

### 3.1.3 Paper I. Stimuli and procedure

In Experiment 1, the participant´s task was to judge the direction of the vibrotactile stimulation: whether the second stimulation was above or below the first one. Four possible stimuli combinations were used: i) low frequency and low amplitude, ii) low frequency and high amplitude, iii) high frequency and low amplitude, and iv) high frequency and high amplitude. Another important independent variable was the position of the second rotor vibration, compared to the first one: higher, lower or the same actuator vibrated as for the first stimulation. There were four possible combinations of the first and second stimuli: strong-strong, weak-weak, weak-strong, strong-weak. There were 1152 trials in the experiment, divided into four blocks. Participants could take pauses between blocks. The experiment took around 90 minutes in total. Figure 4 shows an example of the stimulation patterns used in Experiment 1.

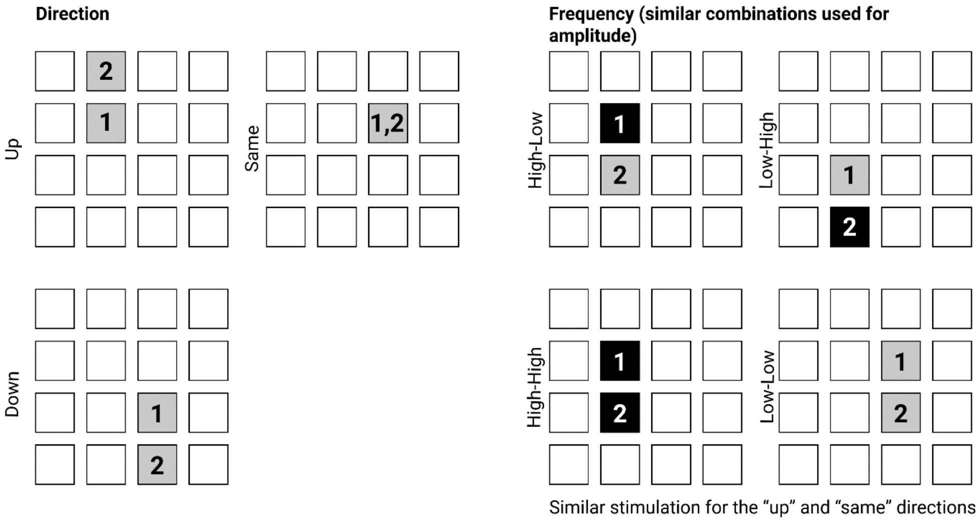

**Figure 4.** Examples of stimulation used in the Experiment 1 of Paper I. The squares represent the 4x4 tactor array, and the marked squares represent the actuators activated in the sequence indicated by the numbers. One trial consisted of 2 successive vibrotactile stimulations, which varied in direction, frequency (high frequency is shown in black and low frequency is shown in grey) and in amplitude (strong-weak, weak-strong, strong-strong and weak-weak, not shown in figure).

In Experiment 2, the participant´s task was to judge whether the second stimulation was to the right or the left from the first one. Based on the results of Experiment 1, only amplitude was used as the independent variable with a constant 200 Hz frequency. The amplitude could be strong or weak. As in Experiment 1, the second stimuli had three possible positions: to the right, to the left, or the same actuator activated as the first one. Similar to the first experiment, the possible combinations of the first and second rotor activation were strong-strong, weak-weak, weak-strong, and strong-weak. There were 280 trials divided into four blocks. Participants could take breaks between blocks. The experiment took around 30 minutes in total. Figure 5 shows an example of the stimulation for Experiment 2.

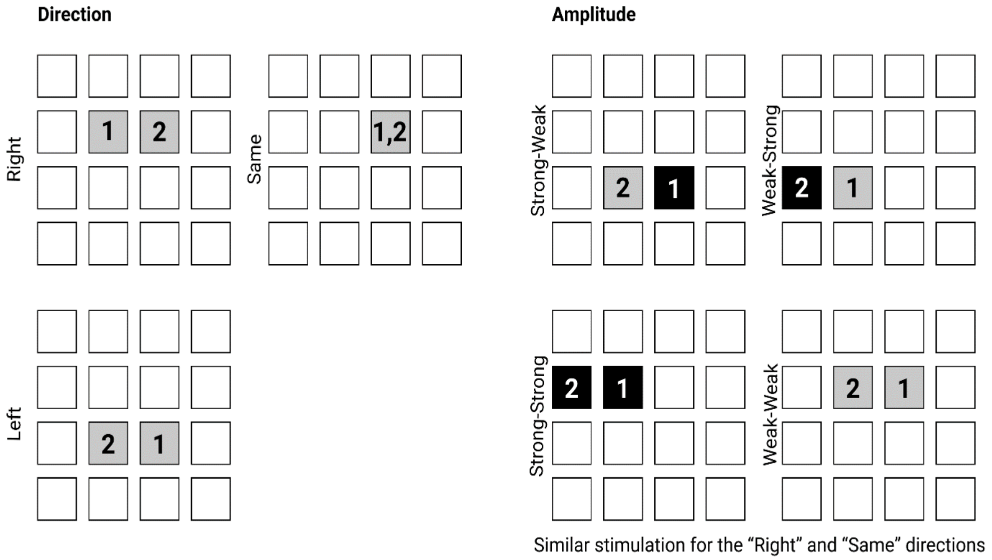

**Figure 5.** Examples of stimulation patterns in Experiment 2 of Paper I. The squares represent the 4x4 tactor array, and the marked squares represent the actuators activated in the sequence indicated by the numbers. One trial consisted of 2 successive vibrotactile stimulations, which varied in direction and in amplitude (strong-weak, weak-strong, strong-strong and weak-weak, shown in black and grey for strong and weak stimulation respectively, on the right in the figure).

Low-frequency values were equal to 100 Hz, and high-frequency values were equal to 200 Hz. For the low amplitude, acceleration was 54.29 $m/s^2$ for the 100 Hz frequency and 48.52 $m/s^2$ for the 200 Hz frequency. For the high amplitude, acceleration was 96.48 $m/s^2$ for the 100 Hz frequency and 91.24 $m/s^2$ for the 200 Hz frequency. These slightly differing acceleration values reflect that to ensure that the vibrotactile stimulus acceleration was equal for the different tested frequencies, the signal amplitudes were adjusted based on vibration acceleration measurements so that both signals (low frequency and high frequency) resulted in the same vibration intensity on the skin.

The procedure was the same for both Experiments. Participants signed an informed consent form and then were asked to sit in the anechoic chamber wearing headphones with white nose to block the noise from the actuators. In the lower thoracic region, participants were wearing a belt, with actuators placed equally on either side of the spine.

After the Experiment began, participants received two consequent stimulations (200 ms each with 50 ms intervals between them). After that, participants had 5 seconds to respond using the usual keyboard (arrows "up" and "down" for Experiment 1; arrows "right" and "left" for Experiment 2).

### 3.1.4 Paper I. Data analysis

One participant was excluded from the final analysis of Experiment 1 and another participant was excluded from the final analysis of Experiment 2. These two people did not seem to understand the task. Both datasets were analysed using the Rstudio program (RStudio Team, 2020).

For Experiment 1, we expected that weak stimulation followed by strong stimulation would increase the probability of up responses and vice versa for the strong stimulation followed by weak. The illusion should occur when the second stimulation is in the same location as the first.

For Experiment 2, we conducted an exploratory study to check whether the IOI occurs horizontally.

All the participants' answers were recalculated as the proportion of the responses as up. Data distribution was checked using the Shapiro-Wilks method. In Experiment 1, a three-way Repeated Measures ANOVA was used: the frequency combinations for the first and second stimulations (high-high, high-low, low-high, low-low), the amplitude combinations for the first and second stimulations (strong-strong, strong-weak, weak-strong, weak-weak), and the direction of the stimulation (up, same, down). In Experiment 2, a two-way Repeated Measures ANOVA was used: the amplitude combinations for the first and second stimulations (strong-strong, strong-weak, weak-strong, weak-weak), and the direction of the stimulation (up, same, down). Pairwise t-tests were used to perform post-hoc comparison tests. Data analysis was made with the R packages called tideverse, ez and rstatix (Kassambara A., 2021; Lawrence MA, 2016; Wickham et al., 2019).

## 3.2 Paper II: The effects of visual and auditory synchrony on human foraging

Three experiments were presented in this study. Data collection for the first and the third Experiment was conducted online, while the second Experiment was run in the lab. All the experimental procedures conformed to the Ethics and Data Protection guidelines of the University of Malta and the University of Iceland for conducting online research. The methodology of the Experiments also conformed to the declaration of Helsinki. All participants confirmed they had read the informed consent and understood the task. Participants signed the hard copy of the informed consent form for Experiment 2.

### 3.2.1 Paper II. Participants

Twenty-four people participated in Experiment 1, 12 for each type of condition: 11 females (19-29 years, M=23.66 years) and 13 males (18-39 years, M=25.21 years). Participants were recruited online from https://prolific.co. They received approximately 300 ISK for a 15-minute session.

For the second experiment, which was run in the laboratory, 12 participants were recruited: females (aged 22-33 years, M=24.5 years) and 4 males (aged 18-23 years, M=21.25 years). Participation was completely voluntary, with no payment provided.

Twelve participants took part in Experiment 3: 7 females (19-39 years, M=28.2 years) and 5 males (22-37 years, M=24.86 years). Participants were recruited online from https://prolific.co as in Experiment 1.

In all three experiments, participants reported being native speakers of English and having normal or corrected to normal vision. Participants were informed that they were free to end their participation in the experiment any time they wanted.

### 3.2.2 Paper II. Apparatus

Experiments 1 and 3 were run online, so there was no control of the precise display or audio parameters. There was also no control of the equipment used. The code was written in JavaScript to run the task in a standard browser. It was impossible to run this code on a smartphone or tablet. A number of studies have shown that the display and response time capabilities in online Experiments can be comparable to lab-based testing (Bridges et al., 2020; Chetverikov & Upravitelev, 2016; Thornton, Tagu, et al., 2021), at least for a task similar to the one tested here. The code checked whether the stimuli were presented in full-screen mode and that the effective frame rate of the display was approximately 60 Hz. Participants were allowed to use trackpads to move the cursor, but their "clicks" on a target were registered only after the participants pressed the space bar. In the current Experiments, participants were required to have working headphones or speakers to hear the sound accompanying the change in the target stimulus. A sound discrimination test was run before participation to check whether the sound on a participant's computer was working properly.

Experiment 2 was run in the laboratory to replicate and check the result patterns observed in Experiment 1. An ASUS Nvidia G-Sync monitor (model PG279Q) was used in the laboratory experiment. The screen dimensions were 60 x 34 cm, with an effective 1920 x 1080 pixels resolution. The monitor was connected to a Windows 10 computer running Google Chrome (Version 108.0.5359.124 (Official Build) (64-bit)), and the monitor refresh rate was fixed at 60 Hz. Participants were seated approximately 60 cm from the screen, with an MI-2500X optical mouse from Trust and a K120 Logitech keyboard in front of them. The sound was played through the default monitor speakers. Experiment 2 used almost the same code, accommodating a within-subject, blocked design with increased repetitions.

### 3.2.3 Paper II. Stimuli and procedure

Eighty stimuli were presented on a grey background (size 800×600 px). In each trial, there were 20 vertical targets among 60 distractors. Half of the distractors were oriented 45° to the left, and the other half were oriented 45 ° to the right. All stimuli were randomly distributed in a non-visible 10 × 8 grid. Stimuli size was 5 × 40 px. An example of a trial is displayed in Figure 6.

