MULTISENSORY PLASTICITY IN ADULT HUMANS: CLINICAL IMPLICATIONS FOR INJURY AND RECOVERY

By

Collins Opoku-Baah

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

In

Neuroscience

May 13, 2022

Nashville, Tennessee

Approved:

David Calkins, Ph.D.

Mark Wallace, Ph.D.

Randolph Blake, Ph.D.

Tonia Rex, Ph.D.

Dedication

To my mentors, friends, and family especially my wife, Abigail and my daughters, Eden, and Aseda.

Acknowledgement

While the beginning and the end of an endeavor usually stir up emotions and capture the spotlight, it is the moments we experience, the lessons we learn, the failures we endure, the successes we celebrate, the people we lose or gain, the relationships we build, that make the journey worthwhile. Of course, although, the PhD journey is ripe with lots of difficulties, I am happy to say that I have enjoyed every moment of it and I am sincerely blessed by the contributions of very intelligent and generous people who made this dream, a reality.

First, I am sincerely grateful to my mentors who have supported and guided me throughout my PhD training. I would like to begin with Dr. Randolph Blake because without his involvement, this achievement would not have begun. Dr. Blake, I look with fond memories of our earliest interactions while I was an optometry student in Ghana seeking to be mentored and trained by world-leading scientists. Your graciousness and approachability made feasible what seemed to many a daunting task. You facilitated my travel to the United States and took me in not only as a mentee but also as a “son”. Throughout the years, I enjoyed working with you on many projects and I have learned a great deal from our scientific and non-scientific conversations. To the Dr Calkins and Dr Rex, I am grateful for accepting to be part of my mentors, for your contributions to my research projects and introducing me to important people at the Vanderbilt Eye Institute. Finally, Dr Mark Wallace, it has been an amazing opportunity to work in your lab and explore these interesting ideas. You cared deeply about not only my academic work and professional growth but also my personal well-being. Your flexibility permitted me to perform both my PhD and parenting responsibilities well. In the end, nothing I have accomplished could have happened without your mentorship, support, and guidance; and for that, I am forever grateful.

Second, I am grateful to a team of wonderful undergraduate researchers – Brian, Julia, Annika, and Amanda – who contributed immensely to my PhD projects from brainstorming ideas to data collection and analysis. I was also fortunate to work in the same lab with very smart and friendly scientists and lab mates. I would like to thank Kevin and Jocelyn for their assistance in learning psychophysics and Psychtoolbox in Matlab. I am grateful to both past (David, Jean-Paul, Aaron, Gabby, Tovar, Lauren, and Iliza) and present lab colleagues (Jacob, Adriana, Sarah, Ansley, Emma) for their collaboration on research projects, help with learning useful techniques and serving as participants for my lengthy and numerous experiments. I would like to thank Bruce, Roger, and Patrick for their contributions in building a clinical perimetry device for my research projects and lastly, I am sincerely grateful to Rosalind Johnson, Beth Sims, Angela Collins, and Darlene for their support in planning numerous academic events and keeping up with the requirements of the PhD degree.

Finally, my PhD experience was enriched by the support and love of numerous friends and family members. To my friends, I am thankful to Felix, Anthony and Bright for their numerous conversations about life and career plans and pushing one another to greater heights. I am grateful to Laura (Kay) Eaves, Richard Zeigler, and all members of the international students' fellowship for their spiritual support and fellowship throughout the years. To my family, I would like to express my sincere thanks to Richard, Elaine, Eden and the entire Newman family and friends for taking me in as son and brother and showing me sincere love and support. I am eternally grateful for everything you have done for my family and me during my stay in Nashville. Furthermore, I am grateful to George and Esther Gyamfi for their love and support for my family and me. To my parents, siblings, and in-laws, I cannot thank you enough for your continued investment, prayer, and support in my life. Dad, you left to be with God very early in my PhD and I hope I have made you proud with this accomplishment. Last and certainly not the least, I am forever grateful to my wife and better half, Abigail. You have certainly been the major pillar in my life throughout my years in the PhD. Your presence through the difficulties brought me comfort and hope. Thank you so much for your care, love, support, and fellowship and for also, giving me two beautiful daughters, Eden, and Aseda, who are the spotlight of my PhD training.

