8-24-2023 1:00 AM

The Roles of Vestibular and Proprioceptive Signals in Updating Spatial Selective Auditory Attention during Head Motion

Erisa Davoudi, *Western University*

Supervisor: Macpherson, Ewan A., *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Neuroscience

Abstract

Spatial Selective Auditory Attention (SSAA) allows individuals to attend to a desired sound's location in an acoustically rich environment. This research project explored the compensatory updating of SSAA focus during head movements and the roles of vestibular, proprioceptive, and visual self-motion signals in this process. A behavioural auditory selective attention task was conducted with and without visual cues and in three different motion conditions that manipulated the availability of vestibular and proprioceptive signals: Static (no updating required), Active head rotation (vestibular and proprioceptive signals available), and Passive whole-body rotation (only vestibular signals available). The findings suggest that listeners do appropriately update their SSAA while moving their heads, although with a slight frontal bias. Performance in the Passive condition was almost equal to the Active condition, indicating that vestibular signals were sufficient for SSAA updating, while proprioceptive signals were not necessary. Visual cues improved SSAA updating somewhat, but were not essential.

Keywords:

Auditory Attention, Selective Spatial Attention, Vestibular, Proprioceptive, Spatial Attention Updating, Self-Motion, Vision, Spatial Hearing.

Summary for Lay Audience

Human beings exhibit a remarkable capacity to selectively attend to particular sounds of interest in complex and noisy environments. Nevertheless, this process of selective auditory attention goes beyond mere identification of sound location. In a world where everything, including ourselves, moves, we must constantly update our selective auditory attention to adapt to new sound locations. We are provided with some internal signals about our movement including vestibular and proprioceptive cues generated from our inner ear and neck muscles respectively. These cues enable our attentional ability called Spatially Selective Auditory Attention (SSAA) to be on duty. Yet, it is not clear which internal signal (vestibular or proprioceptive) is mainly responsible to continually update our SSAA to a specific sound while we move our head's direction during our everyday lives. In this research project we explored this question through a behavioral listening-attending task, in which individuals were required to update their focus of SSAA to a stationary target loudspeaker while undergoing voluntary (neck) or passive (whole-body) rotations.

Our findings suggest that among all the internal cues at our disposal, the signals generated by the vestibular system are dominant and sufficiently support our ability to update SSAA during movement, but that visual and proprioceptive signals did not play a necessary role in this process. This study pioneers the investigation of the weighting of internal cues on SSAA during motion.

Acknowledgment

I would like to express my heartfelt gratitude to my great supervisor, Dr. Ewan Macpherson, for his unwavering support and care throughout this master's thesis. His expertise and willingness to share knowledge have shaped my research approach. I have learned a lot from his patience and perspectives in life.

I extend my sincere thanks to my advisory committee members for their insightful feedback. Their dedication to reviewing my work and offering expert advice have significantly enriched the quality of this thesis.

My heartfelt appreciation goes to my dear friends, M.Sh who gave me statistical insights and stayed up late with me, and A.B who kept me grounded and motivated during this entire master’s/immigration journey.

Finally, I am the most grateful for my loving family (Neda, Mehrdad, Parsa) who has always stood by my side and believed in me, giving me unconditional support and love in all situations.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLES

1 Introduction

1.1 Overview

In everyday listening situations, amidst multiple conversations and various sound stimuli, we possess the ability to select topics of interest, focus our attention on specific sound sources, and avoid distractions from others. We are not static observers in these scenarios; rather, we actively engage with our surroundings. We are involved in identifying a particular sound source, determining its location, orienting our attention towards it, and ultimately maintaining focus on the sound source. This process engages various cognitive functions, such as attention and spatial orientation, utilizing our visual and auditory systems, as well as sensorimotor information about and derived from our normal movements. As informative cues change with our movements, our perception of this information must also adapt. Thus, during head motions we need to continuously update the direction of our attention in head-centred coordinates to maintain that attention on an ongoing sound source's location or to maintain pre-allocated attention to an expected location. The main aim of this study is to determine the roles of vestibular, proprioceptive and visual cues for head motion in listeners' ability to change the direction of spatial selective auditory attention (SSAA) to the desired sound's location. We refer to this process as "SSAA updating".

The roles of these cues have been previously investigated in updating the spatial representation of already heard sounds ("spatial updating") and in dynamic sound localization, yet not in the SSAA updating. Also, there is limited research on the mechanisms of selective auditory attention in relation to moving listeners or moving sound sources to effectively focus on specific sound sources in dynamic spaces (Davis et al., 2016)

The initial step in directing attention to a desired sound involves localizing the sound source. When the sound source location is static, a mental representation of that location is formed. However, sound localization can be challenging when the listener is stationary, requiring head rotations to enhance the reception of accurate auditory cues. Head movements are particularly crucial for front/back localization, which will be elaborated on in detail in the following section.

In addition to the advantages associated with head movements for sound localization, there is a noteworthy challenge pertaining to attending to the selected sound source. In order to generate appropriate responses and compensate for head motion in terms of attention, information regarding the behavior and actions of the head must be integrated with the systems involved in attention and localization. This information serves as a foundation for the modulation and maintenance of spatial auditory attention, as well as the dynamic localization cues (as discussed in section 1.2.2), in relation to the representation of sound source location.

The challenge of focusing SSAA during motion has been investigated in previous research studies from our laboratory and it has been shown that head movement can somewhat decrease listeners' ability to selectively attend to a frontal target in the presence of flanking distractors (Macpherson & Ellis, 2016; Macpherson & Ransom, 2018). However, SSAA when head movement is involved remains possible in such situations and indeed contributes to the target identification (Macpherson et al., 2019).

For the purpose of following this study's objectives and design, the background information included in the rest of this introduction discusses:

- Sound localization: To understand how the auditory system makes a representation of sound source location from auditory cues.
- Perception of self-motion and signals associated with it: To understand the sensory systems sensitive to head rotation and how they interact with each other.
- Auditory attention and more specifically SSAA: Introducing the spatial and non-spatial aspects of selective auditory attention.
- The importance of self-motion signals in audition: To understand what information these signals carry about motion and how they integrate with the systems involved in auditory attention.

1.2 Sound Localization Cues

1.2.1 Static Cues: Binaural and Spectral

Sound localization plays a crucial role in human auditory perception as it enables navigation in the acoustic environment, identification of sound sources, and effective interaction with surroundings. The accurate perception of sound location relies on the positioning of the ears on opposite sides of the head. This arrangement allows the auditory system to capture sound from two spatial locations simultaneously. When sounds are not in line with the listener's midline, they are experienced under slightly different acoustic conditions. The information collected from both ears constitutes the binaural cues used to determine the location of sounds.

Humans with normal hearing rely on binaural cues that rely on the comparison of sound arriving at each ear to assess the side-to-side position of a sound source in terms of its azimuth, contributing to sound localization abilities (Rayleigh, 1907). These interaural difference cues serve as static sound localization cues, providing information about the stationary position of a sound source. One binaural cue is the interaural time difference (ITD). ITD refers to the slight time delay between a sound arriving at one ear versus the other ear. The brain processes this time difference and uses it to estimate the lateral position of the sound source. ITD is most effective for low-frequency sounds, for which the wavelength is larger than the dimensions of the head. For higher-frequency sounds, another important binaural cue, known as the interaural level difference (ILD), comes into play (Feddersen et al., 2005; Macpherson & Middlebrooks, 2002). ILD is the difference in sound intensity between the ears, resulting from the head acting as a barrier to sound. The head shadow causes a reduction in sound intensity on the side opposite the sound source, providing valuable information for sound localization (Mills, 1958).

Static binaural cues (ILD, ITD) can only determine the lateral position of a sound source with respect to the median plane (in other words, constraining the location to a particular “cone of confusion”), but spectral cues associated with the shape of the pinnae are specific to vertical directions, enabling the perception of the elevation and front/back location of a sound source on the vertical plane. The unique shape of the pinna and its interaction with the head and shoulders cause modifications to the spectral content of incoming sounds. These modifications result in frequency-dependent amplitude changes and phase shifts, which vary with the sound source location (Shaw, 1997). By analyzing these spectral modifications, the auditory system can extract information about the vertical and front/back position of the sound source. Spectral cues in frequency ranges beyond approximately 4 kHz play a role in differentiating sound sources positioned at the front versus those positioned at the back of a listener's head (Hebrank & Wright, 1974; Zhang & Hartmann, 2010).

1.2.2 Dynamic Sound Localization Cues

During sound localization, interaural level and time differences (ILD and ITD) provide lateral location information for a sound source. However, when binaural cues become ambiguous within the cone of confusion, head movements can aid in resolving localization ambiguities. As mentioned above, spectral cues above 4 kHz, resulting from direction-dependent sound filtering by the pinnae, contribute to this process, but using narrowband stimuli can lead to localization errors (Hebrank & Wright, 2005). Active sound localization, which involves head movement, assists in resolving front/back confusions. By considering the relationship between head motion and self-generated changes in ILD and ITD, dynamic sound localization cues clarify sound source locations (Macpherson & Kerr, 2008) and also the utilization of dynamic localization cues relies on the vestibular and proprioceptive systems since the auditory system needs precise information about the head's orientation and motion in space to interpret these dynamic cues effectively (Wallach, 1940).

On the other hand, Carlile and Leung (2016) conducted experiments shedding light on the costs associated with head movement during localization. During rapid head turns, systematic mislocalization of lateral angles, particularly at the rear, was observed. Interestingly, the smallest localization errors occurred when participants turned their heads towards the point of attention, indicating the role of attention in mitigating cue smearing.

1.3 Perception of Self-Motion

In the world we inhabit, movement is an inherent aspect of our lives, and our ability to perceive it accurately is important as we attempt to comprehend our surroundings. As we move, this movement is sensed by a stream of sensory information that informs the brain about our motion and orientation; Our visual system detects our movement and direction through patterns of optic flow, while proprioceptors perceive the torsion of our neck muscles with the involvement of efference copy signals sending motor commands to these muscles, and the vestibular organs sense the accelerations of our head. Additionally, our auditory system detects changes in the sound waves reaching our ears to provide information about our locomotion. These sensory inputs are integrated by the central nervous system to form an internal representation of our own movement. (St George & Fitzpatrick, 2011). In the following subsections, we will explore the specific roles of these signals generated from sensory systems and their integration in the brain for the perception of self-motion.

1.3.1 Vestibular System: How Self-Motion and Orientation are Sensed by Vestibular Stimulation.

The vestibular system holds a fundamental significance in our ability to perceive balance and spatial orientation, making essential contributions to our subjective sense of movement and position in space. The vestibular information is then integrated with other sensory signals, such as visual and proprioceptive signals, to create a coherent perception of self-motion. (Colclasure & Holt, 2007).

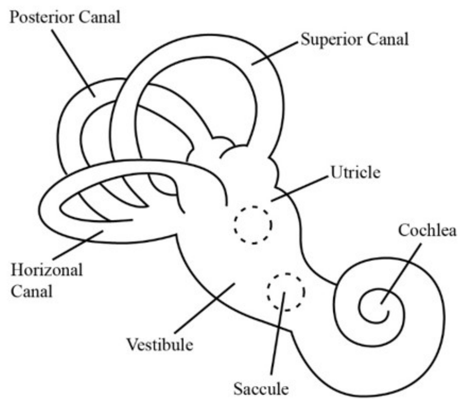

Figure 1.1 Representation of the basic anatomy of the peripheral vestibular system. Image from medlink.com website.

The vestibular system is located in the inner ear and consists of two types of motion-sensing organs: the semicircular canals and the vestibule. The semicircular canals, called superior, posterior, and lateral canals, detect rotational head movements in different directions. On the other hand, the vestibule, made up of the utricle and saccule, senses linear acceleration, like moving forward or up and down. Together, these organs help us measure our head movements in different ways (Highstein & Holstein, 2012). Inside the semicircular canals and the utricle and saccule, there are tiny hair cells that convert mechanical signals into electrical signals. When we rotate our head, the fluid in the semicircular canals moves and bends the hair cells, changing their activity. Similarly, when we experience linear acceleration, the hair cells in the utricle and saccule get bent, sending sensory information to the nerves and the central nervous system. The vestibular system plays a role in various neural and cognitive functions (Day & Fitzpatrick, 2005). Studies have shown that the vestibular system is particularly important for perceiving self-motion in the absence of visual cues. For example, when we are in complete darkness or when we are moving in a vehicle with no windows, our perception of self-motion relies heavily on vestibular signals (Dichgans & Brandt, 1978; Merfeld et al., 1999).

Furthermore, the results of a study demonstrate a clear pattern in motion thresholds by the vestibular organs, whereby they show an upward trend at frequencies below 0.5 Hz. This pattern suggests that sensitivity to motion declines gradually as frequencies decrease, resembling the characteristic high-pass response observed in the semicircular canals of vestibular system (Grabherr et al., 2008). However, vestibulo-ocular reflex (VOR) testing shows that the vestibular system accurately represents head oscillations from 0.1 to 10 Hz (Zalewski, 2017).

1.3.2 Proprioceptive and Efference Copy Signals.

Proprioceptive signals, originating from specialized sensory receptors found in muscles, tendons, and joints, play a crucial role in detecting changes in body position and movement. These signals contribute vital information about the relative positions and movements of body parts, aiding in the accurate perception of self-motion. The integration of proprioceptive cues with visual and vestibular information enhances the overall perception of self-motion (Cullen & Zobeiri, 2021). In addition to proprioceptive signals, efference copy signals are generated by the motor system as duplicates of the command signals sent to muscles. These efference copy signals enable the brain to anticipate the sensory consequences of movement. By providing predictions about incoming sensory information, these signals alleviate the processing load on sensory systems and facilitate efficient processing by allowing them to anticipate and prepare for the expected sensory input. The brain integrates multiple sources of sensory information, including proprioceptive and efference copy signals, to form a coherent perception of self-motion, which allows for a more accurate perception and better control of movement. In laboratory settings, researchers can manipulate proprioceptive and efference copy signals to study involuntary motion situations (Crowell et al., 1998; Cullen et al., 2011). However, in everyday life, these signals naturally occur together during movements, leading to the consideration of them as a single piece of information for voluntary motion. In the current study, we refer to the combined proprioceptive and efference copy signals together (Lackner & DiZio, 2005).

1.3.3 Visual Information for Self-Motion

According to (Gibson, 1950) the observer's motion is a critical source of information about the environment, and it plays a central role in the perception of objects and events. Gibson's theory of perception emphasizes the active role of the observer and the visual system in detecting the motion and perceiving the environment. A phenomenon called "vection" happens when purely visual stimuli induce the sensation of self-motion because the brain processes the changes in the visual field and interprets it as motion (Howard & Howard, 1994). Other studies conducted by Lappe et al. (1999) and Cutting & Vishton (1995) also indicate that visual cues play a crucial role in perceiving self-motion. These visual cues, such as optic flow and local relative motion cues, can be used to estimate the direction of self-motion (heading) quickly and effectively.

Furthermore, the interaction between the visual and vestibular information can explain the powerful sensations of self-motion and there is demonstration of the significance of both signals in the perception of self-motion in the literature (for review see Ramkhalawansingh et al., 2018; DeAngelis & Angelaki, 2011; Fetsch et al., 2009). A noteworthy investigation conducted by Dichgans and Brandt (1978) examined the impact of the interaction between visual and vestibular inputs on the perception of self-motion and control of posture. Employing a rotating chair to induce self-motion, the researchers manipulated visual cues during the experiment. The presence of visual cues exerted a substantial influence on the perception of self-motion and the ability to control posture, even when conflicting vestibular cues were present. Another study that investigated how we perceive self-motion with visual information was conducted using functional magnetic resonance imaging (fMRI) to measure brain activity while participants performed a task involving self-motion perception. The authors found that there are specific regions in the brain, including the Sylvian fissure, that are involved in integrating visual and vestibular information for accurate self-motion perception (Frank et al., 2016). On the other hand, recent studies have suggested that certain neurons involved in mediating self-motion perception may not be strongly driven by visual self-motion cues. For example, Gu et al. (2006) investigated how neurons in the medial superior temporal area respond to different types of sensory cues during three-dimensional heading selectivity tasks. The researchers recorded neural activity from monkeys while presenting them with combinations of visual and vestibular stimuli that simulated different types of self-motion. They found that while many neurons in this area responded to both visual and vestibular stimuli, some neurons were more strongly driven by vestibular cues than by visual cues.