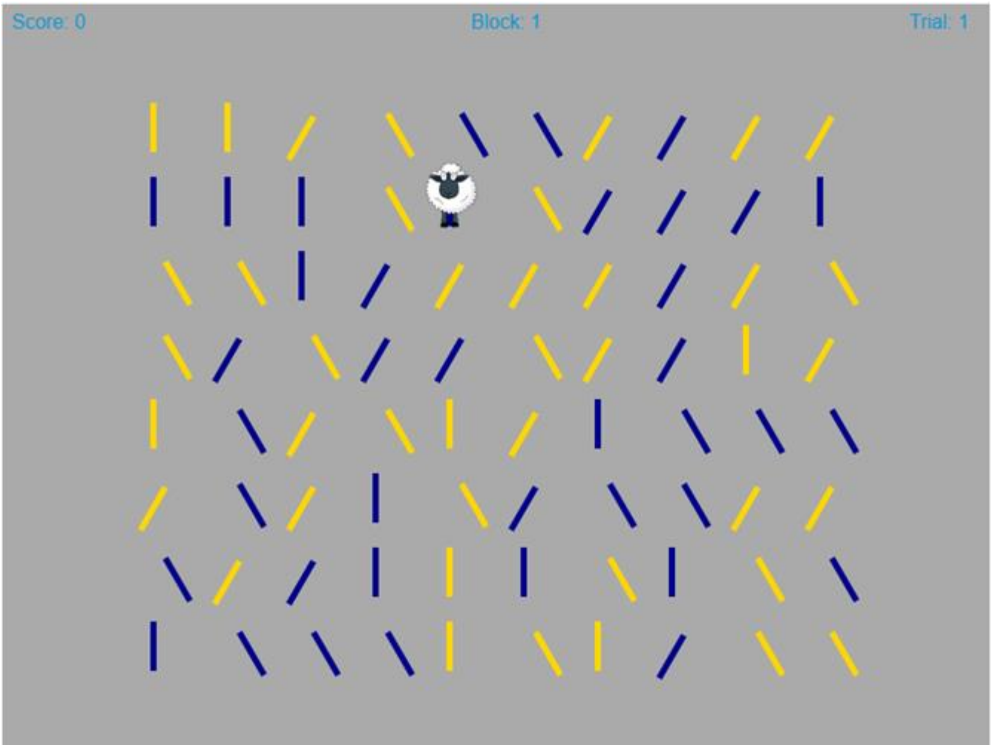

**Figure 6.** Example of one trial in Experiments 1 and 2 of Paper II. Eighty stimuli were presented on each trial (20 vertical targets and 60 distractors). Participants' task was to move the sheep (which served as a cursor) to the vertical lines and then press the spacebar to select targets. The number of targets already selected was shown next to the text "score" on the screen. Block number was denoted in the middle top position of the screen (1 block out of 4). The number of the current trial was presented in the right top corner. A demo can be accessed at https://maltacogsci.org/synchsearchdemo/.

Experiment 1 was run online. The experimental design included one between-subject factor and one within-subject factor. The within-subject factor was called synchrony type. It included two levels: vision only (trials where targets changed only their colour in synchrony) and vision + sound (trials where their synchrony colour change was accomplished with non-spatial sound). The between-subject factor was called synchrony presence and had two levels: synchronous trials (trials where targets changed their colour in synchrony with each other or accomplished with non-spatial sound) and asynchronous trials (trials where the targets were changing without synchrony). Four blocks were presented in Experiment 1: two training blocks (5 synchronous trials and 5 asynchronous trials) and two main blocks (20 alternated synchronous and asynchronous trials in each block). Participants completing the vision + sound condition had an audio pre-test, where their task was to report how many beeps (odd or even numbers) were played. This pre-test was made to check whether they had working speakers.

In the second Experiment, participants came to the laboratory in a dark room for two consecutive days to reduce the amount of time they spent on the experiment. There were six sessions in total (three per day). Each session had three blocks of trials for each experimental condition: asynchronous, vision only (synchronous), and vision + sound (synchronous). The blocks were presented in a random order, but the trial type was constant within a block. There were 3 practice trials and 12 experimental trials within each block, taking approximately 10 minutes per block to complete. There were 270 trials (6 sessions x 3 blocks x 15 trials).

The experimental procedure was similar for the two Experiments. In asynchronous trials, half of the stimuli were coloured blue and half yellow. Each stimulus changed its colour at a randomly determined time from the gap between 180ms and 500ms, independently from each other. This value was randomly assigned. In synchronous trials, target stimuli had the same colour at the beginning of each trial and changed colour synchronously. In contrast, distractors changed colour in the same way as in the asynchronous trials. The colour of the vision + sound condition targets changed with a 12 ms auditory beep (approximately 1930 Hz).

The participant's task was to collect 10 vertically oriented stimuli as fast as possible. After each completed trial, feedback with the total time taken was provided. The trial was abandoned and repeated in all three experiments if a distractor was selected. Participants could rest between the trials.

It is important to mention that participants in Experiments 1 and 3 used different screens because they were run online. Hence, stimuli cannot be described in visual angle values. In Experiment 2, participants sat in the laboratory with a constant viewing distance (60 cm) and a standard monitor (1920 x 1080). The size of the grey background was 800 × 600 px (24 x 18° visual angle). The size of the item was 5 × 40 px (0.2 x 1.2° visual angle). The size of the sheep, which was used as a cursor, was 50 × 60 px (1.4 x 1.6° visual angle.)

Figure 7 shows an example of the display for Experiment 3. The main characteristics, time difference, and experimental design were similar to those in Experiments 1 and 2. There were differences in the task and the targets. The task was to find as many targets as possible within a 10-second limit. Targets were also different: for 8 participants, targets were oriented 45° clockwise, while for the other 8 participants, targets were oriented 45° counterclockwise. These targets were randomly placed between 60 distractors randomly and independently assigned on each trial. Distractors had random orientation in the space, but there were two constraints: no closer than 30° to vertical and no closer than 30° to target orientation.

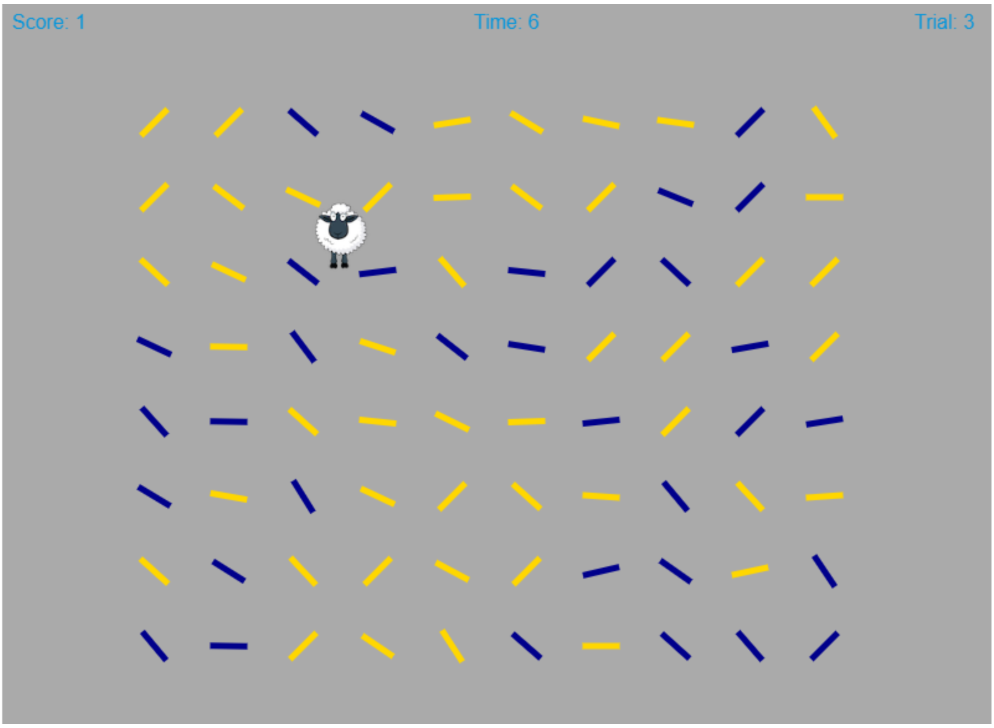

**Figure 7.** Example of one trial in Experiment 3 of Paper II. Eighty stimuli were presented on each trial (20 targets and 60 distractors). The participants' task was to move the sheep (which served as a cursor) to the targets (the 45° lines) and press the spacebar to select targets. The number of targets which were already selected were reported next to the score. The time left before the trial ended was reported in the middle top position of the screen (here 6 seconds left out of 10). The number of completed trials was presented in the right top corner.

### 3.2.4 Paper II. Data analysis

Dependent variables were similar for the first two experiments. In the case of the task in Experiment 3, the average number of targets found per condition was added to the analysis. The main speed-related variables for all Experiments were: the time to find the first target (RT1); the intercept of the ITT (inter-target times) function (an estimate of per-item speed during the so-called cruise phase of foraging, computed across the last 9 target items); the slope of the ITT function (a measure of foraging efficiency during the cruise phase, computed across the last 9 target items) and; the total completion time (TCT) for the trial (an estimate of overall speed during the cruise phase of foraging, computed by summing the ITTs for the last 9 target items). Another dependent variable was the error rate (number of errors made per condition). Note that in the current study, the foraging differs from many other studies (Á. Kristjánsson et al., 2014). Participants always search for the same category and do not switch between them, so assessing "run" behaviour is impossible. We estimated how organized the foraging was using the best-r parameter (Woods et al., 2013). This measurement can help to distinguish between global segregation and local grouping influences of synchrony. To compute the best-r, we initially assessed the correlation between each target position and serial selection position along the x-axis and y-axis individually in each trial. Subsequently, we identified the larger coefficient of the two correlations, regardless of the correlation's sign. Lastly, these best-r coefficients were averaged across trials for each participant, resulting in a singular measure for each condition. A low best-r coefficient means disorganized foraging, while a high coefficient indicates organized foraging (Jóhannesson et al., 2016).

Experiment 1 analysis was performed on all dependent measures separately using a 2 (synchrony presence: synchronous/asynchronous) x 2 (synchrony type: vision only/vision + sound) mixed ANOVA model. In experiments 2 and 3, all dependent measurements were analysed using separate one-way repeated measures ANOVA with the factor condition, having three levels: asynchronous, vision only (synchronous), and vision + sound (synchronous). Pairwise t-tests with Bonferroni corrections were used to analyse the difference between these conditions, where this test was applicable.

Bayesian analyses (Masson, 2011) were used to check the tests where the null hypothesis was not rejected. We reported BF01 and p(H0|D) for the cases where weak evidence was obtained – specifically BF01 < 3.0 and p(H0|D) < 0.75 (Jeffreys, 1998; Raftery, 1995; Wagenmakers et al., 2011).

Data analysis was performed in RStudio using the rstatix, ez, and tidyverse packages (Kassambara A., 2021; Lawrence MA, 2016; RStudio Team, 2020; Wickham et al., 2019). Analysis scripts and raw data files have also been made available on the OSF page associated with this paper at osf.io/xu5rd/.

## 3.3 Paper III: Cross-modal cues improve the detection of synchronized targets during human foraging

Two experiments were performed in this study. Experimental procedures conformed to the declaration of Helsinki. All participants confirmed they had read the informed consent and understood the task. Participants signed the hard copy of the informed consent form for both Experiments.

### 3.3.1 Paper III. Participants

Twenty-four people participated in both experiments (18 females, Mean age = 26.7), and there were 12 in each experiment. People participated voluntarily and were recruited via convenience sampling from the wider University of Iceland community. All participants provided written, informed consent forms and knew they could end participation at any time.

### 3.3.2 Paper III. Apparatus

The current study was conducted in the same laboratory as Experiment 2 in paper II. See section "Paper II. Apparatus" for monitor parameters, Google Chrome version, computer system, optical mouse model, and keyboard model. Both experiments were programmed in JavaScript. The default monitor or computer speakers were used to present the sound. In the condition where vibrations were used as a haptic cue, we used a custom-made bracelet which consisted of a Lofelt L5 voice-coil actuator (same as in the Paper I) that responds to audio signals, that was placed in 3D-printed plastic casing, placed flat up against the skin (Ævarsson et al., 2022; Yeganeh et al., 2023). We set up this condition with an external music card (Logilink 7.1 channel USB sound box). During this condition, participants wore headphones playing auditory white noise to block the sound coming from the actuator.

### 3.3.3 Paper III. Stimuli and procedure

Thirty-six stimuli were presented on each trial in a 6 x 6 grid on a grey background. The diameter of each black circle was 1.45°. A small dot (diameter approximately 0.1°) was placed on each circle to show its movement. One step of the dot was equal to 30°, starting from a random location. The dots could move in both directions. There was 1 template (upper left corner), 23 distractors, and 12 targets. On the targets, the position of the dots rotated in synchrony (at the same rate). The remaining 23 items came from the two distractor sets, moving faster or slower than the targets by a fixed offset of 167 ms (10 refresh cycles).

The participants' task was to collect 6 targets as fast as possible. There were two conditions in Experiment 1. The first condition was called the "No cue" condition. Participants searched for 6 targets relying only on the template. The second condition was called "Sound cue." Every movement of the targets was accompanied by an auditory beep (same as in Paper II, 1930 Hz). Two additional conditions were added in Experiment 2. The first two conditions were the same as in Experiment 1, the "No cue" and "Sound cue" conditions. The third condition in Experiment 2 was the "Visual cue" condition. The target movement was synchronized with the colour change of the frame from white to black. The last condition was called the "Haptic cue" condition. The target movement was synchronized with a haptic pulse lasting 12 ms. Its frequency was equal to 150 Hz, close to the optimum frequency for vibrotactile perception (Ævarsson et al., 2022).

The experiment began after participants received instructions via consent form, their questions, if any, were answered and they signed consent forms. If participants clicked on the target, the dot on the circle disappeared, and the "Score" rate increased. If they clicked on a distractor, the dot on the circle remained visible, and the "Errors" counter at the top-right of the screen increased. The trial ended after the successful selection of 6 targets. At the end of each trial, participants received feedback on the number of errors made and total completion time.