TABLE OF CONTENTS

DEDICATION .... **ii**

ACKNOWLEDGEMENT .... **iii**

LIST OF FIGURES .... **viii**

1. Introduction **1**

- 1.1 General overview of dissertation work: background, motivation and aims 1
- 1.2 An Introduction to Multisensory Processing 4
- 1.3 Temporal Aspects of Multisensory Processing 11
- 1.4 Neuroplasticity 21
- 1.5 References 39

2. Brief Period of Monocular Deprivation Drives Changes in Audiovisual Temporal Perception **63**

- 2.1 Abstract 63
- 2.2 Introduction 64
- 2.3 Methods 68
- 2.4 Results 75
- 2.5 Discussion 85
- 2.6 References 91

3. Neural Correlates of Short-Term Monocular Deprivation Effects on Multisensory Temporal Perception **102**

- 3.1 Abstract 102
- 3.2 Introduction 103
- 3.3 Method 107
- 3.4 Results 116
- 3.5 Discussion 123
- 3.6 References 128

4. A Temporally Correlated Sound Facilitates Visual Learning of Contrast Detection **136**

- 4.1 Abstract 136
- 4.2 Introduction 137
- 4.3 Methods 140
- 4.4 Results 154
- 4.5 Discussion 168
- 4.6 References 176

5. General Discussion 185
5.1 Low Vision: Clinical model of adult-onset visual deprivation 185
5.2 Clinical Implications for Injury in Low Vision 185
5.3 Clinical Implications for Recovery and Rehabilitation in Low Vision 186
5.4 References 197

A. Appendix Ch A: Binocular enhancement of multisensory temporal perception 208
A.1 Abstract 208
A.2 Introduction 209
A.3 Methods 215
A.4 Results 225
A.5 Discussion 231
A.6 References 237

LIST OF FIGURES

Figure Page

1. 1 Psychophysical Study of Multisensory Temporal Perception.................................. 15

2. 1 Experimental procedure. 73
2. 2 Mean proportion of synchrony reports... 76
2. 3 Effects of deprivation on the point of subjective simultaneity (PSS). 79
2. 4 Effects of deprivation on the temporal binding window (TBW). 82
2. 5 Correlation between the absolute change in point of subjective simultaneity (PSS) and the relative change in temporal window of simultaneity (TWS)........................ 84

3. 1 Experimental procedure. 114
3. 2 Effects of monocular deprivation on the psychophysical measures of audiovisual temporal perception. 118
3. 3 Effects of monocular deprivation on alpha (8-12), low beta (12-20) and high beta (20-30) post-stimulus oscillatory power... 120
3. 4 Effects of monocular deprivation on alpha (8-12), low beta (12-20) and high beta (20-30) post-stimulus oscillatory inter-trial phase coherence (itpc)....................... 122
3.5 Plots of deprivation-induced (post-pre) changes for the psychophysical measures of audiovisual temporal perception.. 135

4. 1 Training paradigms and predictions from PTM..................................... 152
4. 2 Baseline Contrast Thresholds.. 155
4. 3 Training results... 161
4. 4 Post training results. 164
4. 5 PTM modelling results 167

5.1 Proposed Neural Mechanisms Underlying Visual and Multisensory Perceptual Learning (PL) in Low Vision.. 196

A. 1 Schematic of the procedure for the (A) flash-beep SJ task and (B) speech SJ task. 219
A. 2 Causal inference model for audiovisual SJ tasks. 223
A. 3 Mean proportions of synchrony reports. 226
A. 4 Effects of viewing condition and stimulus type on PSS and TBW........................ 228
A. 5 The effects of viewing condition of sensory noise for the flash-beep SJ task. 230

Chapter 1

Introduction

1.1 General overview of dissertation work: background, motivation and aims

The human brain is remarkably plastic, that is, it possesses the ability to change its morphology and function to meet the demands of the changing environment (Pascual-Leone, Amedi, Fregni, and Merabet, 2005). This intrinsic property of the nervous system is highest during development, in a time period generally referred to as the "critical" or "sensitive" period (Berardi, Pizzorusso, and Maffei, 2000). During this period, normal sensory experience plays a crucial role in shaping brain connectivity and function (Berardi et al., 2000). Consequently, changes in sensory experience can cause changes in the morphology and the function of the brain in ways that can have lasting effects on sensory function, perception, and behavior. In the visual domain, earlier work by Hubel and Wiesel discovered that a few weeks to months of visual (i.e., monocular) deprivation through lid suturing or eye occlusion changed the receptive field properties of neurons and reorganized ocular dominance columns in favor of the open eye (Hubel and Wiesel, 1970; Wiesel and Hubel, 1963a, 1963b). These findings were fundamental to our understanding of the role normal visual sensory experience plays in the development of the visual system.