1.4 Selective Auditory Attention

1.4.1 Auditory Scene Analysis

“Attention is a set of processes that modulate what information gets represented in the brain. These processes act similarly—and are even shared—across auditory and other sensory modalities” (Shinn-Cunningham, 2008). Nevertheless, to understand auditory attention, we need to have a good grasp of auditory perception. Auditory perception allows us to perceive and make sense of multiple sounds around us. To accomplish this, our auditory system creates a representation of auditory streams, enabling us to segregate and distinguish different sources of sound (Bregman, 1994; Cherry, 1953). The ability to segregate auditory streams is often referred to as Auditory Scene Analysis (ASA) (Bregman, 1994) (for reviews see Bronkhorst, 2000; Haykin & Chen, 2005; Masanao, 2001).

Auditory Scene Analysis involves utilizing various cues and perceptual processes to differentiate between different sound streams, which can be spatial or non-spatial in nature. Spatial separation has been demonstrated to be generally beneficial, although not the most dominant among the cues available to listeners for segregating sound streams in the presence of competing sound sources. The non-spatial cues, such as pitch contour and distinctive characteristics of the speaker or transmission channel, play an important role in maintaining attention on specific talkers over time (Broadbent, 1952; Egan et al., 1954). Additionally, research has shown that certain aspects of speech, like word emphasis or vocal tract size, can override spatial cues and contribute to selective attention and comprehension of speech (Darwin & Hukin, 2000). In essence, the availability of non-spatial cues reduces the reliance on spatial cues, nonetheless, the spatial cues would be more important when non-spatial cues are minimized. By integrating spatial and nonspatial cues, our auditory system and the brain create a coherent representation of the auditory scene, allowing us to isolate and focus on specific sound sources of interest (Shinn-Cunningham, 2008).

1.4.2 Spatial Selective Auditory Attention (SSAA)

The ability that allows individuals to focus their attention on a specific location of the target sound source and filter out other sounds from irrelevant directions is Spatial Selective Auditory Attention. Research on auditory spatial attention reveals that the way we allocate our attention can be explained by a gradient model. This implies that the concentration of attentional resources is highest at the location we are focusing on and gradually decreases as we move away from it and also the time required to fully engage attention to a spatial location falls within the range of 150 ms to 600 ms (Mondor & Zatorre, 1995). It has been shown that when the competing sound sources are far enough apart, there is a spatial release from masking, resulting in improved speech identification performance (e.g Plomp, 1976). Marrone et al. (2008) found that speech identification performance improved from the collocated condition to where the target and masker were 15 degrees apart. As the angle of separation increased beyond 15 degrees, the additional improvement in speech identification performance became less pronounced. In addition to orienting SSAA to the sound stimulus once they are heard, the pre-allocation of attention to the target location and active engagement of SSAA is also helpful to effectively separate and attend to specific auditory stimuli in complex environments (Kidd et al., 2005) and can also improve the detection and identification of single targets (Mondor & Zatorre, 1995; Quinlan & Bailey, 1995).

1.4.3 Maintaining and Updating SSAA

To maintain attention to the same location in world-centered coordinates during self-motion (updating SSAA), the auditory system needs to update the perception of the sound source location in head-centred coordinates (Fig. 1.2). Previous experiments have indicated that SSAA is less focused under dynamic head movements compared to a static position, although participants could still effectively update their SSAA when provided with sensorimotor cues (Macpherson et al., 2019; Macpherson & Ransom, 2018).

Maintaining attention on a source during self- or source motion has benefits over switching attention to the new location. Although voluntary attentional control that shifts attention toward specific locations or spectral features can be beneficial and fast, there are also some costs associated with it. For instance, a study conducted by Lin & Carlile (2015) revealed that switching auditory spatial attention during conversational turn-taking resulted in a decline in word recall and discourse comprehension. This finding highlights the challenges associated with shifting auditory selective attention and suggests potential limitations in the efficiency and ease of this process. Furthermore, another study by Koch et al. (2011) emphasizes the temporal inflexibility and potential time costs associated with intentionally switching auditory selective attention. We can quickly switch our attention to the next source of sound, however, maintaining attention on a moving source that needs our ongoing attention is beneficial as well.

1.5 Auditory Weighting of Vestibular and Proprioceptive information about Self-motion.

We have discussed the challenge faced to update the SSAA while we are in motion. Also, we have mentioned the sensorimotor cues (vestibular and proprioceptive/efference copy) that are available to us following the movement. While other studies have examined the weighting of vestibular and proprioceptive signals in representational auditory updating and front/back localization, little to no evidence exists on how these cues are weighted in updating of Spatial

Selective Auditory Attention. In the following section, we will discuss relevant findings pertaining to the weighting of vestibular and proprioceptive cues.

1.5.1 The Weighting of Self-Motion Signals in Sound Localization

Kim et al. (2013) conducted experiments to measure the influence of vestibular and proprioceptive cues on the ability of normally hearing individuals to accurately localize low-frequency sounds during head rotations. They asked participants to localize dynamic auditory targets while the two modalities (vestibular and proprioceptive) were individually, congruently, or incongruently stimulated. The results indicate that front/back localization using dynamic cues showed comparable performance in both active and passive rotation conditions. This finding suggests that vestibular inputs were sufficient in facilitating accurate localization. Additionally, when the proprioceptive input from the neck was isolated, performance was near chance, suggesting that neck proprioceptive cues are not sufficient for effective utilization of dynamic auditory cues. These results suggest that between the two sensorimotor cues (vestibular and proprioceptive), only the vestibular inputs are necessary and sufficient to inform the auditory system about head movement, highlighting the crucial role of vestibular signals in the integration of auditory information with head motion.

1.5.2 The Weighting of Self-Motion Signals in Spatial Updating.

Spatial updating in the auditory system refers to the brain's capability to maintain a current spatial representation by continuously updating the mental location of a previously heard auditory target as the head moves. As discussed above, this spatial updating following head motions is necessary to understanding our orientation and perception of movement.

Genzel et al. (2016), investigated the weighting of vestibular signals in spatial updating using a two-interval forced-choice (2-IFC) task. Participants were presented with two sound sources before and after a head rotation. They were then asked to compare the azimuthal position of the sound sources in a world-centered coordinate. By manipulating the motion conditions, the researchers were able to examine the contributions of different sensory signals. The study revealed that the gain on auditory updating was reduced during passive head rotations, where only vestibular signals indicated rotation. This suggests that both vestibular and proprioceptive/efference copy signals contribute to auditory spatial updating, even if one of the signals is set to zero. These findings highlight the complex interplay between different sensory signals in the process of auditory spatial updating.

Furthermore, the results of the study indicated that vestibular signals play a significant role in auditory spatial updating with vestibular weight of 0.85 in a weighted linear combination model and the vestibular signal being given a higher weight than the proprioceptive/efference copy signals. Their weighting model suggests that vestibular signals contribute to the updating of representation of sound-source position in a head-centered coordinate system (Genzel et al., 2016).

1.6 The Structure of the Current Study

1.6.1 Experimental Design Considerations

Drawing upon our research questions and the literature review presented in Chapter 1, several pivotal factors were taken into account when designing our task and experiments:

- In the research design to study Spatial Selective Auditory Attention, multiple sound sources are incorporated, where participants need to select the target among distractors.
- Non-spatial sound cues such as pitch, loudness and timbre can be really powerful in distinguishing specific sound sources, and therefore they were intentionally removed by using same-talker sound stimuli to assess the reliance on spatial cues for selective attention.
- The separation of the loudspeakers was selected via pilot testing to make the SSAA task difficult enough for performance to be reduced in the more challenging motion conditions while avoiding floor effects.
- To examine the necessity and sufficiency of vestibular and proprioceptive signals, the study varied the availability of these signals in each motion condition, which allowed for an assessment of the role they play in updating SSAA. In an Active motion condition participants voluntarily oscillated their head side to side; both vestibular and proprioceptive cues were available. In a Passive motion condition, participants were passively oscillated side to side in a swiveling chair; the availability of proprioceptive cues was minimized and vestibular cues maintained.
- Visual cues are commonly regarded as potent sources of information for self-motion perception. However, previous studies on dynamic SSAA have indicated that visual cues may not be essential. In our investigation, we included the manipulation of visual information to assess whether the findings would be consistent with prior research. The manipulation of the visual information involved changing the lighting of the testing room.

1.6.2 Objectives

The findings of the studies described above on spatial updating and localization, prompted us to examine the significance of vestibular, proprioceptive and visual signals and assess their necessity and sufficiency for updating SSAA. My study is, as far as I know, the first investigation of this topic.

The following is an outline of specific research questions to examine:

1) Are people able to update their allocation of attention (SSAA) to a specific location while moving their heads?
2) Is proprioceptive and efference copy information necessary for updating SSAA?
3) Is Vestibular information sufficient for updating SSAA?
4) How do visual cues affect updating of SSAA?

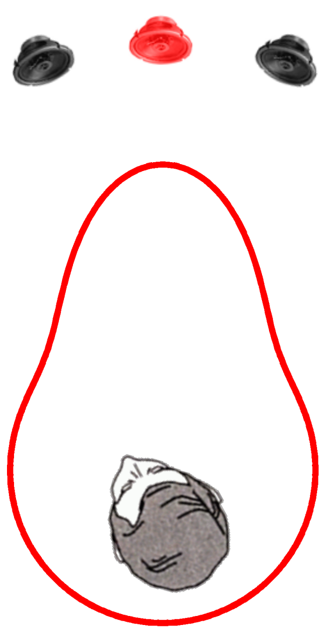

Figure 1.2

This figure captures how we want to look at the SSAA performance during head motion. The avocado shape represents the allocation of SSAA and the red loudspeaker is the selected target sound source. The goal is to update (in head-centred coordinates) the allocation of attention to the target during movement for best performance.

1.6.3 The Hypothesis and Expectations

Evidence from previous literature and results from a pilot study has led us to our main hypothesis about the ability to update Spatial Selective Auditory Attention in the availability of different sensorimotor cues (Vestibular and Proprioceptive signals). Previous studies have shown that proprioceptive signals seem unimportant for spatial updating and mostly vestibular signals are dominant in doing the tasks. Also, lighted or dark rooms have not shown any significant differences in updating SSAA in both static and dynamic conditions when the target is already localized. In general, the performance of maintaining SSAA to a stationary target should be the best and easiest in static conditions because there is no movement and therefore no need to update the SSAA. Moreover, it was anticipated that in Active conditions updating SSAA should be significantly better than in Passive conditions since all sensorimotor cues are available to be used. However, it was also expected that in Passive conditions, in which only vestibular information is available, the performance should be good enough to effectively update and then maintain the SSAA. According to what has been mentioned so far, in the current research project, we hypothesized that:

1) Although updating SSAA is challenging during head movements, people still are able to effectively update their allocation of SSAA.
2) Proprioceptive and efference copy signals are not necessary for updating and maintaining SSAA on a stationary target during head motions.
3) Vestibular signals are sufficient for updating and maintaining SSAA on a stationary target during head motions.
4) Visual cues do not assist with updating SSAA in both static and dynamic head conditions.

2 Methods

2.1 Overview and Procedure

This study consisted of a main experiment and two follow-up experiments designed to test some interpretations that derived from the result of the main experiment. A summary of all motion and lighting conditions used in these experiments is provided in Table 2.1 at the end of this chapter.

The participants' activities in the main experiment occurred in the following sequence:

Hearing screening:

Before going to the actual experiment, participants first went through a Hearing Screening test that ensured their hearing sensitivity was in the normal range. Hearing Screening was completed with an iPad-based audiometer (SHOEBOX Ltd.) at frequencies of 500, 1000, 2000, and 4000 Hz at 20 dB HL. The results were recorded as PASS (all thresholds ≤ 20 dB HL in both ears) or REFER.

Practice phase:

Following the successful completion of the hearing screening, participants engaged in task practice in each of the main experiment's three motion conditions (described below in Section 2.6) to establish familiarity. The practice phase for each motion condition consisted of 10 trials of the SSAA task described below in Section 2.5 and practicing the intended speed for voluntary head movement, during which participants acclimated themselves to the task requirements.

SSAA screening:

Following the practice trials, participants underwent Spatial Selective Auditory Attention (SSAA) screening to assess their ability to selectively attend to a designated target sound source. The screening consisted of 40 trials of the SSAA task described below in Section 2.5, conducted under lighted conditions and without self-motion. A performance criterion was set at a minimum of 75 percent accuracy, for eligibility to proceed with the subsequent experiments.

Main experiment blocks:

Once the participants passed the SSAA screening, they commenced the main experiments, which involved 9 counterbalanced blocks, with breaks provided between each block. Each block comprised 40 trials, taking approximately 10-12 minutes to complete. Participants received no feedback regarding their performance in any of the experiments.

2.2 Participants

Eighteen normally hearing listeners (13 females, 5 males; age range = 19 - 30) took part in the main experiment. Among these participants, nine individuals also participated in follow-up experiment (a), while six participants (4 females, 2 males; age range = 25-35, mean age = 28) were involved in follow-up experiment (b), with two of them having also taken part in the main experiment and experiment (a). They gave informed consent according to the ethical standards of Western's Research Ethics Boards. As a result of word-of-mouth recruitment and snowball sampling, participants were all students at Western University where this research was conducted. All participants were compensated $15/hour for their participation.

Participants needed to fulfill several criteria to be eligible for the study. Firstly, they had to be first-language English speakers between the ages of 18 and 35, capable of comprehending task instructions. All participants were required to undergo pure-tone audiometric screening to confirm normal hearing, defined as thresholds of 20 dB HL or lower at standard audiometric octave frequencies ranging from 250 to 8000 Hz. No one recruited failed the hearing screening. Additionally, they had to perform adequately in a preliminary test of SSAA as described below. Certain conditions would render participants ineligible for the study. These conditions included a history of vestibular/balance disorders or dizziness, as these individuals could be at risk of discomfort during tests involving head movement. Participants lacking flexibility in their neck and/or back, which could impede their ability to turn their heads left and right during the task, were also excluded. Furthermore, individuals who had difficulty sitting for extended periods of time were ineligible since the SSAA task required remaining in a seated position for two or more hours with intermittent breaks between blocks.