An example of one trial can be found on Figure 8.

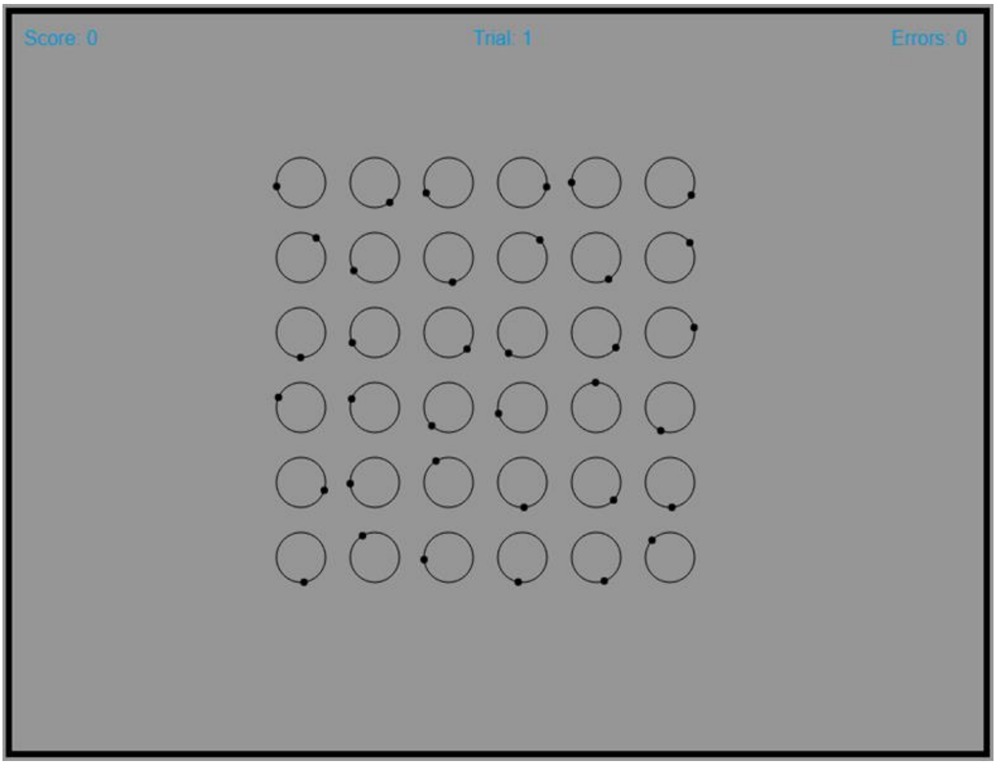

**Figure 8.** Example of one trial in Experiments 1 and 2 of Paper III. The small black dots changed position around the circumference of their hosting circles at a specific rate. There were three groups of 12 objects that moved in synchrony at a fast, medium or slow rate. The 12 target dots were always defined as the dots moving at the medium speed (an example target was always present at top left). See text for further details and https://maltacogsci.org/circleSynch/DEMO/ for a dynamic demonstration of the display and task.

Both experiments were organized into blocks equal to the number of conditions in each experiment (two in Experiment 1 and four in Experiment 2). The presentation of the blocks was randomized. Experiment 1 took approximately 25 minutes to complete (2 blocks x (2 practice + 10 main trials)). Experiment 2 took approximately 55 minutes to complete (4 blocks x (2 practice + 10 main trials)).

### 3.3.4 Paper III. Data analysis

The following dependent variables were analysed in the current paper: first target selection time (RT1), the slope of the RT x target number function, and the average number of errors. The main purpose of the first variable was to estimate response properties to later selection (Á. Kristjánsson et al., 2014, 2020; Ólafsdóttir et al., 2016; Thornton & Horowitz, 2004, 2020) and to compare the results directly with Paper II. The slope function characterises search patterns in the trial. Error analysis provides insights into the speed/accuracy tradeoffs in the search.

In both experiments, repeated measures ANOVA was chosen for the analyses of slope functions. Paired t-tests were used in Experiment 1 to compare error rates and RT1, and a One-way Repeated Measures ANOVA was used to analyse error rates in Experiment 2. Violations of sphericity were corrected using Greenhouse-Geisser adjustments to the degrees of freedom, and we used Holm corrections for multiple comparisons.

As in all previous studies, data analysis was performed in Rstudio, using the rstatix, ez, and tideverse packages (Kassambara A., 2021; Lawrence MA, 2016; RStudio Team, 2020; Wickham et al., 2019). Analysis scripts and raw data files have been made available on the OSF page associated with this paper at OSF page (https://osf.io/yrmp8/).

# 4 Results

As described in the "Data analysis" sections, many dependent variables were assessed. Hence, many statistical tests were conducted to answer research questions. The detailed results can be found in the papers in the appendix. Some extra analyses and raw data can be found on the OSF projects' pages. In the current section, only the main results of the papers will be discussed.

## 4.1 Paper I: The Haptic Intensity Order Illusion Is Caused by Amplitude Changes

In this paper, our main goal was to investigate whether the IOI is caused by amplitude or frequency changes. All the participants' responses were recalculated as a proportion of "up" responses. All possible frequency and amplitude value combinations were analysed and compared using three-way Repeated Measures ANOVA.

Our results provide clear evidence that amplitude changes are the primary cause of the IOI. This is evident from the descriptive Figure 9, which illustrates the role of frequency in Experiment 1 of the first paper. Furthermore, our statistical analysis confirms that the significant difference in IOI appeared only in the groups where amplitude was manipulated. These findings contribute to our understanding of the IOI phenomenon and have potential implications for future research in this area.

Secondly, the Lofelt L5 actuators were successfully tested and used in the current research. After this positive result, our group used them several times including Paper III that is presented in the current work.

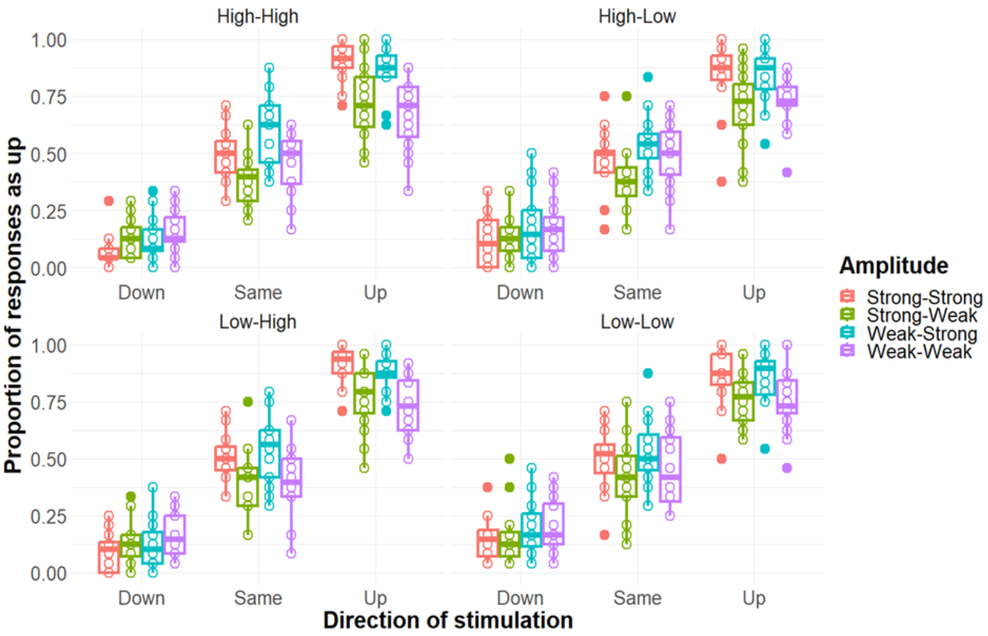

**Figure 9.** Results of Experiment 1 in Paper I. The proportion of "up" responses as a function of amplitude differences between the first and second stimulation (strong-strong, strong-weak, weak-strong or weak-weak). The data are separated by the direction of stimulation (whether the second stimulation was up, down or in the same location, relative to the first stimulation) and by the combinations of frequencies of the first and second stimulation (high-high, high-low, low-high or high-low). The horizontal lines in the box plots represent medians. The top horizontal line of the boxplot represents the upper quartile (75%), the bottom horizontal line represents the lower quartile (25%). The top vertical line going from the boxplot represent the maximum value, the bottom line represents the minimum value. The filled dots represent outliers. The hollow dots represent the means for individual participants.

Another result shows that no IOI appears for differential stimulation in the horizontal direction, as there was seen for differential vertical stimulation. Current results imply that the IOI may be linked to effects in which specific types of stimulation have a directional component, where, in this case, amplitude increases are potentially connected with the "up" direction.

Finally, our results show how the IOI was replicated using different actuators than in the original report by Hoffmann et al. (2019). Nevertheless, the IOI appears much stronger when the second stimulation is higher than the first stimulation. The IOI effect tends to be very weak when the second stimulation is lower than the first stimulation. Current results are less clear than they were in the previous work of our research group in Hoffman et al. (2018). However, the main effect of the IOI when the same actuator activity follows the first stimulation tends to be strong. Overall, we mostly replicated our group's previous findings.

## 4.2 Paper II: The effects of visual and auditory synchrony on human foraging

The main goal of the research in Paper II was to understand whether attention can be driven by visual synchrony and whether this could be reflected in faster foraging performance. Additionally, we studied the multimodal effects of the combination of visual and auditory synchrony. Three experiments were conducted to understand these topics.

In Experiment 1, our primary inquiry involved whether foraging behavior would demonstrate variability based on the synchronicity of visual and auditory stimuli. Our findings provided compelling evidence that the presence of synchrony does indeed impacts foraging. Specifically, we observed reductions in both the time taken to locate the first target (RT1) and the duration of the search time of the trials after first target was found (TCT), alongside an increase in the level of search organization (best-r) when target colour changes were synchronous compared to asynchronous changes. Analysis of the inter-target interval (ITT) pattern during the cruise phase indicated that synchrony offers relatively modest benefits through local grouping rather than altering target detection via global segregation. In the Figure 10 results for the TCT can be found.

Despite expectations based on the "pip-and-pop" effect (Van der Burg et al., 2008) found for single target visual search, adding a cross-modal auditory cue to target synchrony did not enhance performance in the current Experiment. However, the between-subject design used and the inherent variability in auditory environments when conducting the study online may have limited our ability to detect subtle cross-modal advantages, particularly given the presence of a robust visual synchrony cue. Null effects associated with synchrony type exhibited weak supporting evidence in our Bayesian analyses, along with trends indicating overall speed differences between participant groups. To address these uncertainties, we opted for replicating Experiment 1 under more controlled laboratory conditions, utilizing a within-subject design, before drawing definitive conclusions regarding any beneficial of cross-modal synchrony on foraging.

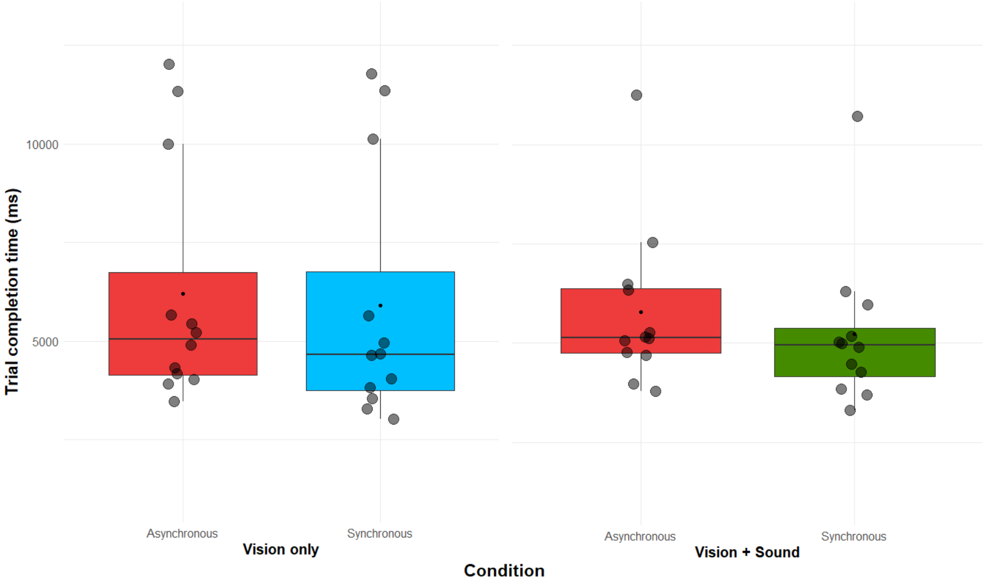

**Figure 10.** Results of the trial completion time (TCT) duting the cruise faze in Experiment 1 in Paper II. The horizontal lines in the box plots represent medians. The top horizontal line of the boxplot represents the upper quartile (75%), the bottom horizontal lines represent the lower quartile (25%). The top vertical line represents the maximum value, the bottom line represents the minimum value. The dots represent individual participants' median TCT's.