From a clinical perspective, visual deprivation can occur during development in humans as a result of conditions such as glaucoma, refractive errors, retinal disorders, cornea disorders, cataracts etc. Visual deprivation can be monocular or binocular, can be short-lived or permanent and can occur in varying degrees ranging from partial deprivation to complete deprivation. Regardless of these differences, visual deprivation in humans is associated with striking reorganization in brain structure and connectivity, and impaired performance on several visual tasks including visual acuity, contrast detection, shape, and depth perception (Fine, Smallman, Doyle, and MacLeod, 2002; Fine et al., 2003; Levi, McKee, and Movshon, 2011; Ostrovsky, Andalman, and Sinha, 2006).

In the past, although the study of sensory perception focused mostly on one modality at a time, a growing body of compelling evidence supports the idea that the different sensory modalities are not separate entities and that, they interact at various brain levels (Foxe and Schroeder, 2005; Ghazanfar and Schroeder, 2006; Stein and Stanford, 2008). Indeed, integrating information from multiple modalities augments our perceptual experience through enhanced detection (Frassinetti, Bolognini, and Làdavas, 2002; Lovelace, Stein, and Wallace, 2003), discrimination (Ernst and Banks, 2002), localization (Hairston, Laurienti, Mishra, Burdette, and Wallace, 2003; Zou, Müller, and Shi, 2012), and speeding response times (Diederich and Colonius, 2004). Interestingly, altered visual experience during development can influence the maturation of aspects of multisensory integration (such as temporal or spatial) during development. Evidence for the role of normal visual sensory experience in the development of multisensory integration has been provided in dark rearing studies in animal models (Carriere et al., 2007; Wallace, Perrault, Hairston, and Stein, 2004) and in patients who experienced short-lived visual deprivation because of congenital cataracts (Chen, Lewis, Shore, and Maurer, 2017; Putzar, Goerendt, Lange, Rösler, and Röder, 2007; Putzar, Hötting, and Röder, 2010) and refractive errors (Richards, Goltz, and Wong, 2017). In blind patients where there is complete visual deprivation, structural and functional reorganization of the brain especially in the visual cortex has been functionally linked with compensatory behaviors observed in various tasks such as braille reading (Cohen et al., 1997; Kupers et al., 2007) and verb-generation (Amedi, Floel, Knecht, Zohary, and Cohen, 2004) and also the use of auditory-based sensory substitution devices (Merabet et al., 2009).

Importantly, visual deprivation can occur also in adults after development has ended. In fact, age is a major risk factor for most of the leading causes of visual impairment such as age-related macular degeneration, primary open angle glaucoma, cataracts etc. (Congdon et al., 2004). Although after development, brain plasticity wanes drastically, the evidence of sufficient plasticity in the adult brain implies the following. First, visual deprivation occurring in the adult can influence multisensory integrative abilities and consequently, such changes can underlie the development and the nature of learned compensatory behavior as observed in people with congenital blindness. Despite this possibility, the impact of visual deprivation on multisensory integrative abilities in adult humans has not been quite explored. Second, the presence of plasticity opens up avenues for visual rehabilitation through therapies such as perceptual learning. Interestingly, perceptual learning paradigms that incorporate information from different sensory modalities have been shown to produce more effective and efficient learning outcomes compared with paradigms that rely on information from one modality (Kim, Seitz, and Shams, 2008; Seitz, Kim, and Shams, 2006; Shams and Seitz, 2008). Nevertheless, currently, it is not clear whether such multisensory facilitated learning benefits can occur for different visual stimuli and tasks including those that are often utilized in conventional visual perceptual learning such contrast detection and discrimination.

To this effect, the goal of my dissertation was two-fold. First, to characterize and understand the nature of multisensory temporal perceptual changes in adult humans after brief periods of monocular deprivation using both psychophysical and neuroimaging techniques. Second, to assess the impact of multisensory-based perceptual learning paradigm in improving visual perceptual outcomes in a contrast detection task, a simple visual task used in assessing visual performance in patients with visual impairment and also, employed in visual perceptual learning paradigms.