2.3 Apparatus and Materials:

All experiments were conducted within the 5.5 x 7 x 3.7 m^3 hemi-anechoic chamber located at the National Centre for Audiology (NCA) at Western University. Participants were seated in an open fronted shoulder-height wooden box that supported their backs and shoulders fixed on a swiveling chair with their heads positioned at the center of a circular 16-loudspeaker array. The floor between the active loudspeakers and the listening position was covered with sheets of sound-absorbing foam. The chair utilized for the experiments was not motorized, which minimized noise and limited auditory distractions unrelated to the task at hand. The wooden platform and chair were capable of silent rotation, facilitated by an assistant positioned behind the chair wearing headphones. This person was provided with continuous voice cues through the headphones, instructing the frequency and direction of head movement oscillation; right and left. Custom stimulus generation and experimental control software was implemented in MATLAB (The Mathworks) running on a Windows 7 PC. Digit stimuli were presented at a sampling rate of 44.1 kHz via a multi-channel audio interface (MOTU 24i/o), audio cross-point switcher (Yosemite 128128A, Sierra Video), power amplifiers (QSC CX168), and compact single-driver loudspeakers (A'Diva, Anthony Gallo Acoustics). The output level of each loudspeaker was individually adjusted so that the average sound level (LAeq) for a concatenation of all digit exemplars was 70 dBA measured at the listener's position using a Type 1 sound level meter (Bruel & Kjaer 2250) and free-field microphone (Bruel & Kjaer 4189). Head position data were monitored and captured using an electromagnetic Fastrak system (Polhemus) interfaced via an HTI3 module (Tucker-Davis Technologies) to an RX6 signal processor (Tucker-Davis Technologies) to enable real-time tracking of head position in space. To restrict neck movement during passive motion conditions, participants were provided with a curved wooden headrest containing foam padding, positioned at the lower part of their heads and leaving the ears unobstructed. This headrest was removed in conditions requiring the participant to make voluntary active head movements. Verbal responses were captured via a microphone, relayed to a set of headphones, and typed using a computer keyboard by the researcher outside the chamber room.

In Experiment (b) (see section 2.7.2), a second Fastrak motion sensor was utilized and attached to a short hand-held stick. For some lighting conditions, an acoustically transparent black fabric composed of 97% polyester and 3% spandex was employed to obscure the frontal loudspeakers. It was at 1.2 m distance in front of the listener. Green stars were projected onto the curtain using a Zenoplige Star Night Light 5-in-1 Projector to provide some optic flow information in the otherwise darkened room (see section 2.5.2)

Real-time tracking of head orientation was continuously employed, ensuring that stimuli were presented to the participants based on their specific head orientation as monitored by the electromagnetic tracker. This tracking method was employed for both the passive and active conditions (see figure 2.3 & 2.5). In the Static condition, the next stimulus was played after the participant's response was entered by the experimenter.

In the main experiment, only the three frontal loudspeakers were used (-22.5°, 0°, and +22.5° azimuth), but in follow-up experiment (a), the active loudspeakers varied from -67.5° to +-67.5° azimuth from trial to trial.

2.4 Sound Stimuli and SSAA Task

The stimuli used in the experiment consisted of three simultaneous sequences of 4 digits spoken by the same male talker. The sound stimuli and task were originally developed by Ruggles and Shinn-Cunningham for their (2011) study and they kindly provided their digit recordings to us. The digits included the integers from one to nine, but excluded the two-syllable digit "seven" and sequences were generated by randomly selecting a digit (one of eight) and a recorded exemplar (one of five for each digit) for each serial position in each stream, with the constraint that the same digit was never simultaneously presented in multiple streams. In the anechoic chamber, the three sequences consisting of four digits were played from three adjacent loudspeakers spaced 22.5° apart, one sequence from the target and two others from the distractors. The participant's task was to report verbally the 4-digit sequence produced by the central target loudspeaker. Each response digit was scored correct only if it matched the target digit at the same position in the sequence. Each digit within the sequence had a duration of 440 ms, resulting in a total sequence duration of 1760 ms. Importantly, all sequences were presented at the same intensity level of 70 dB SPL. Testing in each experimental condition involved one block consisting of 40 trials, such that 160 target digits were presented in each block.

2.5 The Main Experiment: SSAA Task Experimental Conditions

In order to evaluate the necessity and sufficiency of vestibular and/or proprioceptive (or efference copy) information in predictive updating of SSAA, the availability of these signals was manipulated through three distinct motion conditions. To evaluate the role of visual self-motion cues on SSAA updating, each motion condition was conducted under three lighting conditions detailed in section (2.6.2). In total, there were nine combinations of conditions (3 motions x 3 lightings) conducted in a counterbalanced order across participants using the double balanced Latin square method (Shuttleworth.,2009). The following sections 2.6.1 and 2.6.2 provide details of the motion and lighting conditions, respectively.

2.5.1 Motion Conditions:

The main experiment comprised three motion conditions, each serving a distinct purpose within the study.

Static: In this condition (illustrated in Fig. 2.1), participants maintained a stationary position centred at 0° azimuth with their head and body facing the three loudspeakers positioned at 1.7 m distance in front of them. There was no motion in this condition therefore no need for updating SSAA, rather, in this condition, participants focused on maintaining their SSAA towards the target utilizing available binaural and spectral cues to selectively attend to the middle target loudspeaker.

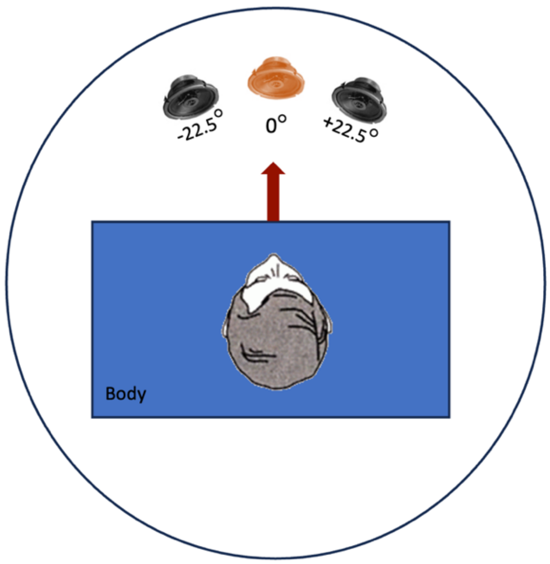

Figure 2.1

Static: An aerial view of the participant during the static condition. The target loudspeaker in shown in orange color. The sequences of digits were played from all three loudspeakers simultaneously as the participant fixated their head and body centered at 0°.

Active: In the Active condition (illustrated in Fig. 2.2), participants were asked to continuously move their heads from -45° to the left and +45° to the right (with the neck rest removed) at a fixed speed of 7 seconds for a full oscillation from left to right and back to left. The selection of a 7-second oscillation period served five purposes. Firstly, it was comfortable and easy to do for the participants. Secondly, it was within the frequency range of natural head movements (Zalewski, 2017) while also thirdly being sufficiently slow to allow for the updating of SSAA that aligns with Mondor and Zatorre's (1995) findings. Our task's duration of 3500 ms for a full sweep from left to right surpasses this range. Fourthly, a slow velocity helps to minimize the forces applied to the body (and therefore proprioceptive signals) in the Passive condition. Lastly, in this frequency of oscillation the vestibular system actually represents head oscillations. Furthermore, our pilot experiment substantiated our anticipation that participants would be able to achieve satisfactory SSAA performance with the 7-second oscillation period.

Prior to the main experiment, participants practiced the head movement speed and angle, guided by verbal cues of "left" and "right" presented from the left and right loudspeaker plus a noise that was panned back and forth too. Participants were instructed to replicate the practiced speed and angle of rotation during the task. The sound stimuli in this condition were triggered based on the head position; a trigger angle (0°, ±15°, or ±30°) was randomly chosen on each trial and the stimulus playback began when the head approached within 5 degrees of trigger angle from either direction. Figure 2.3 illustrates the 40 head sweeps recorded during the 1760-ms digit playback periods for one Active block. Digit 1 began playing at 0 ms and the thick dashed lines mark the onsets of digits 2, 3, and 4. The shaded regions show the 100-ms time periods in the middle of each digit over which we calculated the average head azimuth and velocity.

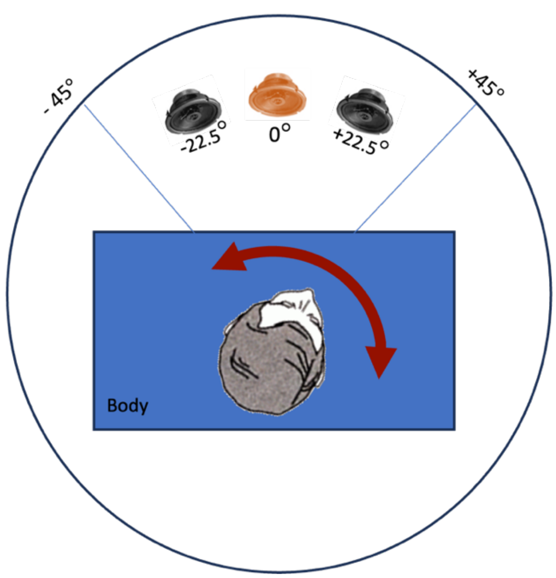

Figure 2.2

Active: An aerial view of the participant during the Active motion condition. The middle loudspeaker is the target, and in this example the subject has rotated their head to the right and is oriented to the right loudspeaker. The voluntary head oscillation to 45 degrees to the left and right was continuously tracked.

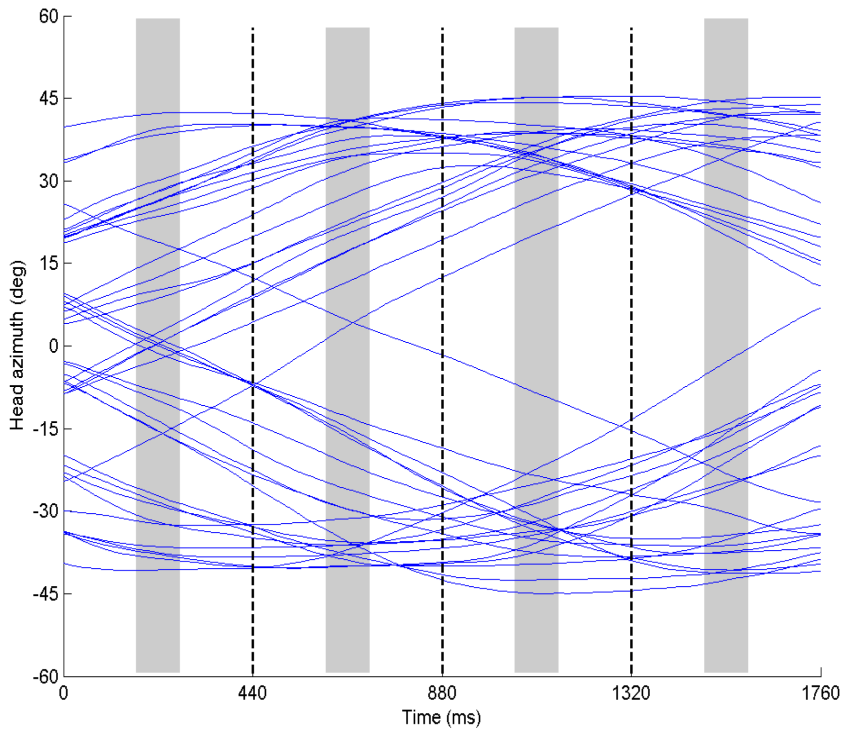

Figure 2.3

Active: Head tracker signal and velocity computation. Illustration of the 40 head sweep done by the participant's head-in-space over a ~ 45° range, as represented by the blue line. Digit 1 began playing at 0 ms and the thick dashed lines are the onsets of digits 2, 3, and 4. The shaded regions show the 100-ms time periods in the middle of each digit over which we calculated the average head azimuth and velocity.

Passive: In the Passive condition (illustrated in Fig. 2.4), the participants' chair was passively rotated to the left and right at 45° by a person positioned behind the chair. This assistant wore a headphone that played the same spatialized verbal cues of "left" and "right" with the same repetition period of 7 s as used during the head movement practice in the Active condition. This encouraged similarity in the rotational experiences between the Active and Passive conditions (see Fig 3.2). The Passive condition intentionally minimized the influence of proprioceptive/efference copy signals, providing participants with exposure to primarily vestibular cues to explore the sufficiency of vestibular signals in updating SSAA. Similarly to Figure 2.3, figure 2.5 illustrates the 40 head sweeps recorded during the digit playback for one Passive block.

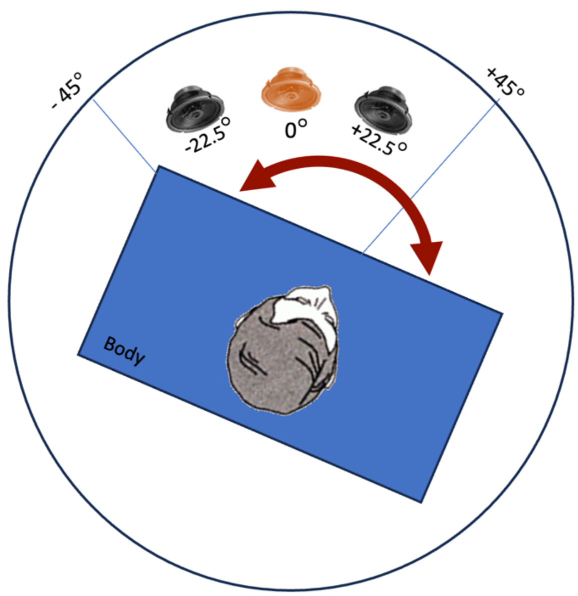

Figure 2.4

Passive: An aerial view of the participant during the Passive motion condition. The rotation of the box to 45 degrees to the left and right led to the whole-body oscillation and it was continuously tracked. The sound stimuli were randomly triggered based on the head position, occurring 5 degrees before or after the trigger angle.

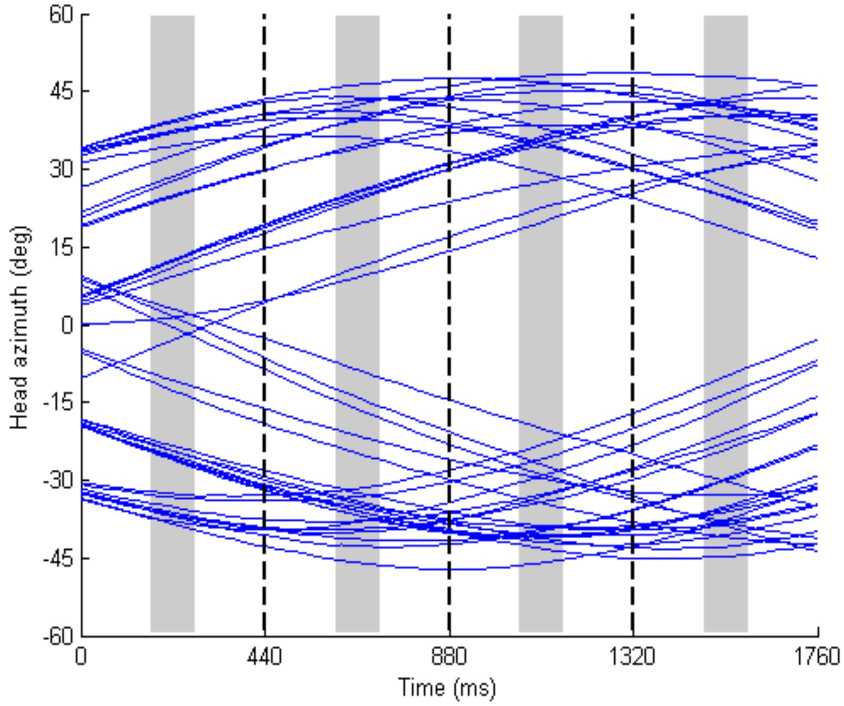

Figure 2.5

Passive: Head tracker signal and velocity computation. Illustration of a whole body (chair) sweep done by the person behind the chair over a ~ 45° range while the head was fixed on body, as represented by the blue line. Digit 1 began playing at 0 ms and the thick dashed lines are the onsets of digits 2, 3, and 4. The shaded regions show the 100-ms time periods in the middle of each digit over which we calculated the average head azimuth and velocity.

Figure 2.6

This figure shows the setting of the anechoic chamber room and the wooden chair during the Passive rotation, which was carried out by an assistant standing behind the chair.