Therefore, Experiment 2 was run in the laboratory conditions to verify the previous results and collect more data. We also made some adjustments that could potentially enhance the detection of any cross-modal synchrony benefits if present. Transitioning to a within-subject design and conducting the Experiment under controlled laboratory conditions with increased trial repetitions notably influenced overall performance levels. Participants exhibited quicker response times and reduced variability across all dependent variables compared to Experiment 1, except for the cruise-phase slopes. Importantly, we successfully replicated the synchronous advantage observed in Experiment 1, wherein performance on synchronous trials surpassed that of asynchronous trials. The results fro TCT can be found in Figure 11. Distinctly positive cruise-phase slopes and spatially organized search patterns further strengthen the notion that synchrony enhances local grouping of targets rather than inducing global segregation, where they would conspicuously "pop out." This absence of evident global segregation remained consistent across both the sound-only and sound + vision conditions of Experiment 2. Despite expectations regarding cross-modal synchrony, there was still no indication of any supplementary advantage from including sound. This replication of patterns observed in Experiment 1 alleviates our previous concerns regarding the methodological limitations of the online, between-subject design employed in Experiment 1.

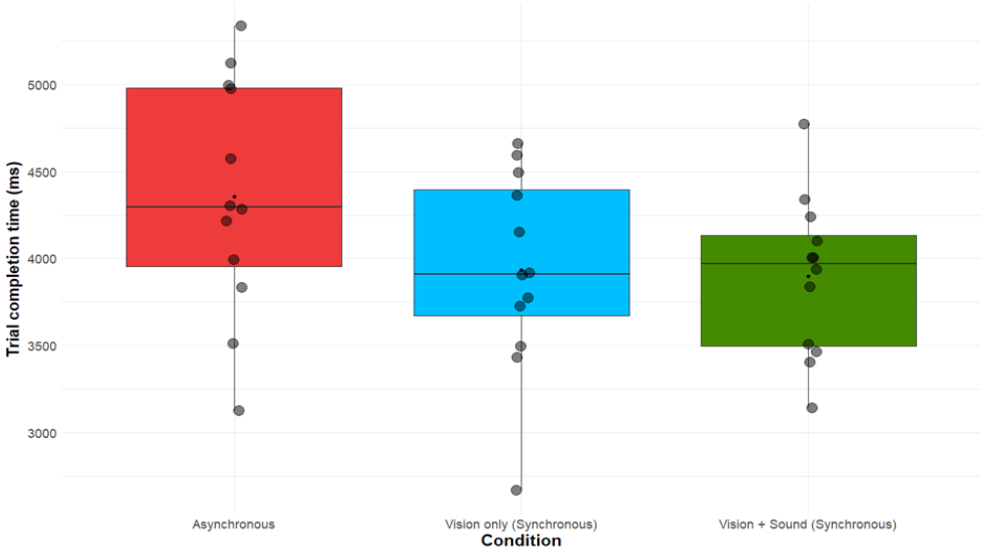

**Figure 11.** Results of the trial completion time (TCT) duting the cruise faze in Experiment 2 in Paper II. The horizontal lines in the box plots represent medians. The top horizontal line of the boxplot represents the upper quartile (75%), the bottom horizontal lines represent the lower quartile (25%). The top vertical line represents the maximum value, the bottom line represents the minimum value. The dots represent individual participants' median TCT's.

Based on the results of the Experiment 2, we tried to increase the difficulty of the task. Our idea was that increased difficulty would push participants towards using cross-modal cues.

The results of Experiment 3 showed that our manipulation to increase the difficulty of the search was successful. Time spent detecting the first target (RT1) and general time patterns of the foraging behavior increased compared to Experiment 1 (as in experiment 1, this was an online study). But even with this increased difficulty, the number of targets detected within 10 seconds was equal for the visual and visual + sound conditions. Results for the number of targets selected in one trial can can be found in Figure 12. These results appeared in all three studies, showing that visual synchrony provides advantages in foraging, while no advantage was observed from the non-spatial sound. The results of Experiment 3 also show that the search time patterns for the RT1 differ from the results of Experiment 1. There was no difference between conditions for RT1. Participants could have used different strategies compared to the previous experiments in this paper. Speculatively, participants tried searching for as many targets as possible and then clicked on them.

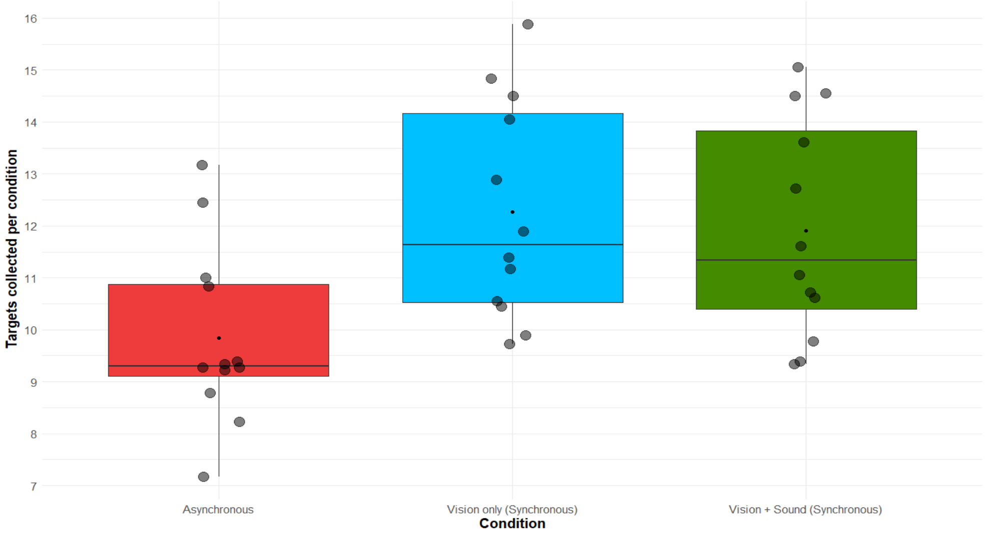

**Figure 12.** Results for the average number of targets collected per condition in Experiment 3 in Paper II. The horizontal lines in the box plots represent medians. The top horizontal lines of the boxplots represent the upper quartiles (75%), the bottom horizontal lines represent the lower quartiles (25%). The top vertical lines represent the maximum value, the bottom lines represent the minimum value. The dots represent individual participants' median results.

To sum up, the results of the current paper were surprising for us, in particular that there were no advantages from the auditory cue, and to clarify some questions our group decided to conduct the studies presented in Paper III.

## 4.3 Paper III: Cross-modal cues improve the detection of synchronized targets during human foraging

In the previous study (paper II), our group showed that visual synchrony can aid foraging performance, even though this synchrony was irrelevant to the task. The results showed that visual synchrony aids the performance, while additional synchrony cueing from sound gave no benefits. One of the possible explanations for the previous results was that the task was too easy to complete, so participants did not need to rely on multimodal cueing.

In the current research, our group conducted two experiments, where Experiment 2 replicated and extended the results of the Experiment 1.

In the first experiment, our group compared the "Sound cue" to the "No cue" condition. In the "Sound cue" condition participants searched for the target stimuli that were moving synchronously with a non-spatial sound. In "No cue" condition participants searched for the target stimuli without any hints. The results showed that participants significantly benefit from the cross-modal cueing. Search time in trials with non-spatial sound was significantly lower than in the trials with no hints.

In the second experiment, we compared cues from the different modalities ("Visual cue", "Haptic cue", "Sound cue") with the baseline ("No cue") condition. "Sound cue" and "No cue" conditions were same as in Experiment 1. In the "Haptic cue" condition the target movement was accomplished with a vibration to the wrist. This vibration did not provide any information about the position of the targets. In the "Visual cue" condition the frame of the search field changed the colour from white to black in synchrony with the target's movement. The results showed that cross-modal cues ("Sound cue" and "Haptic cue") can aid foraging when the targets defined by synchrony. In addition, cross-modal cueing gave a far larger advantage than within-modal cueing. Interestingly, visual cueing did not differ from the baseline level. It did not aid visual foraging. Results for Experiment 2 can be found in Figure 13.

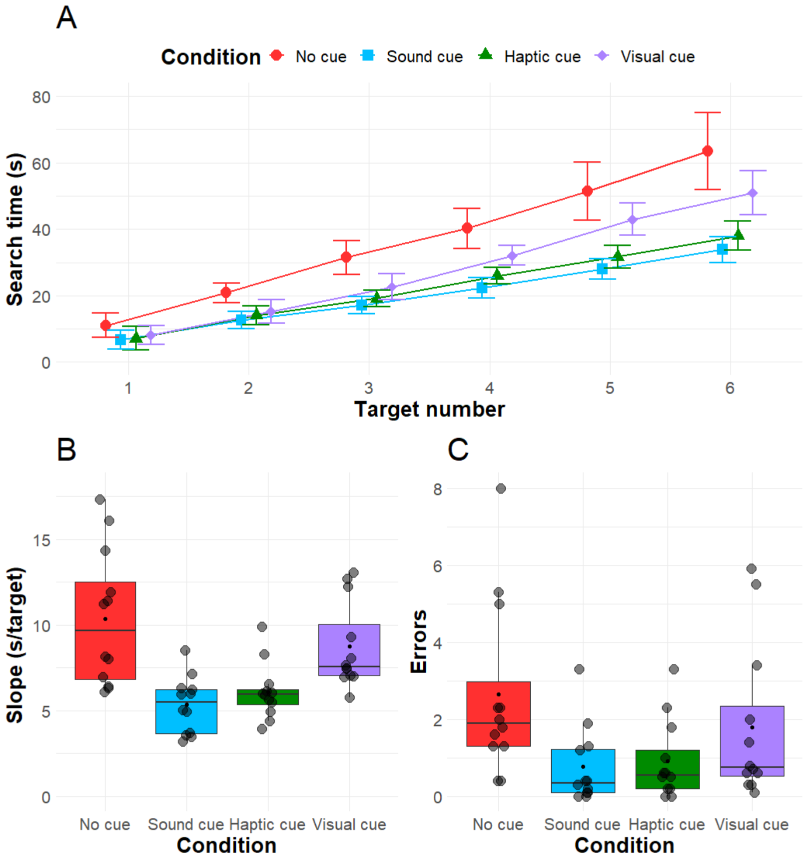

**Figure 13.** Results for several variables for Experiment 2 in Paper III. A – Search time as a function of target number in Experiment 2. Error bars represent 95% confidence intervals. B-C – The slope of the RT x target number function and the error rates in Experiment 2. The horizontal lines in each box plots represent the median. The top horizontal line represents the upper quartile (75%) while the bottom horizontal line represents the lower quartile (25%). The top vertical lines represent the maximum value, while the bottom lines represent minimum values. The large dots represent individual mean slope values. The small dot shows the overall mean.

These results differed from what was observed in all Experiments in Paper II, where visual synchrony as cue, was enough to aid foraging performance. A possible explanation for this difference is that visual synchrony (same rotation rates) identified the targets in Paper III. Rotation rates were relevant to the task, while in Paper II, the targets were identified through the spatial orientation of the stimuli. In paper II visual synchrony was used as a cue. This change in the feature defining the targets increased the benefits of cross-modal synchrony cues.

Another possible explanation, that the task in Paper III was more complicated than in Paper II (Experiment 3). Hence, participants relied more on cross-modal cues.

# 5 Discussion

Our brains create perceptual multimodal representations of the surrounding world by combining information from different senses. The current model of perception works for people with fully functional perceptual systems. However, in situations where individuals are born without a functioning cochlea, the essential sensory organ of the auditory system, or with an eye disconnected from the brain, how does this impact the development of perceptual representations? Here, the sensory substitution plays a role. Sometimes, if we lose our eyes, that does not mean that we have lost our ability to see (Bach-y-Rita & W. Kercel, 2003; Á. Kristjánsson et al., 2016).

The general theme behind the studies presented in this thesis is the investigation of multimodal attention, with the aim of leveraging this understanding to design sensory substitution devices (SSD). Additionally, our research explores the most effective methods of presenting information through SSDs, taking into account the limitations of human perception.

In Paper I, our group aimed to understand what factors cause the intensity order illusion (IOI). Another important goal was to test the actuators. Our group planned to use these actuators in the design of the SSD. These actuators were also used in Paper III.

In Paper II, our group studied the effects of visual synchrony on multiple target foraging. Another important topic our group studied was the mechanisms underlying the search improvement. Finally, our group was trying to find how additional cross-modal cues affect performance.

In Paper III, our group investigated whether cross-modal cues aid foraging more effectively than unimodal cues. Another research question was dedicated to the effect of visual synchrony on multiple-target foraging.