To preview, this dissertation contains five chapters detailing the work I carried out on these aims during my PhD training. For the remainder of chapter 1, I take a deeper dive and review the literature on relevant topics and concepts surrounding this work. Chapter 2 focusses on the behavioral study to determine the impact of short-term monocular deprivation on the psychophysical measures of audiovisual temporal perception. Chapter 3 investigates the neural mechanisms underlying the observed changes in audiovisual temporal perception after monocular deprivation, focusing on the role of neural oscillatory activity recorded using electroencephalography (EEG). Chapter 4 investigates the impact of multisensory perceptual learning paradigms on visual perceptual outcomes compared with visual only training paradigms. Also in chapter 4, we employ a popular model based on signal detection theory to understand how the perceptual system changes following these two types of perceptual learning. Lastly, in chapter 5, I discuss these findings zooming in on the clinical implications for injury and recovery in low vision patients.

1.2 An Introduction to Multisensory Processing

1.2.1 General Overview of Multisensory Interactions

In psychology and neuroscience, the study of how the brain encodes, decodes, and interprets sensory information for perception, behavior and cognition has received a great deal of attention in the past century. Traditionally, these investigations focused heavily on studying one sensory modality at a time e.g., vision, audition or somatosensory. However, such an approach belies reality: that is the fact that most of the events we encounter in the world are multisensory in nature. For example, events such as seeing and touching an object in our hands; seeing and hearing a person talk or a car move; hearing and feeling the direction of the wind; and tasting and smelling food, suggest that our brain routinely receives a barrage of information concurrently from many senses. The phrase "multisensory integration" has been commonly used to describe the fusion of information from different sensory modalities into a single unified percept (Stein and Meredith, 1993; Stein and Stanford, 2008). However, this phrase by definition precludes situations where one modality influences another without necessarily forming a single unified percept. A more encompassing term "multisensory interactions" can be used to generally describe the influence of one sensory modality on the activity of another modality.

Early views on multisensory interactions suggested that the different types of sensory information were first processed separately and extensively in dedicated pathways and cortical areas before they were relayed to specialized higher-level association brain areas that supported the integration of these multisensory signals (Foxe and Schroeder, 2005). However, recent evidence from neural recording studies in macaques and brain imaging studies in humans have challenged this traditional view and have highlighted the presence of significant multisensory processing in brain regions classically considered "unisensory." (Ghazanfar and Schroeder, 2006)

1.2.2 The Guiding Principles of Multisensory Interactions

Dictating the interaction between the sensory systems are a set of fundamental principles closely yoked to the physical characteristics of the stimuli that are combined. Thus, multisensory interactions/integration and the associated neural, behavioral, and perceptual gains are most prevalent (and often largest) when the spatial and temporal disparities between the combined stimuli are small (Stein and Meredith, 1993). In addition, the largest gains accompany the pairing of two weakly effective stimuli (the inverse-effectiveness principle) (Stein and Meredith, 1993). These principles make a great deal of intuitive sense, as stimuli that are spatially- and temporally proximate are likely to be derived from the same object or event. Hence, the brain makes use of this statistical information to make probabilistic judgments about the likelihood that the stimuli co-occurred. Equally intuitive is inverse effectiveness, which shows that highly effective stimuli need little or no amplification. Conversely, where multisensory systems are most important are when both stimuli are weakly effective when presented on their own.

1.2.3 Perceptual and Behavioral Consequences of Multisensory Interactions

The ability of our perceptual system to accommodate consistent interaction between the different sensory signals at relevant stages along the stretch of the processing hierarchy ensures the conservation of the multi-modal picture of events in our world. Consequently, such interactive processes should have advantageous impact on perception and behavior. A great deal of work has highlighted that our brains combine these different sensory signals in order to enhance the detection and discrimination of relevant events (Stein and Meredith, 1993; Stein and Stanford, 2008; Fister, Stevenson, Nidiffer, Barnett, and Wallace, 2016; Seitz et al., 2006; Shams and Seitz, 2008; Van der Burg, Cass, Olivers, Theeuwes, and

Alais, 2010; Philippi, van Erp, and Werkhoven, 2008; Diederich and Colonius, 2004; Hershenson, 1962). For example, in a noisy coffee shop, our ability to comprehend the information from the audible speech signal is greatly aided by the presence of the appropriate lip movements, and these multisensory-mediated improvements in speech comprehension grow as the magnitude of the background noise escalates (Ross, Saint-Amour, Leavitt, Javitt, and Foxe, 2006). This phenomenon can boost comprehension by an amount equivalent to making the intensity of the auditory signal alone 15-20dB louder (Sumby and Pollack, 1954).