2.5.2 Lighting Conditions:

To examine the role of visual self-motion information in SSAA updating, the lighting conditions were manipulated. The lighting conditions consisted of Lighted, Dark, and Curtain settings. In the Lighted condition, participants had a full visual perception of the loudspeakers and other equipment in the room during the task and had rich optic-flow cues to self-motion. In the Dark room, visual cues were totally absent throughout the task. In the Curtain condition, an acoustically transparent black fabric curtain was placed in front of the loudspeakers, obscuring them. Participants could not directly see the sound sources themselves, but small stars were projected onto the curtain to provide visual cues related to their rotation and environment, while only the loudspeakers remained concealed. The Curtain lighting condition aimed to assess the potential costs of concealing the loudspeakers while still having some visual cues about self-motion and orientation.

2.6 Follow-up Experiments

2.6.1 Experiment (a): Unpredictable Sound Location Test

To further understanding of participants' use of *predictive* updating of Spatial Selective Auditory Attention (SSAA) rather than relying solely on sound localization and then attention allocation (i.e., *reactive* updating), Experiment (a) was conducted. This experiment aimed to evaluate SSAA performance in a condition where updating was necessary but predictive updating was practically challenging or impossible because the target location was randomized trial to trial.

In this task, participants encountered a modified configuration of seven loudspeakers spanning from 67.5 degrees left to 67.5 degrees right, replacing the previous three-loudspeaker setup.

Within this arrangement, on each trial a random selection of an array of three loudspeakers generated the sequences of digits. Participants were instructed to stay stationary and to concentrate their attention specifically on the middle loudspeaker (which could be located at 0°, ±22.5°, or ±45°), designated as the target, and to accurately report the digit sequence originating from that loudspeaker. This experiment happened in the dark room to simplify the availability of the information.

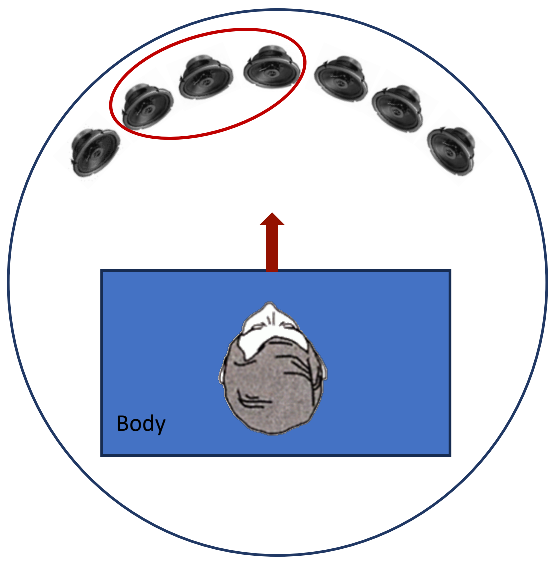

Figure 2.7

Experiment (a) Unpredictable: An aerial view of the participant during the unpredictable experiment in Static condition. The sequences of digits were presented from a random array of three adjacent loudspeakers among the seven spanning from 67.5 degrees left to 67.5 degrees right. The target was the middle loudspeaker of that specific array. The red oval shows an example selection of three loudspeakers with the target located at 22.5 degrees left.

2.6.2 Experiment (b): Cognitive Load of Attention Test

In order to investigate the potential cognitive load induced by continuous head movement in the Active condition and its impact on performance, Experiment (b) was conducted. The objective was to assess whether asking participants to engage in continuous head movements in the Active condition might have resulted in reduced performance compared to the Passive condition where they are not asked to do anything other than the SSAA task. The experimental design included the introduction of a new condition, referred to as Passive-Tap, where participants were required to hold a sensor-attached stick and to move it up and down in synchrony with the changes in motion direction while undergoing passive oscillations. An up and down motion was chosen because it was not pointing to the left and right, thus not interfering with the directional aspects of the SSAA task while still maintaining comfort similar to that of the head rotation. It was hypothesized that the addition of this cognitive load in the Passive-Tap condition would lead to decreased performance in the Spatial Selective Auditory Attention (SSAA) task, when compared to the passive condition without added motion. All participants completed the Static condition first, followed by a Passive block, Passive-Tap block, and Active block in a counterbalanced order (two of them participated in the main experiment and the rest were new to the study). To familiarize participants with the task, practice trials were conducted prior to the main experiment.

2.7 Summary of Experimental Motion Conditions

Table 2.1 provides a summary of the motion conditions used in the main experiment and in follow-up experiments (a) and (b). For each condition, the tables note the procedural details for that condition, the available head-motion cues, and the necessity of updating for optimal performance of the SSAA task.

Motion Conditions	Procedural details	Head-motion cues	Updating required
Static	Head and body orientation maintained at 0 azimuth	None	✗
Active	Head actively oscillated via neck motion while body faced 0 azimuth.	 Vestibular and Proprioceptive/efference copy	✓
Passive	Head held straight relative to the body. Head stabilized with foam neck rest to minimize neck movement. Head and body oscillated together passively by a research assistant behind the chair.	 Vestibular	✓
Unpredictable	Head and body orientation maintained at 0 azimuth	None	✓
Passive - Tap	Similar to the Passive motion condition, with the addition of holding a stick and moving it up and down.	 Vestibular	✓

Table 2.1

Summary of the experimental motion conditions.

3 Results

3.1 Overview of Analysis

The dataset obtained from all experiments encompassed several variables, including PC (Percent Correct), which represents the percentage of correct responses, PD (Percent Distractor) indicating the percentage of digits mistakenly reported from distractor loudspeakers, and PDHS (Percent Distractor Head Side), which is the proportion of head-side distractor errors (see Figure 3.1) that will be discussed further. Plots in this chapter are derived from two main sources of data: the correct answers (PC) and head-side distractor errors (PDHS) observed during all the experiments conducted with 18 participants. Additionally, figures derived from the head tracker data, which depict the observed head behavior, are plotted. The participants' responses were analyzed digit by digit rather than at a sequence level, such that order matters, leading to three types of mistakenly reported digits in addition to correct responses, as illustrated in Fig. 3.1. The first type of error arises when digits emitted from the distractor loudspeakers are incorrectly reported instead of those from the target loudspeaker. This category encompasses errors made from both left and right distractors. Within this category, there are two subtypes of errors. The first subtype occurs when the listener's head is oriented towards either the left or right loudspeaker, and digits from that specific loudspeaker in front of the head are mistakenly reported. This subtype is referred to as head-side distractor error. The second subtype encompasses distractor errors that do not correspond to the loudspeaker towards which the listener's head is oriented, and these are known as non-head-side distractor errors. Finally, there are non-distractor errors, which occur when mistakenly reported digits are not presented by either of the distractors.

Figure 3.1 *Types of errors illustration:*

This constructed response to the target sequence “2 3 8 1” includes one correct response and the three types of errors that can occur. The number '2' represents the correct response, '1' signifies a non-distractor error, the first '9' indicates a head-side distractor error due to the participant's head orientation towards the right loudspeaker. In this example, the right loudspeaker generates the digits '9'. Lastly, the final digit '9' represents a non-head-side distractor error, originating from the left loudspeaker.

Main experiment:

The main experiment encompassed responses from 9 blocks of trials in different motion and lighting conditions. Additionally, data on head and head/body movement were collected from the head tracker for the Active and Passive conditions, respectively. Furthermore, the head tracker recordings allowed us to compute information such as the velocity and azimuth of the head at the time each digit was presented. These comprehensive datasets were utilized to address the following research questions about the results:

1. How does the performance in the Static motion condition compare to the Dynamic motion conditions (Active and Passive) in terms of Percent Correct for reported digits? In other words, do participants perform comparable Spatial Selective Auditory Attention (SSAA) in dynamic conditions where their head is in motion, as compared to their SSAA performance in the Static condition?
2. Was there a difference between Active and Passive motion conditions on the SSAA performance? If not, we would conclude that proprioceptive signals are not necessary and that vestibular signals are sufficient for SSAA updating.
3. What was the effect of lighting conditions on the SSAA performance? If there was not an effect, we would conclude that having access to visual cues is not necessary for SSAA updating.
4. What does the relationship between reported digit errors, head position, and velocity reveal about the allocation of SSAA as a function of head position?

Experiment (a):

1. Based on PC scores, how well could participants update their SSAA to the middle target loudspeaker when predictive updating was impossible? This allowed us to estimate the separate contributions of predictive and reactive SSAA updating in the motion conditions.

Experiment (b):

1. Was the SSAA performance reduced in Passive–Tap conditions compared to the Passive condition? If so, it would suggest that there is a cognitive load of head movement during the Active motion condition.

3.2 Screening for SSAA Ability

Prior to the commencement of the actual experiment, participants underwent a preliminary screening block consisting of 40 trials in the Static-Light condition. This screening block aimed to ensure that individuals possessed the ability to effectively concentrate their selective auditory attention. In other words, it aimed to determine if they were proficient enough and reached 75% performance in the task to proceed to other experimental conditions. Notably, no one recruited for the main experiment failed the screening. Individuals who participated in Experiment (a) were all from the main experiment who had already passed the screening. Those who were recruited for Experiment (b) but not the main experiment went through the screening and two of them were excluded. In total among 25 people who got screened, two people failed at it and didn't proceed to the experiments.

3.3 Head Movements during Dynamic (Active and Passive) Conditions.

Head movement was continuously tracked and was recorded over each 1760-ms stimulus presentation (see Figures 2.3 & 2.5). The average azimuth and velocity were computed during the middle 100 ms of each digit in order to determine whether the participant's heads or oscillatory chair were moving in an approximate same motion range (azimuth deg) and speed (velocity deg/s) as instructed (Section 2.6.1).

In the Active condition, participants received instructions about the desired motion, but their head oscillation was not under our control. In the Passive condition, however, the voice cue through the headphones for the assistant oscillating the chair ensured the correct frequency of oscillation. By comparing the head behavior between the Passive and Active conditions, we can determine whether they align or differ from each other. Hence, the following plot [Figure 3.2] summarizes and compares the head movement behavior of each participant in the Active condition and the Passive condition with the oscillatory chair.

The distribution of digit-by-digit head azimuths and head movement velocities across all lighting conditions were converted to RMS values, which were plotted for each participant in order to compare their behaviour in the Active and Passive motion conditions [Figure 3.2]. The expected RMS head azimuth, considering a +/- 45-degree sinusoidal amplitude for both conditions was calculated as 31.8 degrees (45°/√2). An arbitrary criterion of +/- 33% was chosen to identify cases in which the RMS amplitude was markedly different from the intended value. It is evident that 16 subjects fell within the +/- 33% square box of the RMS value for both Active and Passive, while (in the Passive condition) two participants displayed over-rotated head movements (attributable to the assistant) that deviated from the 45-degree motion range. For the intended sinusoidal oscillation at 1/7 Hz and a 45-degree amplitude, the expected RMS velocity was calculated as 28.6 degrees per second (45 × 2π × 1/7 × √2). In the right graph, it can be observed that five individuals moved too quickly in the Active conditions, as their values were above the +/- 33% square box.

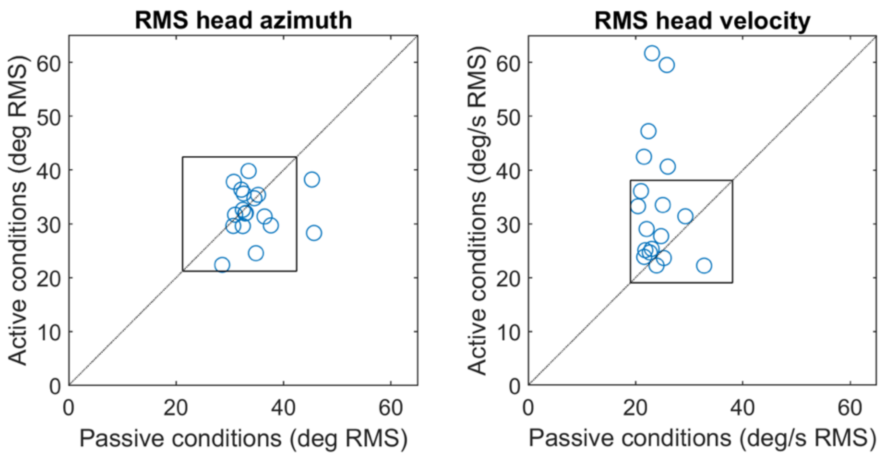

Figure 3.2 *Comparison of RMS head Azimuth (deg, left panel) and head Velocity (deg/s, right panel) during Active and Passive movements for each of 18 participants.*

The horizontal axis shows the RMS head azimuth (left) and velocity (right) during the Passive conditions. The vertical axis shows the RMS head azimuth (left) and velocity (right) during the Active conditions. The square boxes represent +/-33% around the intended RMS values.

3.4 Main Experiment: SSAA Performance Comparison between Different Movement and Lighting Conditions.

The data obtained from the main experiment is presented in Figure 3.3, displaying the percentage of correctly reported digits (PC) across various motion and lighting conditions for each of the 18 subjects. The initial three conditions correspond to Static conditions in Light, Curtain, and Dark. Notably, the majority of participants demonstrated a high level of performance with a mean of 88.27% PC value in these static conditions. In the case of the two dynamic motion conditions (Active and Passive), a decline in performance relative to Static is observed; however, the decline from Active to Passive condition appears to be less pronounced. Notably, two participants (L274 and L279) exhibited somewhat distinctive responses. Specifically, L274 displayed a more significant decline in performance from the Static to Active and Passive conditions compared to other participants. Conversely, L279 exhibited only a modest decline.

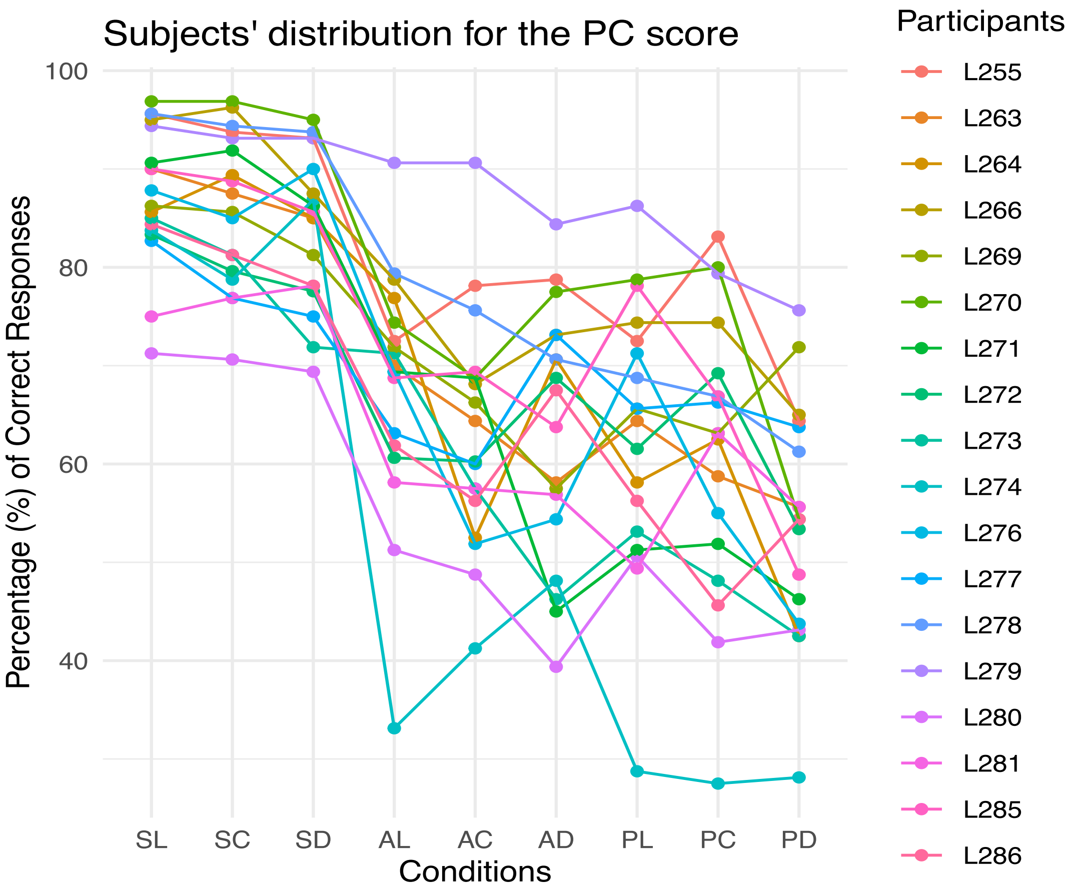

Figure 3.3 *Percent Correct scores for all 18 participants across all motion and lighting conditions.*

The x-axis represents the different conditions, ranging from (SL) to (PD). The conditions are coded as: S, Static; A, Active; P, Passive; L, Light; C, Curtain; D, Dark.
The y-axis represents the percentage of correct digits reported from the target loudspeaker (PC).