In what follows, I am going to discuss main results of these papers. The first part, Tactile Perception and Sensory Substitution Devices,' (section 5.1) covers Paper I and Paper III discussions and highlights how tactile perception can affect sensory substitution devices. The second part, 'Multimodal attention and Sensory Substitution Devices,' (section 5.2) discusses Paper II and Paper III and highlights the potential of using multimodal attention in developing more effective SSDs.

## 5.1 Tactile attention and sensory substitution devices

For the development of high-quality sensory substitution devices, you need to consider many things such as: the most effective way to perceive information through substituting modality, the amount of information that can be sent to the substituting modality so that it does not become overloaded, and many more. One of the possible problems that you may face involves illusions. Tactile illusions involve the interpretation of the object through the touch that does not directly reflect the stimulation and can therefore provide clues about how the nervous system interprets tactile information. There are several examples of tactile illusions where people underestimate the distance between the stimuli (Carter et al., 2008; Lechelt & Borchert, 1977) or even geometrical illusions that have mostly been studied for the visual modality (Gentaz & Hatwell, 2004; Ziat et al., 2014).

Perceptual illusions, like those observed in sensory substitution, offer valuable insights into sensory processing. Hotting and Röder (2004) found that a single tactile stimulus accompanied by multiple auditory sounds can create the perception of multiple touches, akin to visual-auditory illusions described by Shams and his collegues (Shams et al., 2000). Notably, congenitally blind individuals were less susceptible to this effect, suggesting differences in multisensory integration. Other examples include the McGurk effect, where lip movements influence perceived speech, and the cutaneous rabbit illusion, generating illusory movement through rapid stimulation. While these phenomena pose challenges for sensory substitution, they also inform our understanding of haptic processing, crucial for improving such technologies. Further exploration of these illusions holds promise for enhancing sensory substitution devices and unravelling underlying perceptual mechanisms.

In Paper I of this thesis we studied the so-called intensity order illusion. The IOI, primarily driven by amplitude changes rather than frequency modulation, suggests that this illusion can be most likely caused by receptors sensing amplitude changes. These findings align with previous research highlighting the role of attention in tactile discrimination and localization tasks (Cholewiak et al., 2001; Cholewiak & McGrath, 2006).

Anisotropies, or direction dependencies, are observed in tactile perception, particularly for pressure stimuli, showing higher horizontal than vertical acuity across various body parts and tasks (Cholewiak, 1999). The Weber illusion further supports this, revealing that perceived distance between two points is larger on sensitive than insensitive skin (Weber, 1996). Recent studies using vibratory stimulation on the back confirm these findings. However, while tactile anisotropies have been extensively studied for pressure stimuli, results regarding vibrotactile stimulation are mixed, with some studies showing significant anisotropies (Sofia & Jones, 2013), while others do not (Van Erp, 2005). Hoffmann et al. reported higher accuracy for horizontal vibrotactile stimulation, suggesting lower precision in the vertical dimension, potentially increasing the likelihood of perceptual illusions (Hoffmann, Brinkhuis, et al., 2018).

Plaisier, Sap & Kappers (2012) found that vibrotactile stimulation using eccentric rotating mass actuators resulted in longer perceived distances for vertical than for horizontal stimulation. Moreover, distances were perceived as longer when vibration occurred bilaterally along the spine compared to unilateral stimulation, potentially due to hemifield effects in vision (Alvarez & Cavanagh, 2004; Luck, S. J. et al., 1989).

Our findings are in line with most of the previous studies and suggest that understanding the role of attention and of perceptual interpretation in tactile perception is essential for optimizing the effectiveness of sensory substitution devices, ensuring that they provide meaningful and actionable feedback to users.

Paper I also showed, that Lofelt 5 actuators provide us with good quality tactile stimuli that can be used for SSD's. The same actuators were also used in Paper III. This paper provides us with interesting insights about the way how we can combine tactile attention and visual attention.

Tactile attention is used a lot as a cue in the modern world. Mobile phones provide people with all kinds of feedback through such vibrations. The results of the research in Paper III shows that haptic cueing could be as effective as visual cueing. These results shows that tactile attention is a very important topic to study in the future and can be broadly used for Sensory Substitution Devices.

Another broad field that was covered in this thesis is the integration of different modalities to complete visual search tasks, which is one of the most common tasks in the real world. How do cues from different modalities affect performance on visual search tasks? How can this information be used in developing SSD's?

## 5.2 Multimodal attention and sensory substitution devices

Human attention is a limited resource, leading to selective filtering of sensory information (Á. Kristjánsson, 2006; Most et al., 2005; Neisser, 1967; Treisman, 1960; Wolfe, 2021). This impacts various sensory modalities; for instance, in vision, observers often miss significant changes that occur during eye movements (Grimes, 1996). Similar constraints apply to haptic perception (Gallace et al., 2006), emphasizing a universal principle of attention and perception. An illustrative example is evident in cockpit displays (Haines, 1991). Pilots focusing on instrument readings often miss vital information, such as obstacles during landing. This scenario has implications for sensory substitution, where prioritizing substitute sensory inputs, like auditory cues via headphones, may detract from attending to real-world sounds (Kärcher et al., 2012).

To address these challenges, sensory substitution devices must minimize demands on attentional resources to prevent sensory overload (Elli et al., 2014; Kristjánsson et al., 2016). Additionally, auditory substitution should complement rather than interfere with the natural auditory environment (Collins, 1985). Neglecting these processing constraints could limit the effectiveness of sensory substitution devices in practical scenarios (Loomis et al., 2018), highlighting the importance of understanding sensory and attentional mechanisms in their design.

In Paper II and Paper III, our group studied the effect of synchrony on the performance during multiple target foraging. In Paper II, synchrony was not a target feature, while in Paper III, participants had to search for synchrony. Our group expected that synchrony would enhance foraging, making the targets more easily discernible from the distractors. There were two main hypotheses as to why synchrony can lead to a "pop-out" effect: global segregation or the "local grouping" effect of the targets. Cross-modal cueing was another thing that was studied in these papers. In paper II, we expected additional benefits from audio cueing, which appeared synchronously with changes of target colour. In paper III, different types of cues and their effect on the foraging compared to the condition with no cues were tested.

In Paper II, results showed clear benefit from visual synchrony, while no additional benefit from cross-modal cueing was found. Those results were observed in all experiments in Paper II.

While our group anticipated synchronous items to pop out among distractors, our findings revealed a more nuanced picture. Contrary to expectations, no additional benefit appeared from non-spatial sound. There was no difference between synchronous trials with and without non-spatial sound for following dependent variables: RT1 (response times for the first selected target), ITT slopes, best-r, number of targets found in Experiment 3. This lack of synchrony "pop-out" aligns with previous visual search studies by Cavanagh et al. (2001), Li et al. (2014), and Thornton and Vuong et al. (2021), which reported inefficient search even with synchronous stimuli. My proposal is that synchrony may aid foraging performance through local grouping rather than global segregation, facilitating more organized search patterns.

Additionally, our group explored whether cross-modal synchrony via non-spatial sound could enhance performance, inspired by Van der Burg et al.'s (2008) "pip and pop" effect. However, our experiments did not replicate these benefits in a foraging task for multiple targets, suggesting a complex relationship between synchrony and task dynamics.

My proposal is that that synchrony aids performance through local target grouping, guiding attention to synchronized items once a target is selected. This mechanism, supported by our findings of organized search patterns, suggests a locally guided foraging process. Despite expectations based on previous studies, cross-modal synchrony via a non-spatial sound did not improve performance in our experiments. Unlike the "pip and pop" effect reported by Van der Burg et al. (2008), synchronous sound did not reduce completion times, challenging the idea that it enhances bottom-up processing. Our results suggest that while auditory and visual cues may bind for single items, this binding may be hindered by grouping, eliminating the expected benefits from auditory synchrony.

Our results align with those of Fujisaki et al. (2006), who demonstrated deteriorating target detection with increasing unsynchronized visual distractors, suggesting sequential processing in spatiotemporal synchrony benefits. Response times in the "pip & pop" paradigm increase with set-size, challenging the idea of auditory cue-induced pop-out. The integration of sight and sound can vary widely based on temporal and spatial factors (Arnold et al., 2005; Driver, 1996; Reisberg, 1978), making it sensitive to specific perceptual and attentional parameters (Talsma et al., 2010). Exploring subtle temporal differences may shed light on the presence of cross-modality benefits in future experiments.

In summary for Paper II, our experiments highlight limitations in the benefits of cross-modally synchronous presentation in multiple-target search scenarios. While this alone may not conclusively question the "pip-and-pop" effect, the absence of a cross-modal foraging advantage, especially in the time taken to find the first target, is clearly unexpected. Instead of solely varying target-distractor similarity, exploring other ways of manipulating task difficulty could provide valuable insights. If we assume that vision dominates in our tasks, reducing the overall reliability of visual input might potentially benefit from sound.

In the next study I focused on manipulating the cross-modal cueing and making our task more difficult, expecting that participants will rely more on the cross-modal cues.

In paper III, despite the synchrony benefits from the cues, foraging remained slow and effortful due to the dynamic nature of the target feature which was rotation rate. Unlike single-target searches for dynamic objects, the cues did not lead to any "global segregation" of targets from distractors, suggesting a serial search process where each item needs to be checked individually (Cavanagh et al., 2001; Rensink, 2000). The synchrony cues aid in target identification but do not seem to facilitate attentional shifts between items like in Paper II.

One question arises: why does the visual cue condition not offer the same benefit as the cross-modal conditions? This discrepancy suggests within-modality conflict, possibly due to resource competition (Navon & Gopher, 1979; Salvucci & Taatgen, 2008; Wickens, 2002). Additionally, the duration of individual cue events varied between modalities, which could influence cue efficiency. The key change from previous studies is that synchrony became a defining feature of target identity, amplifying the cross-modal advantage (Van der Burg et al., 2008, 2011).

Current findings also challenge the notion that cross-modal cueing benefits are exclusive to biological stimuli, as the observed benefits in displays had litttle obvious biological relevance. Further investigations will explore if these advantages extend to biologically relevant displays (Nguyen et al., 2021; Shen et al., 2023).

In summary, the current studies provide us with the insights about the effectiveness of the cross-modal cueing and how the information should be sent through the SSD using different modalities not to overload perception mechanisms. These findings were proved by the results that occurred in Paper III, showing that cross-modal cueing was more effective than unimodal cueing. More than that, the use of Lofelt 5 actuators for tactile stimulation shows promise for SSD development, emphasizing the importance of tactile and visual attention integration.

# 6 Conclusions

The results of all three papers in the current thesis highlight different aspects, that need to be considered when developing Sensory Substitution devices. This thesis is divided into two main parts: the fundamental research of multimodal attention and the development and testing of SSDs. The primary goal is to deepen our understanding of perceptual mechanisms and apply this knowledge to create more effective and usable SSDs.

The first part of the thesis focuses on fundamental research, beginning with the study of the Intensity Order Illusion (IOI) in tactile perception. This section aims to answer the question: What are the underlying mechanisms of the IOI? Specifically, it examines whether changes in amplitude or frequency of sequential stimulations cause the IOI. Advanced tactile actuators that allow independent manipulation of these parameters are used to isolate the contributing factors and assess the generalizability of this phenomenon. In Paper I the first aim was addressed. It was found that intensity order illusion is caused by amplitude changes, not frequency manipulations. Also, the previous results were partially replicated (Hoffmann, Brinkhuis, et al., 2018), using a different type of actuators from the original study (Lofelt 5). These actuators were used in our subsequent research and in our ongoing development of Sensory Substitution devices involving tactile stimulation (Yeganeh et al., 2022, 2023, 2024). The results of experiment 2 show that this effect is specific to the vertical dimension and does not occur with differential stimulation in the horizontal dimension. Moreover, our findings suggest a perceptual anisotropy: the intensity order illusion for vibrotactile stimulation is less pronounced when the second stimulus is below the first, compared to when it is presented in the same location or above it. The answer for the first research question is that the possible mechanisms underlying are connected to the mechanoreceptors which react to the amplitudes. This finding highlights the importance of amplitude in shaping tactile perceptual experiences.

The second part of the thesis investigates multimodal attention, particularly how visual and auditory synchrony influences attention and perceptual performance in foraging tasks. The key research question here is: How do visual and auditory synchrony influence attention and perceptual performance in foraging tasks? Inspired by the "pip & pop" phenomenon, this study examines whether synchrony can guide attention, improve search accuracy, and reduce search time. To address this question, we perfromed the experiments in the second paper. The findings consistently showed a significant advantage of visual synchrony across all three experiments. Surprisingly, contrary to the expectations based on the "pip-and-pop" effect (Van der Burg et al., 2008), synchronous auditory cues did not provide any additional benefits upon foraging performance, above visual synchrony. While the benefits of visual synchrony contribute to our understanding of spatial and temporal grouping effects in vision, the absence of benefits from auditory cues highlights significant limitations in exploiting synchrony for performance enhancement. The answer for the second research question is that visual and auditory synchrony influence attention and perceptual performance in foraging tasks by enhancing the saliency of targets and improving detection accuracy through multimodal integration.