Generally, the perceptual and behavioral benefits of multisensory interactions can be divided into two broad categories. The first type involves perceptual enhancements that occur when redundant information received from the multiple senses about one environmental property are combined. In this category, information provided individually by the two senses can be used to make judgments about the property being estimated and thus, combining the information from the senses decreases sensory uncertainty and enhances the reliability of perceptual judgments (Ernst and Banks, 2002). For instance, estimating the height of an object using both visual and haptic exploration reduces discrimination thresholds more than using either visual or haptic information alone (Ernst and Banks, 2002). In the temporal domain, Murai and Yotsumoto (2018) demonstrated that combining information between the visual and auditory senses enhanced the estimation of the timing of events.

The second category comprises perceptual enhancements that are observed when task irrelevant or relevant information from one modality influences perceptual judgments specifically related to another modality. In such circumstances, the task-irrelevant information usually shares simple spatiotemporal correspondence with the information used to make the perceptual judgments while the task-relevant formation share complex task-relevant features. In addition, while the impact of task-irrelevant information in another modality mostly occur for low level tasks such as detection, localization, and frequency discrimination, that of task relevant information may occur for both low level and complex tasks such as motion and speech perception respectively. For instance, in terms of task-irrelevant formation, Frassinetti et al. (2002) investigated the effects of cross-modal cues on visual detection sensitivity in a study where they presented auditory white noise bursts and/or flashes of light in various spatial locations on the azimuth to normally sighted individual. Their findings showed that at any spatial location, spatially and temporally coincident cross-modal cues increased visual detection sensitivity compared to the unimodal visual condition (Frassinetti et al., 2002). Similarly, Lovelace et al. (2003) showed that participants' ability to detect a sound stimulus was enhanced by a task-irrelevant light.

On the other hand, for task-relevant information, Møller et al. (2018) showed that presenting a visual stimulus that varied in vertical position— with vertical position known for its correspondence with auditory pitch (Parise, Knorre, and Ernst, 2014; Parise, Spence, and Deroy, 2016) —facilitated the detection of subtle pitch changes in auditory targets. In another study, Su (2014) showed that a bouncing human point-light figure conveying visual beat information enhanced the ability to perceive and synchronize to auditory rhythms. Lunghi and Alais (2013) discovered that concurrent presentation of a haptic signal matched in orientation to one of the visual gratings boosted and reduced the dominance and suppression periods respectively of that grating during a binocular rivalry task. In the area of perceptual tasks with more complex and ecologically valid stimuli such as speech, similar perceptual enhancements arising from audiovisual stimulus correspondence have been demonstrated. During speech perception, the area of the mouth opening and the acoustic envelope of the speech sound share robust spatial and temporal correspondences (Chandrasekaran, Trubanova, Stillittano, Caplier, and Ghazanfar, 2009). Several studies have demonstrated that being able to visualize the talker’s lip movements significantly enhances comprehension of the auditory speech signal under both good (Arnold and Hill, 2001; Reisberg, Mclean, and Goldfield, 1987) and noisy listening conditions (Ross, Saint-Amour, Leavitt, Javitt, and Foxe, 2007; Sumby and Pollack, 1954). In addition, visual information from the talker’s mouth movements can aid in the detection of spoken sentences masked by acoustic white noise under noisy conditions (Grant and Seitz, 2000). Following the principle of inverse effectiveness, earlier studies on the effect of different levels of noise on the magnitude of visually facilitated speech comprehension and intelligibility reported a monotonic relationship where greater multisensory gains were achieved under very low signal-to-noise conditions (Erber, 1969; Sumby and Pollack, 1954). However, a recent study used a relatively larger stimulus set compared with the previous studies and demonstrated that maximal multisensory gains were achieved within a range of intermediate signal-to-noise ratios (Ross, Saint-Amour, Leavitt, Javitt, et al., 2007). Unlike the previous studies, the findings from Ross et al. (2007) suggest that there may be a “sweet spot” for multisensory gain at intermediate SNRs.