Although there were individual differences in overall performance, the pattern across conditions was similar for most participants, so the mean is fairly representative of the group. Nonetheless, the two outlier participants did not exhibit a reverse trend compared to the others. The mean of data collected from the main experiment is depicted below in [Figure 3.4]. The plot presents the mean SSAA performance (PC) over the 18 participants in each of the nine combinations of motion conditions and lighting conditions.

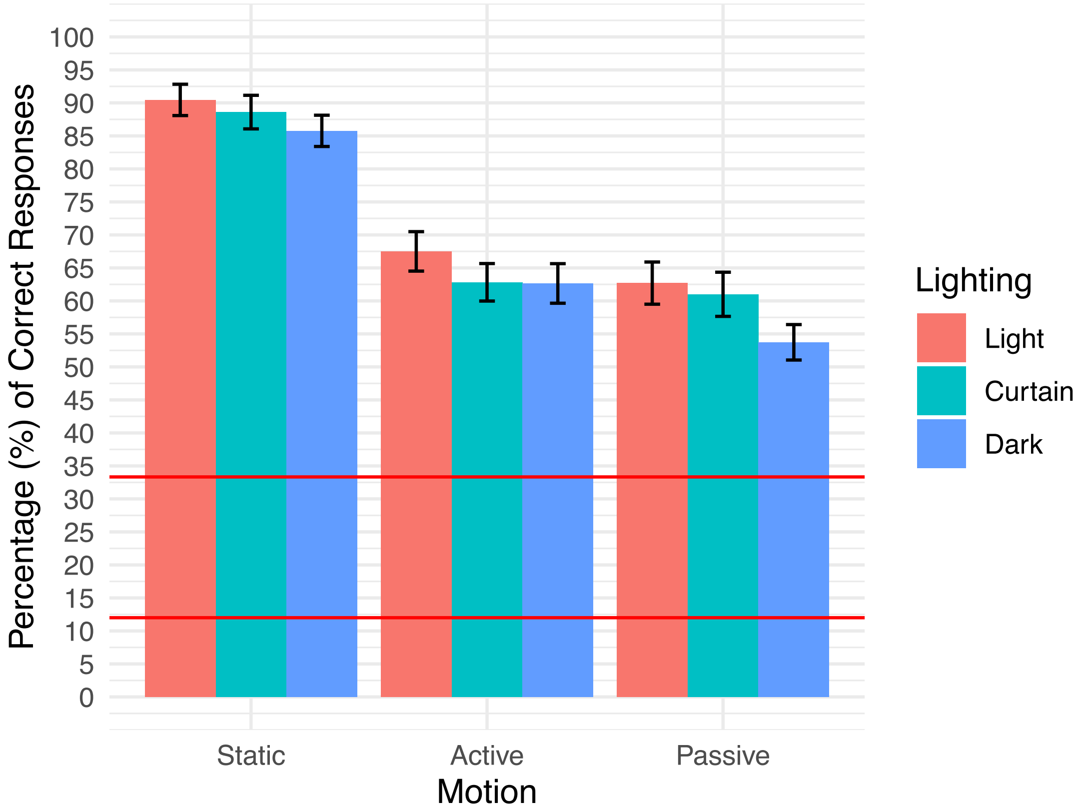

Figure 3.4 *Mean Correct-Response Rate:*

The Y-axis shows the across-participant mean percentage digits correctly reported from the target loudspeaker (PC), and X-axis is lighting conditions grouped within motion conditions. The two red lines represent the source-guessing and digit-guessing chance percentages (33.3% and 12.5%, respectively). Error bars are standard error of mean.

The results indicate that the Static motion condition consistently yielded the highest SSAA performance across all conditions, in contrast to the two dynamic motion conditions (Active and Passive). Furthermore, the performance during Active and Passive motion conditions appears to be quite similar, without either condition surpassing the performance of the Static condition. Regarding lighting conditions, it appears that participants demonstrate slightly improved performance when visual cues are available (Light and Curtain), as compared to the Dark room condition.

The data demonstrate that participants consistently (with the exception of L274) performed significantly above the source-guessing (33.3%) and digit-guessing (12.5%) chance levels in all conditions. We used one sample t-tests to compare participants' mean PC scores within each of nine conditions against source-guessing chance level (33.3%). The result was significant in all conditions (all nine $t(17) \geq 7.592$, all $p \leq 7.377e-07$, 95% CI [min>48.05, max<95.46]). These findings highlight the robustness of participants' performance in the main experiment, indicating updating of SSAA during Active and Passive motion conditions.

To determine the effects of motion and lighting conditions on performance, we conducted a two-way repeated measures analysis of variance (ANOVA), with RAU-transformed percent-correct scores (PC) as the dependent variable. This yielded significant main effects for motion type ($F(2, 34) = 146.4$, $p < 0.001$) and lighting ($F(2, 34) = 14.22$, $p < .001$), indicating significant differences among the levels of these factors. The interaction effect between motion type and lighting was not significant ($F(4, 68) = 2.031$, $p = 0.0998$), suggesting that the effects of motion type and lighting on PC were independent.

Subsequent post hoc tests on the main effect of motion (with Bonferroni correction) revealed significant performance differences between all pairs of motion types (Static>Active, Static>Passive, Active>Passive), indicating varying levels of performance depending on the type of motion (collapsed across lighting). Specifically, participants exhibited significantly higher performance during the Static condition compared to both the Active condition (mean difference

= 23.954, 95% CI [19.380, 28.529], $p < .001$) and the Passive condition (mean difference = 29.135, 95% CI [24.561, 33.709], $p < .001$). Furthermore, a significant difference in performance was observed between the Active and Passive conditions, with the Active condition yielding higher performance than the Passive condition (mean difference = 5.181, 95% CI [0.606, 9.755], $p = 0.022$). To interpret the absolute differences between conditions, Cohen's d statistic was calculated: the effect size (Cohen's d) between the Static and Active conditions was 2.09, indicating a very strong effect, while the effect size between the Active and Passive conditions was 0.39, which is considered a small effect.

Similarly, post hoc tests for lighting conditions indicated significant performance differences between all pairs (Light>Curtain, Light>Dark, Curtain>Dark), demonstrating the impact of lighting (collapsed across motion) on performance. Participants achieved significantly higher performance in the Light condition compared to both the Curtain condition (mean difference = −2.743, 95% CI [−5.663, 0.177], $p < 0.024$) and the Dark condition (mean difference = −6.172, 95% CI [−9.092, −3.252], $p < .001$). The Cohen’s d effect size for the difference between Light and Curtain was calculated as 0.07 or very small. Additionally, the Curtain condition yielded significantly higher performance than the Dark condition (mean difference = −3.429, 95% CI [−6.350, −0.509], $p = 0.011$) with a small d effect size calculated as 0.13.

3.4.1 Analysis of Errors in Updating SSAA

In order to analyze the mistakes as a function of digit-by-digit head orientation, the Percent of Distractor Head-Side errors (PDHS) value was calculated by dividing the number of head-side distractor errors by the total number of distractor errors. Only digits for which the head was turned left or right by at least 15° were included in the PDHS calculation. If PDHS exceeded 50 percent, it indicated a higher occurrence of head-side distractor errors compared to non-head-side errors. Figure 3.5 below shows the percentage of distractor digits that were reported mistakenly from the head-side distractor loudspeaker (PDHS) during Active and Passive motion conditions.

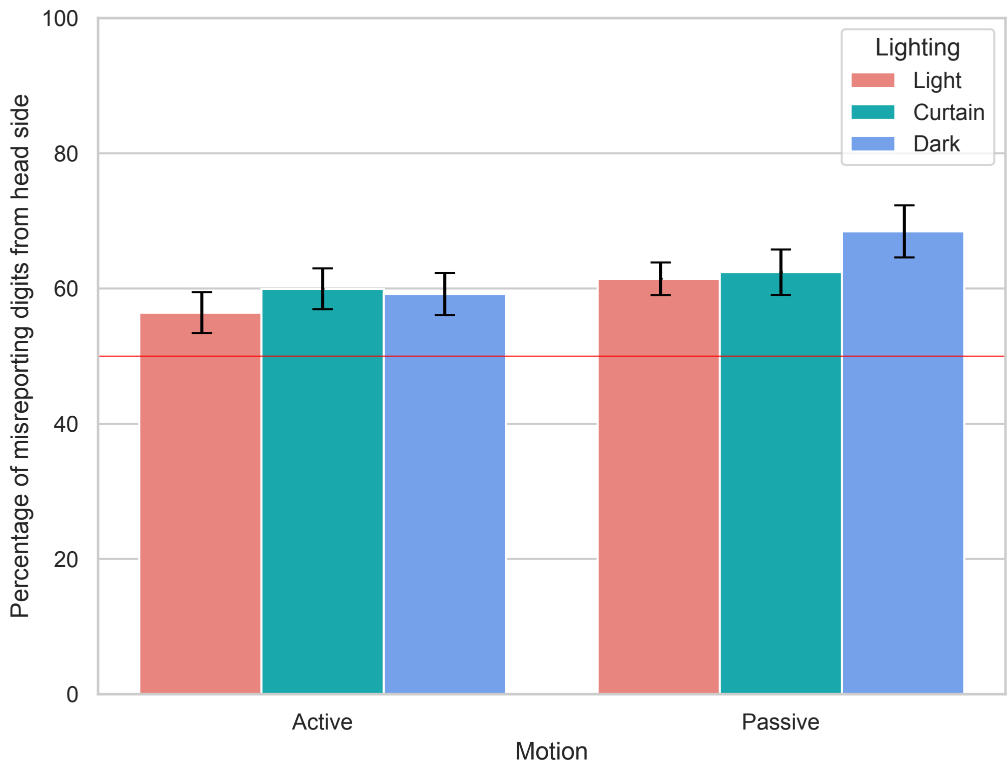

Figure 3.5 *PDHS value, the percentage of head-side distractor errors among all distractor errors while the head in Active and Passive motion*:

The y-axis is PDHS. The x-axis is categorized into Active and Passive motion conditions and also lighting conditions. The red line represents 50 percent PDHS, which would occur if distractor errors were equally common on the head-sides and non-head-sides. Error bars are standard error of mean.

To assess the significance of difference between the misreported head-side distractor digits and non-head-side misreported digits, six one-sample t-tests were conducted comparing the mean

PDHS value among all participants in each of the six dynamic motion and lighting conditions to a 50 percent value. The t-test results between the mean of PDHS scores across the dynamic motion conditions and 50 percent, revealed a significant difference between the observed mean PDHS in 3 out of 6 motion-lighting conditions after the Bonferroni correction (Table 3.1). This indicates that participants' PDHS values in those conditions significantly exceeded the bias-free value of 50 percent, and thus that head-side distractor errors were significantly more common than non-head-side distractor errors.

Conditions	t-value	p-value
Active Light	1.79	0.090
Active Curtain	3.50	0.003 < 0.05/6 *
Active Dark	2.35	0.030
Passive Light	4.09	0.00076 < 0.05/6 *
Passive Curtain	2.71	0.014
Passive Dark	4.31	0.00047 < 0.05/6 *

Table 3.1 Test statistic of the mean PDHS for dynamic motion conditions.

Also, when comparing the PDHS values between the participant's mean across lightings in Active and Passive, a paired t-test revealed no significant difference between the motion conditions ($t(17) = 2.0351$, $p = 0.05774$, 95% CI [-0.165, 91.647]). It is important to note that the observed mean difference in PDHS between Active and Passive conditions was minimal (mean difference = 4.499) because the frontal bias occurred similarly in both the Active and Passive conditions.

3.4.2 Spatial-Attentional Map

To further illustrate the incorrectly reported digits trend, the frequency of correct responses, head-side distractor errors and non-head-side distractor errors as a function of head orientation are depicted in [Figure 3.6] pooled across all the participants through the Static and Dynamic motion conditions separately. The orientation of head can be toward the left or right loudspeaker (distractors), and also toward the front loudspeaker (target) while head was rotating to the left and right. Because each participant made a relative small number of distractor errors, the pooled results shows trends not visible in individual participants' data. We looked for trends in the pooled data, but did not perform statistical analysis.

The upper plot shows the errors reported during the Static condition when the head was always oriented to the front and remained still. The lower plots indicate the errors during dynamic motion conditions (Active and Passive) when the head was oriented to either the left (< -15), right (> +15), or front (-15 > > +15). In the lower panel, during both dynamic motion conditions, when the head is oriented to the left, the misreported digits occurred more frequently from the left loudspeaker compared to the right loudspeaker (pink bars). However, the correctly reported digits (green bar) were still more frequent than both, indicating the ongoing updating of SSAA. Similarly, in the lower right panel, it is evident that when the head was oriented towards the right loudspeaker during head motion, the errors were predominantly reported from the right loudspeaker. In contrast, the middle lower panel shows that when the head was oriented towards the frontal loudspeaker during Active and Passive motion, the misreported digits were equally distributed between the right and left distractor loudspeakers. This last pattern is also observed in the upper plot, which represents the Static condition when the head was always stationary towards the front. In both cases, the misreported digits (errors) were equally reported from both distractors without any bias towards either of them. Hence, consistent with the PDHS analysis above, the results in this figure indicate that more distractor digits were reported from the loudspeaker on the head turn side, showing a head-centred frontal bias phenomenon for the spatial attention. This result is aligned with Macpherson et al., (2019) findings.