Based on the research in Paper II, we designed the experiments presented in paper III. The goal was to explore the effects of cross-modal cues on the detection of synchronized targets in complex visual foraging tasks, addressing the question: What are the effects of cross-modal cues on the detection of synchronized targets in complex visual foraging tasks? This involves comparing the effectiveness of unimodal versus multimodal cueing. In summary, the findings demonstrate a significant advantage conferred by cross-modal audio and vibrotactile synchrony cues, whereas this advantage is absent for a non-spatial visual cue. These results underscore the critical role of cross-modal interactions in elucidating visual attention, temporal attention, and scene comprehension. Additionally, these findings have strong implications for the development and refinement of sensory substitution devices aimed at enhancing sensory perception and information processing. The answer for the third research question can be formulated in the following sentence: cross-modal cues significantly enhance the detection of synchronized targets in complex visual foraging tasks by providing additional sensory information that helps to highlight the relevant stimuli, thereby improving accuracy and speed of target identification.

The thesis primarily presents research led by the author, focusing on the aims and findings of these studies. A detailed description of collaborative research can be found in the introduction. Overall, this work contributes valuable insights into tactile perception mechanisms, sensory interaction, and practical applications for SSDs and multimodal interfaces.

All these results highlight the importance of studying tactile attention and multimodal attention from various angles, using different methods. Multimodal attention studies will provide lots of important insights about the developing and increasing the quality of Sensory Substitution devices.

# References

Abboud, S., Hanassy, S., Levy-Tzedek, S., Maidenbaum, S., & Amedi, A. (2014). EyeMusic: Introducing a "visual" colorful experience for the blind using auditory sensory substitution. *Restorative Neurology and Neuroscience*, *32*(2), 247–257. https://doi.org/10.3233/RNN-130338

Ævarsson, E. A., Ásgeirsdóttir, T., Pind, F., Kristjánsson, Á., & Unnthorsson, R. (2022). Vibrotactile Threshold Measurements at the Wrist Using Parallel Vibration Actuators. *ACM Transactions on Applied Perception*, *19*(3), 1–11. https://doi.org/10.1145/3529259

Alais, D., Blake, R., & Lee, S.-H. (1998). Visual features that vary together over time group together over space. *Nature Neuroscience*, *1*(2), 160–164. https://doi.org/10.1038/414

Alvarez, G. A., & Cavanagh, P. (2004). The Capacity of Visual Short-Term Memory is Set Both by Visual Information Load and by Number of Objects. *Psychological Science*, *15*(2), 106–111. https://doi.org/10.1111/j.0963-7214.2004.01502006.x

Amedi, A. (2002). Convergence of Visual and Tactile Shape Processing in the Human Lateral Occipital Complex. *Cerebral Cortex*, *12*(11), 1202–1212. https://doi.org/10.1093/cercor/12.11.1202

Amedi, A., Malach, R., Hendler, T., Peled, S., & Zohary, E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nature Neuroscience*, *4*(3), 324–330. https://doi.org/10.1038/85201

Arno, P., De Volder, A. G., Vanlierde, A., Wanet-Defalque, M.-C., Streel, E., Robert, A., Sanabria-Bohórquez, S., & Veraart, C. (2001). Occipital Activation by Pattern Recognition in the Early Blind Using Auditory Substitution for Vision. *NeuroImage*, *13*(4), 632–645. https://doi.org/10.1006/nimg.2000.0731

Arnold, D. H., Johnston, A., & Nishida, S. (2005). Timing sight and sound. *Vision Research*, *45*(10), 1275–1284. https://doi.org/10.1016/j.visres.2004.11.014

Bach-y-Rita, P., Collins, C. C., Saunders, S. A., White, B., & Scadden, L. (1969). Vision substitution by tactile image projection. *Nature*, *221*(5184), 963–964.

Bach-y-Rita, P., & W. Kercel, S. (2003). Sensory substitution and the human–machine interface. *Trends in Cognitive Sciences*, *7*(12), 541–546. https://doi.org/10.1016/j.tics.2003.10.013

Banissy, M. J., Jonas, C., & Cohen Kadosh, R. (2014). Synesthesia: An introduction. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01414

Bedny, M., Richardson, H., & Saxe, R. (2015). "Visual" Cortex Responds to Spoken Language in Blind Children. *The Journal of Neuroscience*, *35*(33), 11674–11681. https://doi.org/10.1523/JNEUROSCI.0634-15.2015

Bella-Fernández, M., Suero Suñé, M., & Gil-Gómez De Liaño, B. (2023). The time course of visual foraging in the lifespan: Spatial scanning, organization search, and target processing. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02345-8

Bizley, J. K., & Cohen, Y. E. (2013). The what, where and how of auditory-object perception. *Nature Reviews Neuroscience*, *14*(10), 693–707. https://doi.org/10.1038/nrn3565

Blake, R., & Lee, S.-H. (2005). The Role of Temporal Structure in Human Vision. *Behavioral and Cognitive Neuroscience Reviews*, *4*(1), 21–42. https://doi.org/10.1177/1534582305276839

Blankenburg, F., Ruff, C. C., Deichmann, R., Rees, G., & Driver, J. (2006). The Cutaneous Rabbit Illusion Affects Human Primary Sensory Cortex Somatotopically. *PLoS Biology*, *4*(3), e69. https://doi.org/10.1371/journal.pbio.0040069

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, e9414. https://doi.org/10.7717/peerj.9414

Broadbent, D. E. (1958). Effect of Noise on an "Intellectual" Task. *The Journal of the Acoustical Society of America*, *30*(9), 824–827. https://doi.org/10.1121/1.1909779

Brooks, J., Kristjansson, Á., & Unnthorsson, R. (2023). Sensory Substitution: Visual Information via Haptics. In *Neuromethods* (Holmes, N.P. (eds) Somatosensory Research Methods., Vol. 196). Humana, New York, NY. https://doi.org/10.1007/978-1-0716-3068-6_14

Bruns, P. (2019). The Ventriloquist Illusion as a Tool to Study Multisensory Processing: An Update. *Frontiers in Integrative Neuroscience*, *13*, 51. https://doi.org/10.3389/fnint.2019.00051

Calvert, G. A., Brammer, M. J., Bullmore, E. T., Campbell, R., Iversen, S. D., & David, A. S. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *NeuroReport*, *10*(12), 2619–2623.

Carter, O., Konkle, T., Wang, Q., Hayward, V., & Moore, C. (2008). Tactile Rivalry Demonstrated with an Ambiguous Apparent-Motion Quartet. *Current Biology*, *18*(14), 1050–1054. https://doi.org/10.1016/j.cub.2008.06.027

Cavanagh, P., Labianca, A. T., & Thornton, I. M. (2001). Attention-based visual routines: Sprites. *Cognition*, *80*(1–2), 47–60. https://doi.org/10.1016/S0010-0277(00)00153-0

Chebat, D.-R., Maidenbaum, S., & Amedi, A. (2015). Navigation Using Sensory Substitution in Real and Virtual Mazes. *PLOS ONE*, *10*(6), e0126307. https://doi.org/10.1371/journal.pone.0126307

Chetverikov, A., & Upravitelev, P. (2016). Online versus offline: The Web as a medium for response time data collection. *Behavior Research Methods*, *48*(3), 1086–1099. https://doi.org/10.3758/s13428-015-0632-x

Cholewiak, R. W. (1999). The Perception of Tactile Distance: Influences of Body Site, Space, and Time. *Perception*, *28*(7), 851–875. https://doi.org/10.1068/p2873

Cholewiak, R. W., Collins, A. A., & Brill, J. C. (2001). *Spatial Factors in Vibrotactile Pattern Perception*. 8.

Cholewiak, R. W., & McGrath, C. (2006). Vibrotactile Targeting in Multimodal Systems: Accuracy and Interaction. *2006 14th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 413–420. https://doi.org/10.1109/HAPTIC.2006.1627100

Collignon, O., Lassonde, M., Lepore, F., Bastien, D., & Veraart, C. (2006). Functional Cerebral Reorganization for Auditory Spatial Processing and Auditory Substitution of Vision in Early Blind Subjects. *Cerebral Cortex*, *17*(2), 457–465. https://doi.org/10.1093/cercor/bhj162

Collins, C. C. (1985). On mobility aids for the blind. In *Electronic spatial sensing for the blind* (pp. 35–64). Springer.

Day, R. H., Mitchell, P., & Stecher, E. J. (1990). The bourdon illusion occurs with straight-, right-angle-, and parallel-edge figures. *Perception & Psychophysics*, *48*(4), 375–381. https://doi.org/10.3758/BF03206690

Driver, J. (1996). Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature*, *381*(6577), 66–68. https://doi.org/10.1038/381066a0

Elli, G. V., Benetti, S., & Collignon, O. (2014). Is There a Future for Sensory Substitution Outside Academic Laboratories? *Multisensory Research*, *27*(5–6), 271–291. https://doi.org/10.1163/22134808-00002460

Evans, K. K., & Treisman, A. (2011). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, *10*(1), 6–6. https://doi.org/10.1167/10.1.6

Fahle, M. (1993). Figure–ground discrimination from temporal information. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *254*(1341), 199–203. https://doi.org/10.1098/rspb.1993.0146

Farid, H. (2002). Temporal synchrony in perceptual grouping: A critique. *Trends in Cognitive Sciences*, *6*(7), 284–288. https://doi.org/10.1016/S1364-6613(02)01927-7

Flach, R., & Haggard, P. (2006). The cutaneous rabbit revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 717–732. https://doi.org/10.1037/0096-1523.32.3.717

Fujisaki, W., Koene, A., Arnold, D., Johnston, A., & Nishida, S. (2006). Visual search for a target changing in synchrony with an auditory signal. *Proceedings of the Royal*

*Society B: Biological Sciences*, *273*(1588), 865–874. https://doi.org/10.1098/rspb.2005.3327

Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception & Psychophysics*, *68*(7), 1191–1203. https://doi.org/10.3758/BF03193720

Gallace, A., Tan, H. Z., & Spence, C. (2006). The failure to detect tactile change: A tactile analogue of visual change blindness. *Psychonomic Bulletin & Review*, *13*(2), 300–303. https://doi.org/10.3758/BF03193847

Gao, M., Chang, R., Wang, A., Zhang, M., Cheng, Z., Li, Q., & Tang, X. (2021). Which can explain the pip-and-pop effect during a visual search: Multisensory integration or the oddball effect? *Journal of Experimental Psychology: Human Perception and Performance*, *47*(5), 689–703. https://doi.org/10.1037/xhp0000905

Gardner, E. P., & Martin, J. H. (2013). Coding of sensory information. In *Principles of neural science. In: Kandel, E.R., Schwartz, J.H., Jessell, T.M.: Vol. 574 (eds)* (5th ed.).

Gardner, E. P., & Spencer, W. A. (1972). Sensory funneling. I. Psychophysical observations of human subjects and responses of cutaneous mechanoreceptive afferents in the cat to patterned skin stimuli. *Journal of Neurophysiology*, *35*(6), 925–953. https://doi.org/10.1152/jn.1972.35.6.925

Geldard, F. A., & Sherrick, C. E. (1972). The Cutaneous 'Rabbit': A Perceptual Illusion. *Science, New Series*, *178*(4057), 178–179.

Gentaz, E., & Hatwell, Y. (2004). Geometrical haptic illusions: The role of exploration in the Müller-Lyer, vertical-horizontal, and Delboeuf illusions. *Psychonomic Bulletin & Review*, *11*(1), 31–40. https://doi.org/10.3758/BF03206457

Gil-Gómez De Liaño, B., & Wolfe, J. M. (2022). The FORAGEKID Game: Hybrid-Foraging as a new way to study aspects of executive function in development. *Cognitive Development*, *64*, 101233. https://doi.org/10.1016/j.cogdev.2022.101233

Grimes, J. (1996). *On the failure to detect changes in scenes across saccades.* (K. Akins (Ed.), Vancouver studies in cognitive science: 5., pp. 89–109). Oxford University Press.

Haines, R. F. (1991). A Breakdown in Simultaneous Information Processing. In G. Obrecht & L. W. Stark (Eds.), *Presbyopia Research* (pp. 171–175). Springer US. https://doi.org/10.1007/978-1-4757-2131-7_17

Hall, U. (1999). *Wavelets, vision and the statistics of natural scenes*.

Härmä, A., Jakka, J., Tikander, M., Karjalainen, M., Lokki, T., Hiipakka, J., & Lorho, G. (2004). Augmented reality audio for mobile and wearable appliances. *Journal of the Audio Engineering Society*, *52*(6), 618–639.