Lastly, beyond perceptual enhancements, multisensory interactions may also result in illusory percepts. These illusions often occur when some degree of conflict is introduced between the multisensory cues being processed. For example, the McGurk effect is a speech-based illusion, which occurs when an auditory syllable (phoneme) paired with an incongruent visual syllable (viseme) results in the perception of a novel syllable (MacDonald and McGurk, 1978; Mallick, Magnotti, and Beauchamp, 2015; McGurk and MacDonald, 1976). Another illusion is the ventriloquist effect where vision captures auditory perception when a spatial conflict is introduced between the cues (Alais and Burr, 2004; Bertelson and Radeau, 1981). Recently, Shams et al. (2000) demonstrated the sound-induced flash illusion (SIFI) which is the perception of two flashes when a single flash is presented with two beeps. Together, these illusions provide evidence of the powerful influence of multisensory interactions on perception.

1.2.4 Mechanistic Principles of Multisensory Influences on Perception

Several studies have suggested that the brain combines sensory signals from multiple modalities relevant to an environmental object or event to, first, reach the most reliable (unbiased) estimate and, second, to minimize the variance associated with the final estimate (Ernst and Banks, 2002; Ernst and Bülthoff, 2004). Importantly, the brain achieves this sensory cue combination, termed maximum likelihood estimation, by weighting the signals according to their relative reliabilities. While this model can account for circumstances where the signals from the different sensory modalities are spatiotemporally coincident, it appears that when there is moderate or large conflict between the signals, the brain has to decide whether to combine or segregate the signals (Körding et al., 2007). This decision is based on the brain's ability to infer the unknown underlying causal structure of the signals, which is whether they originate from a common source or different sources (Körding et al., 2007). Based on this, Shams and colleagues developed the causal inference model, which has been applied to several perceptual tasks including spatial localization (Körding et al., 2007; Odegaard and Shams, 2016; Odegaard, Wozny, and Shams, 2017), temporal numerosity (Odegaard and Shams, 2016), heading estimation (De Winkel, Katliar, and Bülthoff, 2017), audiovisual temporal simultaneity judgments (Magnotti, Ma, and Beauchamp, 2013); and perceptual phenomena including the spatial ventriloquist effect (Körding et al., 2007; Odegaard and Shams, 2016).

1.3 Temporal Aspects of Multisensory Processing

1.3.1 Defining Multisensory Temporal Perception

For us to benefit from multisensory integration, the brain must construct an accurate representation of the world by determining which signals originate from a single source and hence should be integrated or which ones are from different sources and should be segregated (Stein and Meredith, 1993). The brain is known to do this by relying on the amodal and physical characteristics of the stimuli that are to be combined such as their spatial and temporal relationship (Stein and Meredith, 1993). Hence, the closer two stimuli are in space and in time, the more likely they belong to the same event (Bertelson, 1999; Stein and Meredith, 1993). This work focused on how the brain constructs a unified representation of the world using the temporal relationship between the stimuli. While temporal factors are not more important than the spatial factors, there is recent accumulation of evidence that has outlined how multisensory temporal function changes during typical development, and also, a growing acknowledgment that multisensory temporal acuity is altered in a number of neurodevelopmental disabilities – three of which, autism, dyslexia, and schizophrenia (Blau, van Atteveldt, Ekkebus, Goebel, and Blomert, 2009; Kwakye, Foss-Feig, Cascio, Stone, and Wallace, 2010; Ross, Saint-Amour, Leavitt, Molholm, et al., 2007; Stevenson et al., 2014; Szycik et al., 2009).

In fact, the brain's task of using the temporal relationship between the cues to determine their source is not straightforward considering the fact there is no known dedicated organ determining time and the physical and neural transmission times between the senses are not the same (Murray and Wallace, 2011). For example, while in the external world, vision travels faster than sound (i.e., 300,000,000 m/s for vision vs. 330 m/s for audition), in the brain, sound is known to be transmitted faster than vision (approximately 10 ms for sound vs. 50 ms for vision) (Murray and Wallace, 2011). Due to these differences, the simultaneous arrival of information from the different senses to their primary cortical areas is dependent on the distance between the events producing these signals and the observer. In fact, it is estimated that only those audiovisual events occurring in the distance of about 10 to 15 meters from the observer (i.e., the so called "horizon of simultaneity) will yield synchronous arrival times between the visual and the auditory information (Pöppel and Artin, 1988; Pöppel, Schill, and von Steinbüchel, 1990). Therefore, for events within 15 m from the observer, sound is likely to arrive at the primary sensory cortices faster than vision whereas for events farther away, vision is more likely to arrive before sound (Pöppel and Artin, 1988; Pöppel et al., 1990).