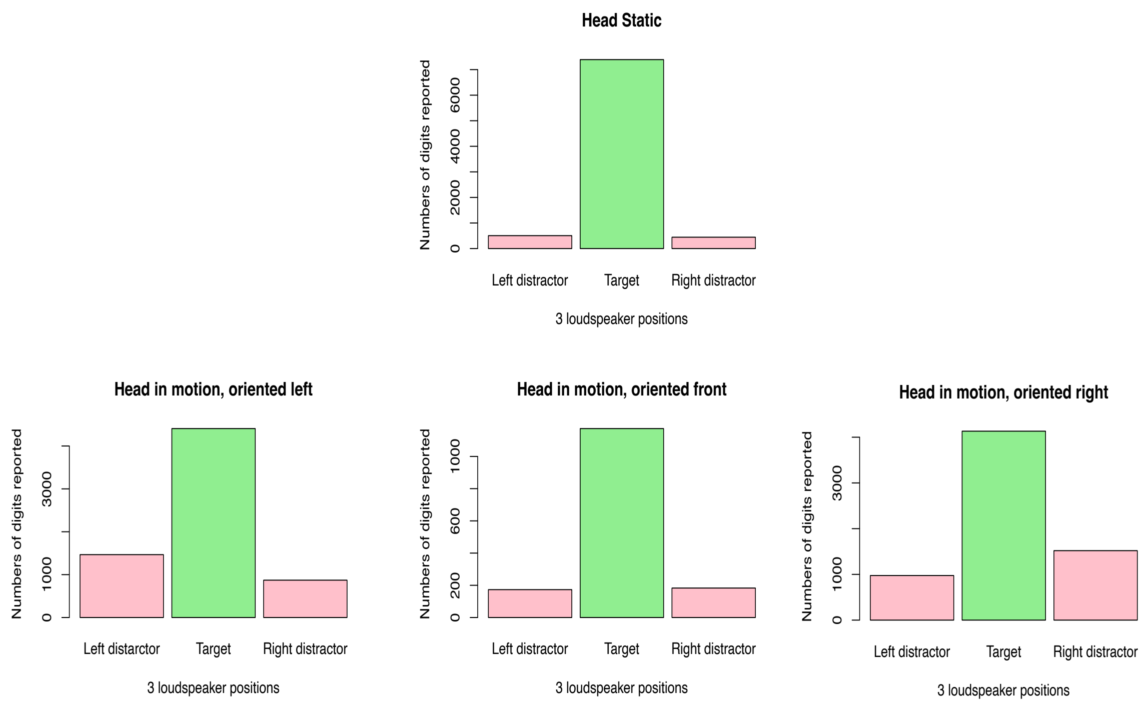

Figure 3.6 *Spatial – attentional Map for each head orientation during dynamic and static motion conditions:*

The total number of correctly reported digits and incorrectly reported digits reported from the distractor loudspeakers are pooled across listeners. The right-distractor and left-distractor errors are the head-side errors when the head is turned right (azimuth > +15) or left (azimuth < -15), respectively. The x-axis represents the position of the three loudspeakers (-22.5°, 0°, +22.5°). The y-axis displays the number of reported digits, categorized as incorrect (pink) or correct (green), across all blocks of experiments.

In Figure 3.7 below, the reported digits are analyzed based on additional factors, including Active and Passive motion conditions, head orientation, head movement direction, and each digit's serial position in a sequence. Notably, there is a consistent bias in distractor errors towards the head-side loudspeaker for all four reported digits, as seen similarly in the pooled data of Figure 3.6. It is also evident that the direction of head movement towards the left or right does not impact the observation of frontal bias in neither of the dynamic motion conditions. If updating were slow, we would expect the loudspeaker on the side the head is moving away from to be more frequently mistaken, however, it is not the case in the frequency of responses reported depicted in Fig 3.7.

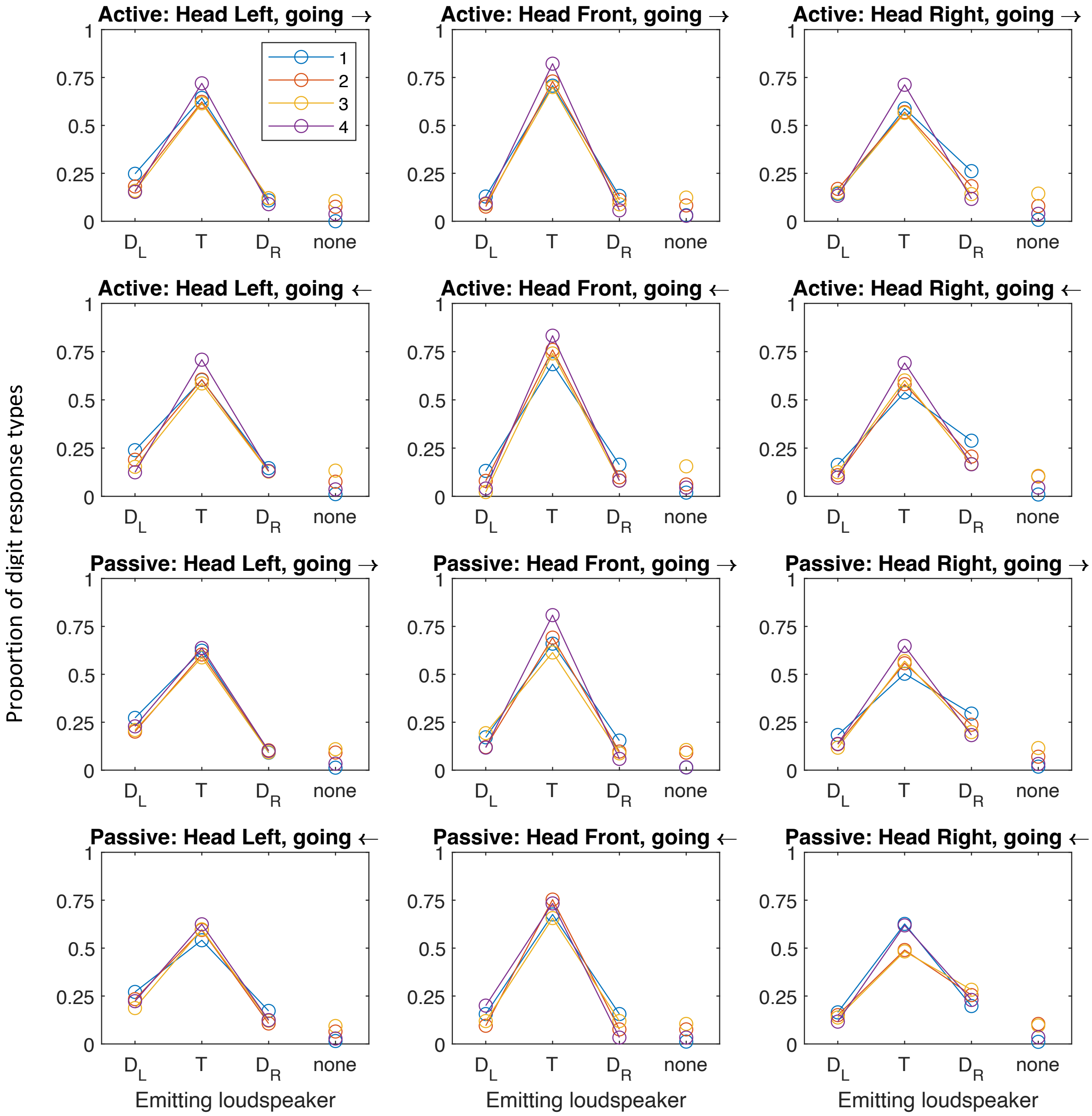

Figure 3.7 *Spatial-attentional map across digits for each head orientation during dynamic motion conditions:*

This figure is depicting the proportion of response types relative to the head position for both Active and Passive motion conditions averaged across all participants and lighting conditions. The y-axis represents the proportion of responses for each position in the digit sequence (1-4) that corresponded to each loudspeaker (or none). The x-axis displays the three loudspeakers from which reported digits were emitted. 'None' indicates that the reported digit was not presented or emitted from any of the loudspeakers. Panel titles indicate the head orientation and direction of motion.

3.5 Follow-up Experiments

3.5.1 Experiment (a) Result: Performance Comparison between Unpredictable and Passive Conditions.

The aim of experiment (a) was to investigate whether reactive updating is achievable under the condition of Unpredictable sound source location. The results obtained of the follow-up experiment (a) are presented in the right-most bar of Figure 3.8, which includes the data from the main experiment for reference. In the Unpredictable motion condition in the dark room, where participants needed to update their allocation of spatial attention to the target due to an unexpected change in its location within a dark environment, the performance was comparatively lower than that in the Passive-Dark condition which was the lowest performance among all other conditions itself [Figure 3.8].

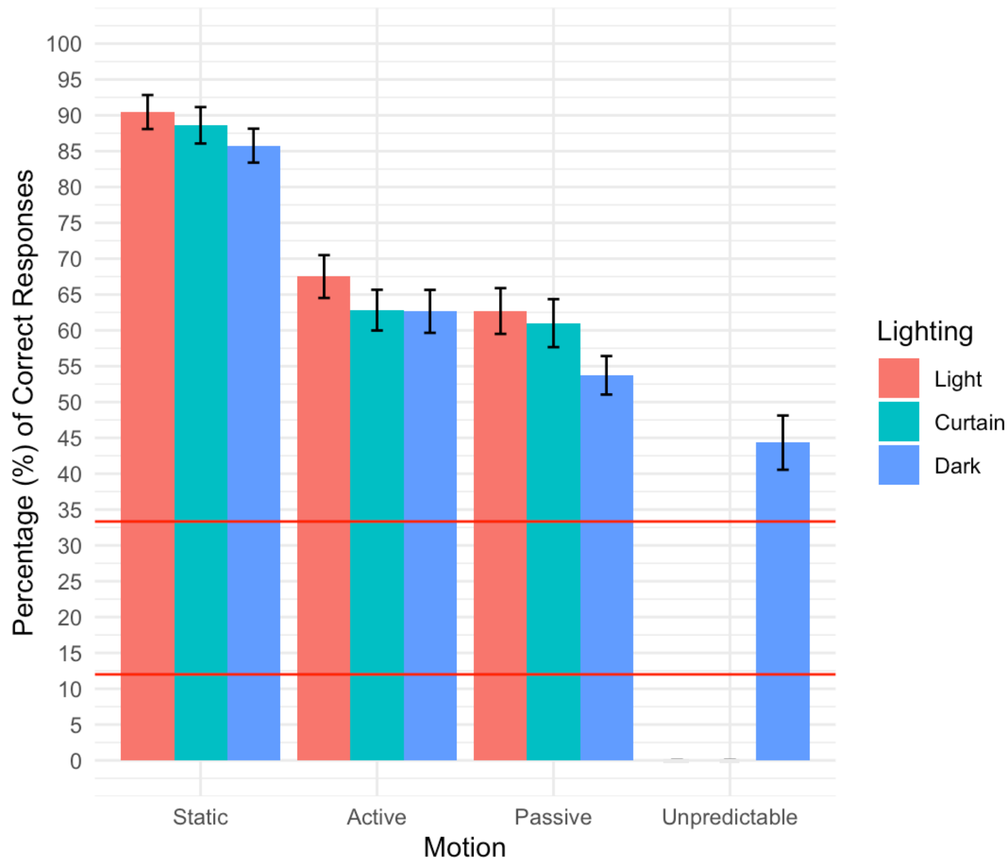

Figure 3.8 *The comparison of the Unpredictable experiment in Static self-motion and Passive Dark condition.*

This graph shows the main experiment data along with the Unpredictable bar to compare this condition with the rest of motion, lighting conditions.

To evaluate whether SSAA performance in the Unpredictable condition was significantly above chance, a one-sample t-test was performed between the Unpredictable condition PC scores and the source-guessing chance level of SSAA performance (33%). The t-test results revealed a statistically significant difference ($t(8) = 2.9092$, $p = 0.01961$) between the observed mean performance ($M = 44.33333$) and the chance level of 33%. Thus, the mean performance in the Unpredictable condition exceeded the chance level, indicating that participants' performance was better than random source-guessing. Also, a paired t-test analysis was performed to compare the performance between the Unpredictable-Dark condition and the Passive-Dark condition (represented by the lowest bar) using data from the nine participants, who provided data in both conditions. The analysis revealed a significant difference between the Unpredictable and Passive-Dark conditions ($t(8) = -2.7893$, $p = 0.02358$), with a lower mean performance observed in the Unpredictable condition (mean difference = -11.11149, 95% CI [-20.297655, -1.925318]) and with a large Cohen's d effect size calculated as 0.82.

To examine the spatio-temporal distribution of responses similarly to those from the main experiment (Figure 3.7), Figure 3.9 presents an analysis of reported digits based on head orientation relative to the changing target azimuth and each specific digit. The PC score derived from the Unpredictable condition encompasses all positions of the target loudspeaker (-45°, -22.5°, 0°, +22.5°, +45°). Notably, the figure highlights that the highest updating scores are achieved when the group of three loudspeakers is situated in front of the static head, with the target (middle loudspeaker) aligned with the Static condition location as in the main experiment. Conversely, when the target is oriented towards other positions, the correct reporting of digits diminishes. Furthermore, in terms of the order of the four digits, it is evident that the performance of the SSAA improves progressively over the course of the sequence. Specifically, the fourth digit is reported correctly more frequently from the target loudspeaker, indicating an enhanced updating of SSAA to capture accurate information from that particular source. Moreover, for targets at +/-22.5°, the responses for digits 1 & 2 looks like source-guessing (approx. equal proportions between the loudspeakers), and for +45 there is a more frontal bias that observed in Active and Passive conditions; head-side distractor responses are approx. equal to correct ones.

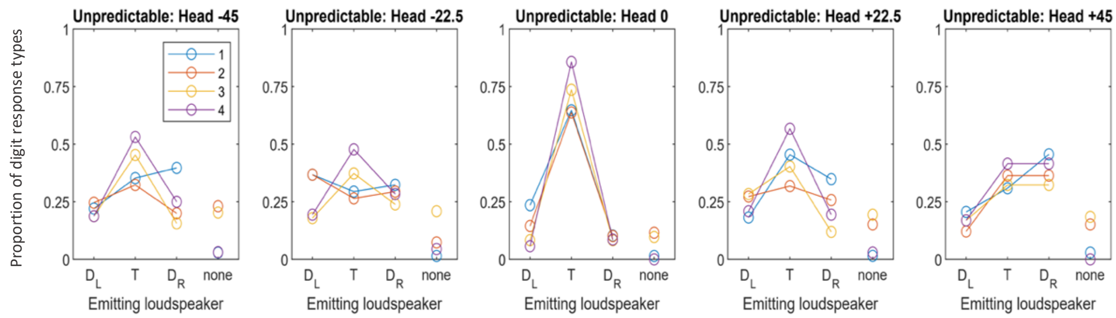

Figure 3.9 *Spatial-attention map across digits for each head orientation during Unpredictable condition.*

The y-axis represents the proportion of responses for each position in the digit sequence (1-4) that corresponded to each loudspeaker (or none). The x axis shows the three loudspeakers from which the reported digits were generated from. The panel titles indicate the head orientation relative to the randomly changing target loudspeaker azimuth.

3.5.2 Experiment (b) Result: Performance Comparison between Passive and Passive-Tap Conditions.

The objective of experiment (b) was to examine the potential impact and reduction in SSAA performance within the Passive-Tap condition, where participants were required to hold and move a stick during doing the task in the passive condition, thereby introducing an additional cognitive task. The data obtained from a sample of six participants are shown in Figure 3.10 and did not indicate any visible trend of reduced performance in the Passive-Tap compared to the Passive condition, however, the trend across Static, Active and Passive motion conditions is similar to that of the participants in the main data plot (Figure 3.3). Although we acknowledge the limited sample size in this experiment, the sample estimate indicates a mean difference of -2.71 between the Passive and Passive Tap conditions and a paired t-test between participants' PC scores in the Passive and Passive Tap conditions did not reveal a significant difference ($t(5) = -1.4419$, $p = 0.2089$). The null hypothesis, suggesting that the true mean difference is equal to 0, cannot be rejected. The 95% confidence interval for the mean difference ranged from -7.541349 to 2.121349.