Harrar, V., & Harris, L. R. (2005). Simultaneity constancy: Detecting events with touch and vision. *Experimental Brain Research*, *166*(3–4), 465–473. https://doi.org/10.1007/s00221-005-2386-7

Harrar, V., & Harris, L. R. (2008). The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Experimental Brain Research*, *186*(4), 517–524. https://doi.org/10.1007/s00221-007-1253-0

Hoffmann, R., Brinkhuis, M. A. B., Kristjánsson, Á., & Unnthorsson, R. (2018). Introducing a New Haptic Illusion to Increase the Perceived Resolution of Tactile Displays: *Proceedings of the 2nd International Conference on Computer-Human Interaction Research and Applications*, 45–53. https://doi.org/10.5220/0006899700450053

Hoffmann, R., Brinkhuis, M. A. B., Unnthorsson, R., & Kristjánsson, Á. (2019). The intensity order illusion: Temporal order of different vibrotactile intensity causes systematic localization errors. *Journal of Neurophysiology*, *122*(4), 1810–1820. https://doi.org/10.1152/jn.00125.2019

Hoffmann, R., Spagnol, S., Kristjánsson, Á., & Unnthorsson, R. (2018). Evaluation of an Audio-haptic Sensory Substitution Device for Enhancing Spatial Awareness for the Visually Impaired. *Optometry and Vision Science*, *95*(9), 757–765. https://doi.org/10.1097/OPX.0000000000001284

Ifukube, T., Sasaki, T., & Peng, C. (1991). A blind mobility aid modeled after echolocation of bats. *IEEE Transactions on Biomedical Engineering*, *38*(5), 461–465. https://doi.org/10.1109/10.81565

Jarocka, E., Pruszynski, J. A., & Johansson, R. S. (2021). Human Touch Receptors Are Sensitive to Spatial Details on the Scale of Single Fingerprint Ridges. *The Journal of Neuroscience*, *41*(16), 3622–3634. https://doi.org/10.1523/JNEUROSCI.1716-20.2021

Jeffreys, H. (1998). *Theory of Probability*. Clarendon Press. https://books.google.is/books?id=_PuRmAEACAAJ

Jóhannesson, Ó. I., Thornton, I. M., Smith, I. J., Chetverikov, A., & Kristjánsson, Á. (2016). Visual Foraging With Fingers and Eye Gaze. *I-Perception*, *7*(2), 204166951663727. https://doi.org/10.1177/2041669516637279

Kärcher, S. M., Fenzlaff, S., Hartmann, D., Nagel, S. K., & König, P. (2012). Sensory Augmentation for the Blind. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00037

Kassambara A. (2021). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests* [Computer software]. https://cran.r-project.org/web/packages/rstatix/index.html

Kolarik, A. J., Timmis, M. A., Cirstea, S., & Pardhan, S. (2014). Sensory substitution information informs locomotor adjustments when walking through apertures. *Experimental Brain Research*, *232*(3), 975–984. https://doi.org/10.1007/s00221-013-3809-5

Kristjánsson, Á. (2006). Rapid learning in attention shifts: A review. *Visual Cognition*, *13*(3), 324–362. https://doi.org/10.1080/13506280544000039

Kristjánsson, Á. (2015). Reconsidering Visual Search. *I-Perception*, *6*(6), 204166951561467. https://doi.org/10.1177/2041669515614670

Kristjánsson, Á., Björnsson, A. S., & Kristjánsson, T. (2020). Foraging with Anne Treisman: Features versus conjunctions, patch leaving and memory for foraged locations. *Attention, Perception, & Psychophysics*, *82*(2), 818–831. https://doi.org/10.3758/s13414-019-01941-y

Kristjánsson, Á., & Egeth, H. (2020). How feature integration theory integrated cognitive psychology, neurophysiology, and psychophysics. *Attention, Perception, & Psychophysics*, *82*(1), 7–23. https://doi.org/10.3758/s13414-019-01803-7

Kristjánsson, Á., Jóhannesson, Ó. I., & Thornton, I. M. (2014). Common Attentional Constraints in Visual Foraging. *PLoS ONE*, *9*(6), e100752. https://doi.org/10.1371/journal.pone.0100752

Kristjánsson, Á., Moldoveanu, A., Jóhannesson, Ó. I., Balan, O., Spagnol, S., Valgeirsdóttir, V. V., & Unnthorsson, R. (2016). Designing sensory-substitution devices: Principles, pitfalls and potential1. *Restorative Neurology and Neuroscience*, *34*(5), 769–787. https://doi.org/10.3233/RNN-160647

Kristjánsson, Á., Ólafsdóttir, I. M., & Kristjánsson, T. (2019). Visual Foraging Tasks Provide New Insights into the Orienting of Visual Attention: Methodological Considerations. In S. Pollmann (Ed.), *Spatial Learning and Attention Guidance* (Vol. 151, pp. 3–21). Springer US. https://doi.org/10.1007/7657_2019_21

Kristjánsson, T., Thornton, I. M., & Kristjánsson, Á. (2018). Time limits during visual foraging reveal flexible working memory templates. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(6), 827–835. https://doi.org/10.1037/xhp0000517

Kupers, R., Chebat, D. R., Madsen, K. H., Paulson, O. B., Ptito, M., & Mishkin, M. (2010). Neural correlates of virtual route recognition in congenital blindness. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(28), 12716–12721.

*L5 Actuator description*. (2019). Lofelt GmbH.

Lawrence MA. (2016). *_ez: Easy Analysis and Visualization of Factorial Experiments_. R package version 4.4-0* [Computer software]. <https://CRAN.R-project.org/package=ez>

Lechelt, E. C., & Borchert, R. (1977). The interdependence of time and space in somesthesis: The Tau effect reexamined. *Bulletin of the Psychonomic Society*, *10*(3), 191–193. https://doi.org/10.3758/BF03329320

Lee, S.-H., & Blake, R. (1999). Visual Form Created Solely from Temporal Structure. *Science*, *284*(5417), 1165–1168. https://doi.org/10.1126/science.284.5417.1165

Leonards, U., Singer, W., & Fahle, M. (1996). The Influence of Temporal Phase Differences on Texture Segmentation. *Vision Research*, *36*(17), 2689–2697. https://doi.org/10.1016/0042-6989(96)86829-5

Levy-Tzedek, S., Hanassy, S., Abboud, S., Maidenbaum, S., & Amedi, A. (2012). Fast, accurate reaching movements with a visual-to-auditory sensory substitution device. *Restorative Neurology and Neuroscience*, *30*(4), 313–323. https://doi.org/10.3233/RNN-2012-110219

Liu, Y.-C. (2001). Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveller information systems. *Ergonomics*, *44*(4), 425–442. https://doi.org/10.1080/00140130010011369

Loomis, J. M., Golledge, R. G., & Klatzky, R. L. (1998). Navigation System for the Blind: Auditory Display Modes and Guidance. *Presence: Teleoperators and Virtual Environments*, *7*(2), 193–203. https://doi.org/10.1162/105474698565677

Loomis, J. M., Klatzky, R. L., & Giudice, N. A. (2018). Sensory substitution of vision: Importance of perceptual and cognitive processing. In R. Manduchi & S. Kurniawan (Eds.), *Assistive Technology for Blindness and Low Vision* (0 ed., pp. 179–210). CRC Press. https://doi.org/10.1201/9781315216935-12

Loomis, J. M., Marston, J. R., Golledge, R. G., & Klatzky, R. L. (2005). Personal Guidance System for People with Visual Impairment: A Comparison of Spatial Displays for Route Guidance. *Journal of Visual Impairment & Blindness*, *99*(4), 219–232. https://doi.org/10.1177/0145482X0509900404

Luck, S. J., Hillyard, S. A., Mangun, G. R., & Gazzaniga, M. S. (1989). *Independent hemispheric attentional systems mediate visual search in split-brain patients. 342*(6249), 543–545.

Lumpkin, E. A., Marshall, K. L., & Nelson, A. M. (2010). The cell biology of touch. *Journal of Cell Biology*, *191*(2), 237–248. https://doi.org/10.1083/jcb.201006074

Maidenbaum, S., Hanassy, S., Abboud, S., Buchs, G., Chebat, D.-R., Levy-Tzedek, S., & Amedi, A. (2014). The "EyeCane", a new electronic travel aid for the blind: Technology, behavior & swift learning. *Restorative Neurology and Neuroscience*, *32*(6), 813–824. https://doi.org/10.3233/RNN-130351

Makarov, I., Stefánsson Thors, S. S., Ævarsson, E. A., Jörgensson, F. K. P., Yeganeh, N., Kristjánsson, Á., & Unnthorsson, R. (2024). The Haptic Intensity Order Illusion Is Caused by Amplitude Changes. *ACM Transactions on Applied Perception*, *21*(1), 1–18. https://doi.org/10.1145/3626237

Makarov, I., Unnthorsson, R., Kristjánsson, Á., & Thornton, I. M. (2024). The effects of visual and auditory synchrony on human foraging. *Attention, Perception, & Psychophysics*, *86*, 909–930. https://doi.org/10.3758/s13414-023-02840-z

Marks, L. E. (1974). On Associations of Light and Sound: The Mediation of Brightness, Pitch, and Loudness. *The American Journal of Psychology*, *87*(1/2), 173. https://doi.org/10.2307/1422011

Marston, J. R., Loomis, J. M., Klatzky, R. L., & Golledge, R. G. (2007). Nonvisual Route following with Guidance from a Simple Haptic or Auditory Display. *Journal of Visual Impairment & Blindness*, *101*(4), 203–211. https://doi.org/10.1177/0145482X0710100403

Marston, J. R., Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Smith, E. L. (2006). Evaluation of spatial displays for navigation without sight. *ACM Transactions on Applied Perception*, *3*(2), 110–124. https://doi.org/10.1145/1141897.1141900

Martino, G., & Marks, L. E. (1999). Perceptual and Linguistic Interactions in Speeded Classification: Tests of the Semantic Coding Hypothesis. *Perception*, *28*(7), 903–923. https://doi.org/10.1068/p2866

Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, *43*(3), 679–690. https://doi.org/10.3758/s13428-010-0049-5

McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics*, *25*(3), 221–224. https://doi.org/10.3758/BF03202990

McGurk, H., & McDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. https://doi.org/doi:10.1038/264746a0

McMains, S. A., & Kastner, S. (2009). Visual attention. In *Encyclopedia of neuroscience* (M. D. Binder, N. Hirokawa, U. Windhorst (Eds.), pp. 4299–4302). Springer.

Merabet, L. B., & Pascual-Leone, A. (2010). Neural reorganization following sensory loss: The opportunity of change. *Nature Reviews Neuroscience*, *11*(1), 44–52. https://doi.org/10.1038/nrn2758

Millar, S., & Al-Attar, Z. (2002). The Müller-Lyer illusion in touch and vision: Implications for multisensory processes. *Perception & Psychophysics*, *64*(3), 353–365. https://doi.org/10.3758/BF03194709

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, *6*, 414–417. https://doi.org/doi:10.1016/0166-2236(83)90190-x

Mondloch, C. J., & Maurer, D. (2004). Do small white balls squeak? Pitch-object correspondences in young children. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(2), 133–136. https://doi.org/10.3758/CABN.4.2.133

Most, S. B., Scholl, B. J., Clifford, E. R., & Simons, D. J. (2005). What You See Is What You Set: Sustained Inattentional Blindness and the Capture of Awareness. *Psychological Review*, *112*(1), 217–242. https://doi.org/10.1037/0033-295X.112.1.217

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, *86*(3), 214–255. https://doi.org/10.1037/0033-295X.86.3.214

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Nguyen, T. T. N., Vuong, Q. C., Mather, G., & Thornton, I. M. (2021). Ensemble coding of crowd speed using biological motion. *Attention, Perception, & Psychophysics*, *83*(3), 1014–1035. https://doi.org/10.3758/s13414-020-02163-3

Nimmons, G. L., Kang, R. S., Drennan, W. R., Longnion, J., Ruffin, C., Worman, T., Yueh, B., & Rubinstein, J. T. (2008). Clinical Assessment of Music Perception in Cochlear Implant Listeners. *Otology & Neurotology*, *29*(2), 149–155. https://doi.org/10.1097/mao.0b013e31812f7244

Ólafsdóttir, I. M., Gestsdóttir, S., & Kristjánsson, Á. (2020). Age differences in foraging and executive functions: A cross-sectional study. *Journal of Experimental Child Psychology*, *198*, 104910. https://doi.org/10.1016/j.jecp.2020.104910

Ólafsdóttir, I. M., Gestsdóttir, S., & Kristjánsson, Á. (2021). The development of foraging organization. *Attention, Perception, & Psychophysics*, *83*(7), 2891–2904. https://doi.org/10.3758/s13414-021-02328-8