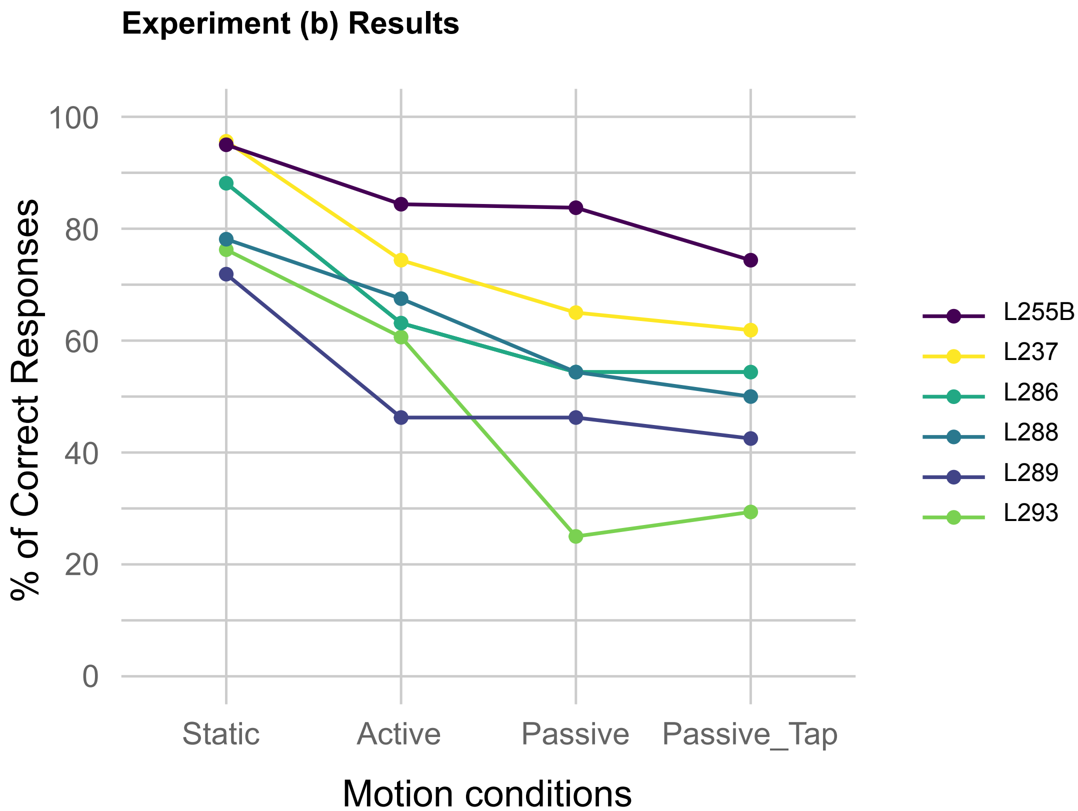

Figure 3.10 *The PC scores of different motion conditions in experiment (b)* (Passive-Tap) for each subject.

The y axis shows the percentage of correct responses reported (PC), and the x axis is the motion conditions (Static, Active, Passive, and Passive-Tap) of this experiment, all in the Dark lighting.

4 Discussion

In this chapter, the results described in Chapter 3 will be used to answer the research questions posed in Chapter1 and the questions arising from the results and the graphs that are listed in the overview of the analysis, section 3.1.

4.1 Does Spatial Selective Auditory Attention (SSAA) Performance Differ between Static Condition and Dynamic Motion Conditions?

4.1.1 Can Individuals Maintain and Update SSAA at a Specific Location while Moving their Heads?

The primary interest of this study was to examine how our auditory attention selectively focuses on a desired sound source despite the presence of other distracting sounds when our head orientation is not fixed. We sought to understand how we use non-auditory self-motion information to update SSAA to the desired sound source location based on the changing binaural cues and head-referenced source location.

This was measured by having the participants move Actively (neck on body) and Passively (whole-body) while they were engaging in an auditory attentional task. We hypothesized that although updating SSAA in motion can be challenging, listeners should be able to perform above chance level, even if not as well as in the Static condition.

The results from the main experiment demonstrate the effective cognitive ability of attention when sound stimuli were competing with each other during constant changes in head orientation. Since in the Static condition, there was no updating required and focusing attention selectively had no challenge, the performance (PC) at this condition is a baseline to compare dynamic motion conditions to see how well the SSAA could be predictively updated like how listeners were focusing their SSAA in the static condition.

Analysis of the digits reported correctly (mean PC = 61.73%) from the target loudspeaker while participants were in dynamic motion conditions (Active and Passive), indicates a significantly superior performance compared to the chance levels for source-guessing (33%) or digit-guessing (12.5%). This finding suggests that updating SSAA during self-motion was possible and exceeded the chance level significantly, yet not perfect. Specifically, participants were maintaining the representation of the target sound source throughout the task when both engaging in Active and Passive motion conditions and predictively updating the involved auditory, attentional, and localization systems (which all are necessary for understanding the head movement behavior and bringing SSAA into action) to align with this maintained representation that was achieved by incorporating cues derived about head movements. Previous studies (Macpherson et al., 2019; Macpherson & Ransom, 2018) have discovered that the focus of SSAA was decreased during head motion, but focusing SSAA was not impossible rather, it was difficult. Here we are suggesting that even though head motion exposes a challenge for spatial selective attention, the integration of head-motion-derived signals (vestibular and proprioceptive/efference copy) with auditory localization cues centered in the head are adequately useful to compensate for the self-motion challenge.

If listeners had been unable to focus their attention on the target loudspeaker while in motion, they would have had to report the digits randomly based on source- or digit-guessing. Rather, we see that there is a consistent trend of correct responses in dynamic motion conditions, not only within all condition blocks but also across almost all individuals [see Figures 3.3 & 3.4].

4.1.2 What Does the Allocation of Attention Look Like during Active and Passive Head Motions?

Additional evidence about the accuracy of predictive SSAA updating during self-motion is derived from the observed spatial frontal bias phenomenon in our results and previous studies (Lerens & Renier, 2014; Macpherson & Ransom, 2018; Macpherson et al., 2019). Our findings indicate that although there was a frontal bias (more incorrect responses from the head-side distractor loudspeaker), this bias was not as prominent as the updating of SSAA towards the target representation during self-motion, and could be interpreted as a slight under-compensation of the head motion. This is evident from the fact that we did not observe more misreported digits from the head-side oriented loudspeaker compared to those from the target loudspeaker. In fact, the opposite was true [see Figures 3.6 and 3.7]. While the digits from the head-side oriented loudspeaker were reported more frequently than those from the other side, they were still significantly fewer in number compared to the digits reported from the target loudspeaker.

The observed frontal bias in misreporting digits from the head-side loudspeaker emphasizes that listeners are not merely attending to the louder loudspeaker, which in the ear closest to the loudspeakers would be the opposite of head-side loudspeaker. Contrarily, the frontal bias suggests that listeners spatially attend to the binaural cues provided by the head-side loudspeaker while in motion, not the louder ones.

In addition to the observation of frontal bias, the spatial-attentional map across digits (see Figure 3.7) provides evidence that the direction of velocity (the side towards which the head is moving), does not impact the incorrect reporting of digits, indicating that the updating process is not sluggish (Macpherson & Ransom, 2018). If it were, then when moving from right to left, we would expect to see a right-side distractor bias persisting in the frontal map, and conversely, a left-side bias when moving from left to right. However, this is not the case (see Figure 3.7). We don't observe this persistent bias opposite the direction of moving, indicating that the updating process is quick enough to compensate for the head motion, at least at the speeds used in this study (target RMS velocity of ~29 deg/s). With regards to the required time for switching spatial auditory attention, Mondor & Zatorre, (1995) reported the time necessary to fully engage attention at a specific spatial position lay between 150 ms and 600 ms, regardless of physical separation of between attentional fixation. Even 600 ms is "quick" compared to the 3500-ms duration of side-to-side head movement in this study.

Although distractor-loudspeaker errors were unbiased when the head was oriented forward during the dynamic Active and Passive conditions, they were more numerous than in the Static condition (Figs. 3.6 and 3.7). This suggests that the focus of SSAA was somewhat widened during motion.

Figure 4.1

This figure shows the likely allocation of SSAA during head motion based on our results and the derived spatial-attentional maps. In this example, the participant's head is oriented to the left distractor loudspeaker while the allocation of attention is appropriately updated toward the target loudspeaker, yet incompletely and with a frontal bias toward where the head is oriented to (left). The allocation of attention is therefore depicted between the two loudspeakers.

4.2 How are Proprioception/Efference Copy and Vestibular Cues Weighted for SSAA Updating during Self-Motion?

The purpose of having different motion conditions in this study was to manipulate the availability of vestibular and proprioceptive/efference copy signals for the purpose of characterizing the relative weighting of those signals. In the Passive condition, the neck was kept still by the neck rest, minimizing proprioceptive signals about head position. If the performance in the Passive motion condition is comparable to the Active motion condition, where all signals are available, it can be argued that proprioceptive/efference copy signals are not necessary for updating SSAA.

To assess the sufficiency of vestibular information, we examined the performance in the Passive motion condition. The results depicted in Fig. 3.4 indicate that SSAA performance in the Passive motion conditions (mean PC 59%) was nearly as good as in the Active motion conditions (mean PC 64%). Performance in Passive was significantly above chance and despite the statistically significant difference observed between the Active and Passive conditions, the effect size (Cohen's d) was small (mentioned in 3.4), indicating only a marginal improvement of the Active condition over the Passive condition. These findings suggest that vestibular cues are sufficient for updating SSAA during self-motion and that proprioceptive/efference copy cues are not necessary. In conclusion, we have shown that the vestibular information is highly weighted in updating SSAA and that the contribution of proprioceptive/efference copy signals appear to be less pronounced. This finding is consistent with previous studies investigating front/back localization accuracy (Kim et al., 2013) and auditory spatial updating (Genzel et al., 2016), that have reported a similar pattern, with a greater weighting and dominance of vestibular signals in front/back localization and auditory spatial updating tasks, respectively. Also, the findings align with another study that investigated how human observers combine proprioceptive and vestibular cues during full-stride curvilinear walking. The study demonstrated that during passive self-motion, where participants experienced vestibular signals without active movement, vestibular signals by themselves were adequate for accurate egocentric updating, while proprioceptive input did not introduce any biases in performance. In contrast, when only proprioceptive information was present (e.g. walking in place), perceived self-motion was underestimated. This suggests that the vestibular system plays a critical role in spatial updating during passive self-motion (Frissen et al., 2011).

4.3 What is the Role of Visual Cues in Updating SSAA

According to the previous discussions regarding the perception of self-motion through optic flow (see section 1.3.3), it might be anticipated that differentiation in lighting conditions would affect updating of SSAA during head motion but based on previous results from our laboratory, our hypothesis was that it would not. Even though there was a statistically significant difference between lighting conditions, the absolute performance differences between lighting conditions were small, and we observed that updating of SSAA was still possible in the absence of visual motion cues. Notably, all lighting conditions exhibited performance significantly above chance source-guessing, even in the Passive Dark block, which had the lowest PC value with an average of 52.5% (Figure 3.4). These findings suggest that lighting does not appear to be necessary for spatially attending to auditory stimuli and updating SSAA. Moreover, having access to the full visual cues in Light conditions, does not seem to add much information to the other sensorimotor information available for updating SSAA given the small Cohen’s d effect size mentioned in section 3.4.

Previous research has explored the effect of visual cues on spatial updating through the integration of visual and vestibular cues during passive linear self-motion. In a study investigating the relative weighting of optic flow and vestibular cues, participants were instructed to remember a visual target while being passively moved in a linear manner while exposed to optic flow. The results indicated that humans integrate optic flow and vestibular self-displacement information using a weighted-averaging process, with visual cues receiving approximately four times more weight than vestibular cues (79% visual weight) (Koppen et al., 2019). These findings suggest a greater reliance on visual cues than vestibular cues for visual, rather than auditory, spatial updating during passive self-motion.

Regarding SSAA, Macpherson and Ransom (2018) reported that having visual access to the target location does not provide assistance in either static or dynamic SSAA when the target location is already known. Additionally, Kim (2011) demonstrated that individuals who are blindfolded can accurately do front/back localization when subjected to passive rotation, meaning that the visual cues are not necessary for dynamic localization.

In the following chapter, we will delve further into the possible investigation of visual flow in SSAA when no other information is available.

4.4 What Interpretations could Explain the Marginal Performance Difference Observed in the Active Condition Compared to the Passive Motion Condition?

The two-way repeated ANOVA analysis comparing the performance in Active motion conditions and the Passive motion condition among all 18 participants revealed a significant but only a slight difference (5%) and small effect size between these two conditions. The performance was slightly better in the Active conditions compared to the Passive conditions. Despite the expectation that SSAA performance would be much better in the Active condition due to the availability of proprioceptive/efference copy signals in addition to the vestibular information, Figure 3.4 from the main experiment indicates that the improvement in the Active condition is only marginal. This observation leads us to consider three possible interpretations.

The first interpretation is the possibility of only localization-based *reactive* updating during dynamic motion conditions, rather than maintaining a representation and engaging in all forms of updating (predictive updating based on vestibular and proprioceptive self-motion cues and reactive updating based on the perceived stimulus location). This interpretation was tested through comparison of SSAA performance in the Unpredictable condition (Experiment a) with performance in the dynamic conditions of the main experiment as shown in Figure 3.8. Comparing the significantly lower updating performance in the Unpredictable condition to the dynamic motion conditions and specifically the lowest-SSAA PC condition (Passive Dark), reveals that predictive updating does occur during self-motion alongside reactive updating. The spatial-attentional maps of Unpredictable experiment (Fig 3.9) and Active and Passive motion conditions in the main experiment (Fig 3.7) display divergent shapes, representing the differences in reported digits when the head is oriented to either side or the front. This is an additional indication of the likelihood of different types of updating mechanisms (predictive vs. reactive) operating in these experiments.

The second interpretation considers the possibility that despite having access to additional sensorimotor cues in the Active condition, the task of continuously oscillating their heads may have placed a higher cognitive load on listeners, reducing their attention. However, the addition of an extra tapping stick task to the Passive condition in Experiment (b) did not show a decreasing trend in performance (see Figure 3.10). This suggests that the PC score in the Active condition was not reduced by the cognitive load of continuous voluntary movements, indicating that the similar performance in the Active and Passive conditions cannot be attributed to the absence of voluntary head movement in the latter. Nevertheless, it is important to note that the sample size for this experiment was insufficient to draw definitive conclusions, necessitating further data collection to validate these findings.

Lastly, an only slightly higher PC value in Active condition compared with Passive condition suggests that although proprioceptive/efference copy information is available, it simply is not weighted as strongly as vestibular signals for updating SSAA. This is consistent with the studies by Genzel et al. (2016) and by Kim et al., (2013), which investigated the weighting of vestibular signals in auditory spatial updating during different rotations and in front/back localization, respectively. Their findings demonstrated a dominance of vestibular signals, with significantly higher weights assigned to vestibular cues compared to proprioceptive/efference copy. However, further research is needed to manipulate and minimize the availability of vestibular signals and determine their necessity for spatially updating selective auditory attention.

5 Conclusion, Significance and Future Directions

This chapter marks the culmination of our research investigation focused on assessing the sufficiency of vestibular and the necessity of proprioceptive signals in updating SSAA during head (Active) or whole-body (Passive) motion. Through a behavioral experiment involving a spatial-auditory task and manipulation of motion and lighting conditions, we have obtained significant findings that shed light on the role of sensorimotor information and the influence of head motion on SSAA. This concluding chapter, will provide a concise summary of these key findings, evaluate their significance, and outline potential directions for future research in this specific field.

To explore our first question which is as the listeners move their heads, what would happen to their focus of SSAA based on the new information from the rotation in the space and the new representation of the location, we observed whether the percentage of corrected responses in Active and Passive condition would be significantly higher than chance level. It becomes evident from the main results (Fig 3.4) and by comparison with the results of experiment (a) (Fig 3.8; reactive updating only) that the focus of SSAA was being predictively updated to the target loudspeaker when participants were moving either actively or passively. In other words, the representation of target location was updated based on the head rotation information (vestibular and proprioceptive/efference copy) and then the SSAA was being updated using the focus of new location representation. Another important conclusion from the main results is that participants performed significantly better than chance in all Passive conditions. This indicates that vestibular information is sufficient for updating SSAA and that proprioceptive/efference-copy information is not necessary. The finding that performance is only marginally improved when proprioceptive/efference-copy information is added to vestibular information suggests that the latter is given more weighting compared to proprioceptive/efference copy signals. Furthermore, having visual information about the sound source location and the head rotation was not necessary for updating SSAA. We also observed a frontal bias (under-compensation) of SSAA during the head motion, in the form of an increase in the frequency of misreported digits from the head-oriented distractor loudspeaker. However, the direction of head movement to the left or right didn't impact the frontal bias, suggesting that SSAA updating was rapid compared to the speed of head motion.

There were certain points that can be desirably improved for future direction of this thesis. For instance, the head movement trajectory, although in most cases well-matched between the Passive and Active motion conditions, wasn't identical. Thus, developing a method that controls the movement of dynamic motion conditions could enhance the procedure of the experiment. Additionally, our method of cue-elimination prevented quantitatively computing the weightings of vestibular and proprioceptive/efference copy signals, but this could be achieved by introducing a counter-rotation condition, which puts these cues in opposition or more realistic sound stimuli. We tried to minimize the availability of neck proprioceptive signals during the Passive conditions by having the neck supported, yet those signals were not completely eliminated during the Passive motion, and having proprioceptive signals from other parts of the body such as legs and trunk might also contribute to self-motion perception in Passive conditions.

In future research, several important steps can be taken to further advance our understanding of spatial auditory attention and its relation to self-motion cues.

First, reactive updating observed during the Unpredictable condition (experiment a) necessitates further investigations, as it could offer valuable insights into the underlying mechanisms of auditory spatial attention in dynamic environments. Second, the use of Galvanic Vestibular Stimulation (GVS) device (Dlugaiczyk et al., 2019) can help further elucidate the role of proprioceptive and visual cues in SSAA. Utilizing Sinusoidal GVS (sGVS), which imparts a sense of artificial rotation to a stationary listener without actual neck movement, could be employed to accentuate the current findings on the unnecessity of proprioceptive/efference copy and sufficiency of vestibular signals. Furthermore, Noisy GVS (nGVS) can be implemented to obstruct the vestibular information derived from the listener's head rotation, aiming to investigate whether vestibular signals are necessary for updating SSAA. Third, in our study, we considered proprioceptive and efference copy signals together due to the challenge of distinguishing their individual impacts on head position and movement. Previous research has attempted to separate these signals using different methods, such as rotating the head without body movement to isolate proprioceptive signals or blocking active head movements to isolate efference copy signals (Crowell et al., 1998; Cullen et al., 2011; Nakamura & Bronstein, 1995). Similar manipulations could be explored in future experiments on auditory spatial attention. Lastly, expanding the sample size, incorporating individuals from diverse age groups, would strengthen the reliability and generalizability of our findings.

To conclude, the significance of our study lies in its novelty in analyzing the relative importance of self-motion cues (vestibular and proprioceptive/efference copy) in updating SSAA during head movements. This research also offers valuable insights into spatial hearing processing and could enhance our understanding of the effects of vestibular and semicircular canal impairments concerning Spatial Selective Auditory Attention.

References

Bregman, A. S. (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press.

Bronkhorst, A. W. (2000). The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple-Talker Conditions. *Acta Acustica United with Acustica*, *86*(1), 117–128.

Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, *25*(5), 975–979. https://doi.org/10.1121/1.1907229

Colclasure, J. C., & Holt, J. R. (2007). Transduction and adaptation in sensory hair cells of the mammalian vestibular system. *Gravitational and Space Research*, *16*(2).

Crowell, J. A., Banks, M. S., Shenoy, K. V., & Andersen, R. A. (1998). Visual self-motion perception during head turns. *Nature Neuroscience*, *1*(8), Article 8. https://doi.org/10.1038/3732

Cullen, K. E., Brooks, J. X., Jamali, M., Carriot, J., & Massot, C. (2011). Internal models of self-motion: Computations that suppress vestibular reafference in early vestibular processing. *Experimental Brain Research*, *210*(3), 377–388. https://doi.org/10.1007/s00221-011-2555-9

Cullen, K. E., & Zobeiri, O. A. (2021). Proprioception and the predictive sensing of active self-motion. *Current Opinion in Physiology*, *20*, 29–38. https://doi.org/10.1016/j.cophys.2020.12.001

Cutting, J. E., & Vishton, P. M. (1995). Chapter 3 - Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth*.

In W. Epstein & S. Rogers (Eds.), *Perception of Space and Motion* (pp. 69–117). Academic Press. https://doi.org/10.1016/B978-012240530-3/50005-5

Davis, T. J., Grantham, D. W., & Gifford, R. H. (2016). Effect of motion on speech recognition. *Hearing Research*, *337*, 80–88. https://doi.org/10.1016/j.heares.2016.05.011

Day, B. L., & Fitzpatrick, R. C. (2005). The vestibular system. *Current Biology*, *15*(15), R583–R586. https://doi.org/10.1016/j.cub.2005.07.053

DeAngelis, G. C., & Angelaki, D. E. (2011). Visual-vestibular integration for self-motion perception. In *The Neural Bases of Multisensory Processes* (pp. 629–649). CRC Press. http://www.scopus.com/inward/record.url?scp=85059470657&partnerID=8YFLogxK

Dichgans, J., & Brandt, T. (1978). Visual-Vestibular Interaction: Effects on Self-Motion Perception and Postural Control. In S. M. Anstis, J. Atkinson, C. Blakemore, O. Braddick, T. Brandt, F. W. Campbell, S. Coren, J. Dichgans, P. C. Dodwell, P. D. Eimas, J. M. Foley, R. Fox, L. Ganz, M. Garrett, E. J. Gibson, J. S. Girgus, M. M. Haith, Y. Hatwell, E. R. Hilgard, … H.-L. Teuber (Eds.), *Perception* (pp. 755–804). Springer. https://doi.org/10.1007/978-3-642-46354-9_25

Dlugaiczyk, J., Gensberger, K. D., & Straka, H. (2019). Galvanic vestibular stimulation: From basic concepts to clinical applications. *Journal of Neurophysiology*, *121*(6), 2237–2255. https://doi.org/10.1152/jn.00035.2019

Feddersen, W. E., Sandel, T. T., Teas, D. C., & Jeffress, L. A. (2005). Localization of High-Frequency Tones. *The Journal of the Acoustical Society of America*, *29*(9), 988–991. https://doi.org/10.1121/1.1909356

Fetsch, C. R., Turner, A. H., DeAngelis, G. C., & Angelaki, D. E. (2009). Dynamic Reweighting of Visual and Vestibular Cues during Self-Motion Perception. *Journal of Neuroscience*, *29*(49), 15601–15612. https://doi.org/10.1523/JNEUROSCI.2574-09.2009

Frank, S. M., Wirth, A. M., & Greenlee, M. W. (2016). Visual-vestibular processing in the human Sylvian fissure. *Journal of Neurophysiology*, *116*(2), 263–271. https://doi.org/10.1152/jn.00009.2016

Genzel, D., Firzlaff, U., Wiegrebe, L., & MacNeilage, P. R. (2016). Dependence of auditory spatial updating on vestibular, proprioceptive, and efference copy signals. *Journal of Neurophysiology*, *116*(2), 765–775. https://doi.org/10.1152/jn.00052.2016

Gibson, J. J. (1950). *The perception of the visual world* (pp. xii, 242). Houghton Mifflin.

Grabherr, L., Nicoucar, K., Mast, F. W., & Merfeld, D. M. (2008). Vestibular thresholds for yaw rotation about an earth-vertical axis as a function of frequency. *Experimental Brain Research*, *186*(4), 677–681. https://doi.org/10.1007/s00221-008-1350-8

Haykin, S., & Chen, Z. (2005). The Cocktail Party Problem. *Neural Computation*, *17*(9), 1875–1902. https://doi.org/10.1162/0899766054322964

Hebrank, J., & Wright, D. (2005). Spectral cues used in the localization of sound sources on the median plane. *The Journal of the Acoustical Society of America*, *56*(6), 1829–1834. https://doi.org/10.1121/1.1903520

Highstein, S. M., & Holstein, G. R. (2012). The Anatomical and Physiological Framework for Vestibular Prostheses. *The Anatomical Record*, *295*(11), 2000–2009. https://doi.org/10.1002/ar.22582

Howard, I. P., & Howard, A. (1994). Vection: The Contributions of Absolute and Relative Visual Motion. *Perception*, *23*(7), 745–751. https://doi.org/10.1068/p230745

Kidd, G., Jr., Arbogast, T. L., Mason, C. R., & Gallun, F. J. (2005). The advantage of knowing where to listena). *The Journal of the Acoustical Society of America*, *118*(6), 3804–3815. https://doi.org/10.1121/1.2109187

Kim, J., Barnett-Cowan, M., & Macpherson, E. A. (2013). Integration of auditory input with vestibular and neck proprioceptive information in the interpretation of dynamic sound localization cues. *Proceedings of Meetings on Acoustics*, *19*(1), 050142. https://doi.org/10.1121/1.4799748

Lackner, J. R., & DiZio, P. (2005). Vestibular, Proprioceptive, and Haptic Contributions to Spatial Orientation. *Annual Review of Psychology*, *56*(1), 115–147. https://doi.org/10.1146/annurev.psych.55.090902.142023

Lappe, M., Bremmer, F., & van den Berg, A. V. (1999). Perception of self-motion from visual flow. *Trends in Cognitive Sciences*, *3*(9), 329–336. https://doi.org/10.1016/S1364-6613(99)01364-9

Lerens, E., & Renier, L. (2014). Does visual experience influence the spatial distribution of auditory attention? *Acta Psychologica*, *146*, 58–62. https://doi.org/10.1016/j.actpsy.2013.12.002

Lin, G., & Carlile, S. (2015). Costs of switching auditory spatial attention in following conversational turn-taking. *Frontiers in Neuroscience*, *9*. https://www.frontiersin.org/articles/10.3389/fnins.2015.00124

Macpherson, E. A., & Ellis, B. K. (2016). Listener head motion can degrade spatial selective auditory attention. *The Journal of the Acoustical Society of America*, *139*(4_Supplement), 2209. https://doi.org/10.1121/1.4950600

Macpherson, E. A., Jeon, M., & Ransom, S. (2019). Updating of spatial selective auditory attention under-compensates for listener head movement. *The Journal of the Acoustical Society of America*, *145*(3), 1724–1724. https://doi.org/10.1121/1.5101334

Macpherson, E. A., & Kerr, D. M. (2008). Minimum head movements required to localize narrowband sounds. *American Audiology Society 2008 Annual Meeting*.

Macpherson, E. A., & Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. *The Journal of the Acoustical Society of America*, *111*(5), 2219–2236. https://doi.org/10.1121/1.1471898

Macpherson, E. A., & Ransom, S. (2018). Effects of vision, listener head movement, and target location on spatial selective auditory attention. *The Journal of the Acoustical Society of America*, *143*(3), 1813–1813. https://doi.org/10.1121/1.5035945

Masanao, E. (2001). Spatial unmasking and attention related to the cocktail party problem. *Acoust. Sci. & Tech.*, *22*, 351–357.

Merfeld, D. M., Zupan, L., & Peterka, R. J. (1999). Humans use internal models to estimate gravity and linear acceleration. *Nature*, *398*(6728), Article 6728. https://doi.org/10.1038/19303

Mills, A. W. (2005). On the Minimum Audible Angle. *The Journal of the Acoustical Society of America*, *30*(4), 237–246. https://doi.org/10.1121/1.1909553

Mondor, T. A., & Zatorre, R. J. (1995). Shifting and focusing auditory spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(2), 387–409. https://doi.org/10.1037/0096-1523.21.2.387

Plomp, R. (1976). Binaural and Monaural Speech Intelligibility of Connected Discourse in Reverberation as a Function of Azimuth of a Single Competing Sound Source (Speech or Noise). *Acta Acustica United with Acustica*, *34*(4), 200–211.

Quinlan, P. T., & Bailey, P. J. (1995). An examination of attentional control in the auditory modality: Further evidence for auditory orienting. *Perception & Psychophysics*, *57*(5), 614–628. https://doi.org/10.3758/BF03213267

Ramkhalawansingh, R., Butler, J. S., Link to external site, this link will open in a new window, & Campos, J. L. (2018). Visual–vestibular integration during self-motion perception in younger and older adults. *Psychology and Aging*, *33*(5), 798–813. https://doi.org/10.1037/pag0000271

Rayleigh, Lord. (1907). XII. On our perception of sound direction. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *13*(74), 214–232. https://doi.org/10.1080/14786440709463595

Shaw, E. A. G. (1997). Spatial perception, the acoustics of the external ear, and interactions with earphones. *The Journal of the Acoustical Society of America*, *102*(5_Supplement), 3116. https://doi.org/10.1121/1.420565

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, *12*(5), 182–186. https://doi.org/10.1016/j.tics.2008.02.003

St George, R. J., & Fitzpatrick, R. C. (2011). The sense of self-motion, orientation and balance explored by vestibular stimulation. *The Journal of Physiology*, *589*(4), 807–813. https://doi.org/10.1113/jphysiol.2010.197665

Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. *Journal of Experimental Psychology*, *27*(4), 339–368. https://doi.org/10.1037/h0054629

Zalewski, C. K. (2017). *Rotational Vestibular Assessment*. Plural Publishing.

Zhang, P. X., & Hartmann, W. M. (2010). On the ability of human listeners to distinguish between front and back. *Hearing Research*, *260*(1), 30–46. https://doi.org/10.1016/j.heares.2009.11.001

Appendix

Date: 30 March 2023

To: Ewan MacPherson

Project ID: 102445

Review Reference: 2023-102445-76828

Study Title: The Role of the Vestibular, Visual, and Proprioceptive Systems in Processing Dynamic Sound Localization Cues - 18956E

Application Type: Continuing Ethics Review (CER) Form

Review Type: Delegated

Date Approval Issued: 30/Mar/2023 10:58

REB Approval Expiry Date: 11/Apr/2024

Dear Ewan MacPherson,

The Western University Research Ethics Board has reviewed the application. This study, including all currently approved documents, has been re-approved until the expiry date noted above.

REB members involved in the research project do not participate in the review, discussion or decision.

Western University REB operates in compliance with, and is constituted in accordance with, the requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The REB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Please do not hesitate to contact us if you have any questions.

Electronically signed by:

Mr. Josh Hatherley, Ethics Coordinator on behalf of Dr.P. Jones, HSREB Chair 30/Mar/2023 10:58

Reason: I am approving this document

Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).

Curriculum Vitae

Name:

Erisa Davoudi

Education:

Zanjan University, Iran

BSc Psychology

Western University, Canada

MSc Neuroscience

Honours and Awards:

First winner of the “7th Seminars of Scientific Achievements” From National Organization for Development of Exceptional Talents, Rasht, Iran | 2015

Neuroscience Travel Award | 2023

Poster Presentations:

Davoudi, E & Macpherson, E.A. *The Roles of Vestibular and Proprioceptive Signals in Updating Spatial Selective Auditory Attention during Head Motion.* 33rd Annual Meeting of the Canadian Society for Brain, Behaviour and Cognitive Science, Guelph. July 2023.

Davoudi, E & Macpherson, E.A. *The Roles of Vestibular and Proprioceptive Signals in Updating Spatial Selective Auditory Attention during Head Motion.* Annual research event London Health Research Day, Western University, June 2023

Davoudi, E & Macpherson, E.A. *The Roles of Vestibular and Proprioceptive Signals in Updating Spatial Selective Auditory Attention during Head Motion.* 41st annual meeting of the Southern Ontario Neuroscience Association, Scarborough. May 2023.