Ólafsdóttir, I. M., Kristjánsson, T., Gestsdóttir, S., Jóhannesson, Ó. I., & Kristjánsson, Á. (2016). Understanding visual attention in childhood: Insights from a new visual foraging task. *Cognitive Research: Principles and Implications*, *1*(1), 18. https://doi.org/10.1186/s41235-016-0016-5

Ortiz, T., Poch, J., Santos, J. M., Requena, C., Martínez, A. M., Ortiz-Terán, L., Turrero, A., Barcia, J., Nogales, R., Calvo, A., Martínez, J. M., Córdoba, J. L., & Pascual-Leone, A. (2011). Recruitment of Occipital Cortex during Sensory Substitution Training Linked to Subjective Experience of Seeing in People with Blindness. *PLoS ONE*, *6*(8), e23264. https://doi.org/10.1371/journal.pone.0023264

Pasqualotto, A., & Proulx, M. J. (2012). The role of visual experience for the neural basis of spatial cognition. *Neuroscience & Biobehavioral Reviews*, *36*(4), 1179–1187. https://doi.org/10.1016/j.neubiorev.2012.01.008

Plaisier, M. A., Sap, L. I. N., & Kappers, A. M. L. (2020). Perception of vibrotactile distance on the back. *Scientific Reports*, *10*(1), 17876. https://doi.org/10.1038/s41598-020-74835-x

Poirier, C., De Volder, A., Tranduy, D., & Scheiber, C. (2007). Pattern recognition using a device substituting audition for vision in blindfolded sighted subjects. *Neuropsychologia*, *45*(5), 1108–1121. https://doi.org/10.1016/j.neuropsychologia.2006.09.018

Prevoteau, C., Chen, S. Y., & Lalwani, A. K. (2018). Music enjoyment with cochlear implantation. *Auris Nasus Larynx*, *45*(5), 895–902. https://doi.org/10.1016/j.anl.2017.11.008

Ptito, M., Matteau, I., Gjedde, A., & Kupers, R. (2009). Recruitment of the middle temporal area by tactile motion in congenital blindness. *NeuroReport*, *20*(6), 543–547. https://doi.org/10.1097/WNR.0b013e3283279909

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology*, *25*, 111. https://doi.org/10.2307/271063

Ratan Murty, N. A., Teng, S., Beeler, D., Mynick, A., Oliva, A., & Kanwisher, N. (2020). Visual experience is not necessary for the development of face-selectivity in the lateral fusiform gyrus. *Proceedings of the National Academy of Sciences*, *117*(37), 23011–23020. https://doi.org/10.1073/pnas.2004607117

Reich, L., Szwed, M., Cohen, L., & Amedi, A. (2011). A Ventral Visual Stream Reading Center Independent of Visual Experience. *Current Biology*, *21*(5), 363–368. https://doi.org/10.1016/j.cub.2011.01.040

Reisberg, D. (1978). Looking where you listen: Visual cues and auditory attention. *Acta Psychologica*, *42*(4), 331–341. https://doi.org/10.1016/0001-6918(78)90007-0

Renier, L., Collignon, O., Poirier, C., Tranduy, D., Vanlierde, A., Bol, A., Veraart, C., & Devolder, A. (2005). Cross-modal activation of visual cortex during depth perception using auditory substitution of vision. *NeuroImage*, *26*(2), 573–580. https://doi.org/10.1016/j.neuroimage.2005.01.047

Rensink, R. A. (2000). Visual Search for Change: A Probe into the Nature of Attentional Processing. *Visual Cognition*, *7*(1–3), 345–376. https://doi.org/10.1080/135062800394847

Ricciardi, E., Handjaras, G., & Pietrini, P. (2014). The blind brain: How (lack of) vision shapes the morphological and functional architecture of the human brain. *Experimental Biology and Medicine*, *239*(11), 1414–1420. https://doi.org/10.1177/1535370214538740

RStudio Team. (2020). *RStudio: Integrated Development for R* [Computer software]. http://www.rstudio.com/

Sadato, N., Pascual-Leone, A., Grafman, J., Ibañez, V., Deiber, M.-P., Dold, G., & Hallet, M. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature*, *380*(6574), 526–528. https://doi.org/doi:10.1038/380526a0

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review*, *115*(1), 101–130. https://doi.org/10.1037/0033-295X.115.1.101

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, *408*(6814), 788–788. https://doi.org/10.1038/35048669

Shen, L., Lu, X., Wang, Y., & Jiang, Y. (2023). Audiovisual correspondence facilitates the visual search for biological motion. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02308-z

Sherrick, C. E., & Rogers, R. (1966). Apparent haptic movement. *Perception & Psychophysics*, *1*, 175–180. https://doi.org/10.3758/BF03215780

Sinnett, S., Costa, A., & Soto-Faraco, S. (2006). Manipulating inattentional blindness within and across sensory modalities. *Quarterly Journal of Experimental Psychology*, *59*(8), 1425–1442. https://doi.org/10.1080/17470210500298948

Sofia, K. O., & Jones, L. (2013). Mechanical and Psychophysical Studies of Surface Wave Propagation during Vibrotactile Stimulation. *IEEE Transactions on Haptics*, *6*(3), 320–329. https://doi.org/10.1109/TOH.2013.1

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971–995. https://doi.org/10.3758/s13414-010-0073-7

Striem-Amit, E., Ovadia-Caro, S., Caramazza, A., Margulies, D. S., Villringer, A., & Amedi, A. (2015). Functional connectivity of visual cortex in the blind follows retinotopic organization principles. *Brain*, *138*(6), 1679–1695. https://doi.org/10.1093/brain/awv083

Suto, Y. (1952). The effect of space on time estimation (S-effect) in tactual space. *Japanese Journal of Psychology*, *22*, 45–57.

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, *14*(9), 400–410. https://doi.org/10.1016/j.tics.2010.06.008

Thornton, I. M., & Horowitz, T. S. (2004). The multi-item localization (MILO) task: Measuring the spatiotemporal context of vision for action. *Perception & Psychophysics*, *66*(1), 38–50. https://doi.org/10.3758/BF03194859

Thornton, I. M., & Horowitz, T. S. (2020). Searching Through Alternating Sequences: Working Memory and Inhibitory Tagging Mechanisms Revealed Using the MILO Task. *I-Perception*, *11*(5), 204166952095801. https://doi.org/10.1177/2041669520958018

Thornton, I. M., Tagu, J., Zdravković, S., & Kristjánsson, Á. (2021). The Predation Game: Does dividing attention affect patterns of human foraging? *Cognitive Research: Principles and Implications*, *6*(1), 35. https://doi.org/10.1186/s41235-021-00299-w

Thornton, I. M., Vuong, Q. C., & Pilz, K. S. (2021). A Search Advantage for Horizontal Targets in Dynamic Displays. *I-Perception*, *12*(2), 204166952110046. https://doi.org/10.1177/20416695211004616

Treisman, A. (1960). Contextual Cues in Selective Listening. *Quarterly Journal of Experimental Psychology*, *12*(4), 242–248. https://doi.org/10.1080/17470216008416732

Treisman, A. (1977). Focused attention in the perception and retrieval of multidimensional stimuli. *Perception & Psychophysics*, *22*(1), 1–11. https://doi.org/10.3758/BF03206074

Treisman, A. (1999). Solutions to the Binding Problem. *Neuron*, *24*(1), 105–125. https://doi.org/doi:10.1016/s0896-6273(00)80826-0

Treisman, A., & Gelade, G. (1980). A Feature-Integration Theory of Attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/doi:10.1016/0010-0285(80)90005-5

Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1053–1065. https://doi.org/10.1037/0096-1523.34.5.1053

Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2009). Poke and pop: Tactile–visual synchrony increases visual saliency. *Neuroscience Letters*, *450*(1), 60–64. https://doi.org/10.1016/j.neulet.2008.11.002

Van der Burg, E., Talsma, D., Olivers, C. N. L., Hickey, C., & Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects. *NeuroImage*, *55*(3), 1208–1218. https://doi.org/10.1016/j.neuroimage.2010.12.068

Van Erp, J. B. (2005). Presenting directions with a vibrotactile torso display. *Ergonomics*, *48*(3), 302–313. https://doi.org/10.1080/0014013042000327670

Violentyev, A., Shimojo, S., & Shams, L. (2005). Touch-induced visual illusion. *Neuroreport.*, *16*(10), 1107–1110. https://doi.org/doi: 10.1097/00001756-200507130-00015.

Vitevitch, M. S. (2003). Change deafness: The inability to detect changes between two voices. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 333–342. https://doi.org/10.1037/0096-1523.29.2.333

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin*, *138*(6), 1172–1217. https://doi.org/10.1037/a0029333

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., & Van Der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*, *100*(3), 426–432. https://doi.org/10.1037/a0022790

Wagner, S., Winner, E., Cicchetti, D., & Gardner, H. (1981). ‘Metaphorical’ Mapping in Human Infants. *Child Development*, *52*(2), 728. https://doi.org/10.2307/1129200

Weber, E. H. (1996). *E.H. Weber on the Tactile Senses* (1st ed.). Psychology Press.

Welch, R. B., DutionHurt, L. D., & Warren, D. H. (1986). Contributions of audition and vision to temporal rate perception. *Perception & Psychophysics*, *39*(4), 294–300. https://doi.org/10.3758/BF03204939

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. https://doi.org/10.1080/14639220210123806

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., ... Yutani, H.

(2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wolfe, J. M. (1994). Visual Search in Continuous, Naturalistic Stimuli. *Vision Research*, *34*(9), 1187–1195.

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, *28*(4), 1060–1092. https://doi.org/10.3758/s13423-020-01859-9

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided Search: An Alternative to the Feature Integration Model for Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419–433. https://doi.org/doi:10.1037/0096-1523.15.3.419

Woods, A. J., Göksun, T., Chatterjee, A., Zelonis, S., Mehta, A., & Smith, S. E. (2013). The development of organized visual search. *Acta Psychologica*, *143*(2), 191–199. https://doi.org/10.1016/j.actpsy.2013.03.008

Wright, R., & Uchanski, R. M. (2012). Music Perception and Appraisal: Cochlear Implant Users and Simulated Cochlear Implant Listening. *Journal of the American Academy of Audiology*, *23*(05), 350–365. https://doi.org/10.3766/jaaa.23.5.6

Yeganeh, N., Makarov, I., Kristjansson, Á., & Unnthorsson, R. (2024). Discrimination Accuracy of Sequential Versus Simultaneous Vibrotactile Stimulation on the Forearm. *Applied Sciences*, *14*(1). https://doi.org/10.3390/app14010043

Yeganeh, N., Makarov, I., Stefánsson Thors, S. S., Kristjánsson, Á., & Unnthorsson, R. (2022). Evaluating the Optimum Distance between Voice Coil Actuators Using the Relative Point Localization Method on the Forearm. *Actuators*, *12*(1), 6. https://doi.org/10.3390/act12010006

Yeganeh, N., Makarov, I., Unnthorsson, R., & Kristjánsson, Á. (2023). Effects of Stimulus Frequency and Location on Vibrotactile Discrimination Performance Using Voice Coil Actuators on the Forearm. *Actuators*, *12*(6), 224. https://doi.org/10.3390/act12060224

Ziat, M., Smith, E., Brown, C., DeWolfe, C., & Hayward, V. (2014). Ebbinghaus illusion in the tactile modality. *2014 IEEE Haptics Symposium (HAPTICS)*, 581–585. https://doi.org/10.1109/HAPTICS.2014.6775520

Zou, H., Muller, H. J., & Shi, Z. (2012). Non-spatial sounds regulate eye movements and enhance visual search. *Journal of Vision*, *12*(5), 2–2. https://doi.org/10.1167/12.5.2

# Original Publications

This thesis is based on the following original publications, which are referred to in the text by their Roman numerals (I, II, III, ... [as needed]):

I. Makarov, I., Stefánsson Thors, S. S., Ævarsson, E. A., Jörgensson, F. K. P., Yeganeh, N., Kristjánsson, Á., & Unnthorsson, R. (2024). The Haptic Intensity Order Illusion Is Caused by Amplitude Changes. *ACM Transactions on Applied Perception, 21*(1), 1–18. https://doi.org/10.1145/3626237

II. Makarov, I., Unnthorsson, R., Kristjánsson, Á., & Thornton, I. M. (2024). The effects of visual and auditory synchrony on human foraging. *Attention, Perception, & Psychophysics, 86*, 909–930. https://doi.org/10.3758/s13414-023-02840-z

III. Makarov, I., Unnthorsson, R., Kristjansson, A., & Thornton, I. M. (2024, January 11). Cross-modal cues improve the detection of synchronized targets during human foraging. https://doi.org/10.31234/osf.io/xrty3. Accepted in Multisensory Research.

All papers are reprinted by kind permission of the publishers.

# Paper I

Paper I

# Paper II

# Paper II

# Paper III