Despite the lags in arrival and processing times between the different senses, the brain can maintain temporal coherence and hence construct an accurate representation of the physical world. This raises an important question of how the brain deals with and maintains temporal coherence, and how do we study this mechanism. Among the theories that exist, the most dominant one is the construct of the temporal window of integration also known as the temporal binding window (Murray and Wallace, 2011; Wallace and Stevenson, 2014). This notion suggests that the brain has a hypothetical window with which an observer judges the perceived simultaneity of multisensory events. Thus, stimuli with temporal delays falling within this window are likely to be perceived as synchronous and vice versa. Furthermore, this indicates a window of time within which the multisensory stimuli are highly likely to be perceptually bound or integrated (Wallace and Stevenson, 2014).

1.3.2 Psychophysical Study of Multisensory Temporal Perception

Psychophysically, several studies have focused on understanding how the brain deals with multisensory temporal factors using simultaneity judgment (SJ) tasks (Zampini, Guest, Shore, and Spence, 2005). In a typical SJ task, participants are presented with paired multisensory stimuli (such as a visual flash and an auditory beep) with varying stimulus onset asynchronies (SOAs) and are asked to determine whether the stimulus pair was "synchronous" or "asynchronous" (Figure 1.1). In other multisensory temporal tasks, subjects are asked to make temporal order judgements (TOJ) as to which stimulus of the multisensory pairing appeared first (Zampini, Shore, and Spence, 2003). Participant's reports of synchrony across the various SOAs can be used to create response distributions and allow the derivation of two important measures of multisensory temporal function - namely the point of subjective simultaneity (PSS) and the temporal binding window (TBW).

The PSS is defined as the SOA at which perceived simultaneity is maximal. Interestingly, the PSS is not always at objective simultaneity (i.e., zero) but is usually found on the visual-leading side of the response distributions; see more discussion, Murray and Wallace (2011). In addition, as opposed to being a fixed construct, the PSS tends to vary dependent upon a variety of factors. These factors can be stimulus related (such as stimulus duration and intensity), (Boenke, Deliano, and Ohl, 2009; Jaśkowski, 1999; Sanford, 1971) task related (such as judging the onset vs the offset in an SJ task) (Wen, Opoku-Baah, Park, and Blake, 2020) or attention related (such as being asked to attend to one modality) (Schneider and Bavelier, 2003; Stelmach and Herdman, 1991; Zampini, Shore, and Spence, 2005).

On the other hand, the TBW is the range of stimulus onset asynchronies within which two stimuli are likely to be perceptually bound or integrated, thus, serving as a proxy measure for multisensory temporal acuity (Wallace and Stevenson, 2014). Experimentally, the TBW is derived as the interval from the SOA at which the probability of synchrony judgments exceeds a criterion (often a report of synchrony on 75% or 50% of trials). Recent studies have highlighted a few important characteristics about the TBW. First, like the PSS, the TBW is modulated by stimulus-related factors such as effectiveness or reliability (Fister et al., 2016) and stimulus complexity (e.g., flash-beep versus speech) (Stevenson and Wallace, 2013). For example, the TBW is smallest for simple stimuli such as flashes and beeps and largest for more complex and naturalistic multisensory stimuli such as speech (Stevenson and Wallace, 2013). In addition, the TBW has been shown to be broader in children than in adults, however, it continues to mature (i.e., narrow) late into development, that is the TBW begin to narrow after adolescence (Hillock-Dunn and Wallace, 2012; Hillock, Powers, and Wallace, 2011). Finally, the TBW is incredibly malleable, narrowing after a period of perceptual training with feedback (Powers, Hevey, and Wallace, 2012; Powers, Hillock, and Wallace, 2009) (Powers et al., 2009, Powers et al., 2012, Stevenson et al., 2013, Schlesinger et al., 2014).

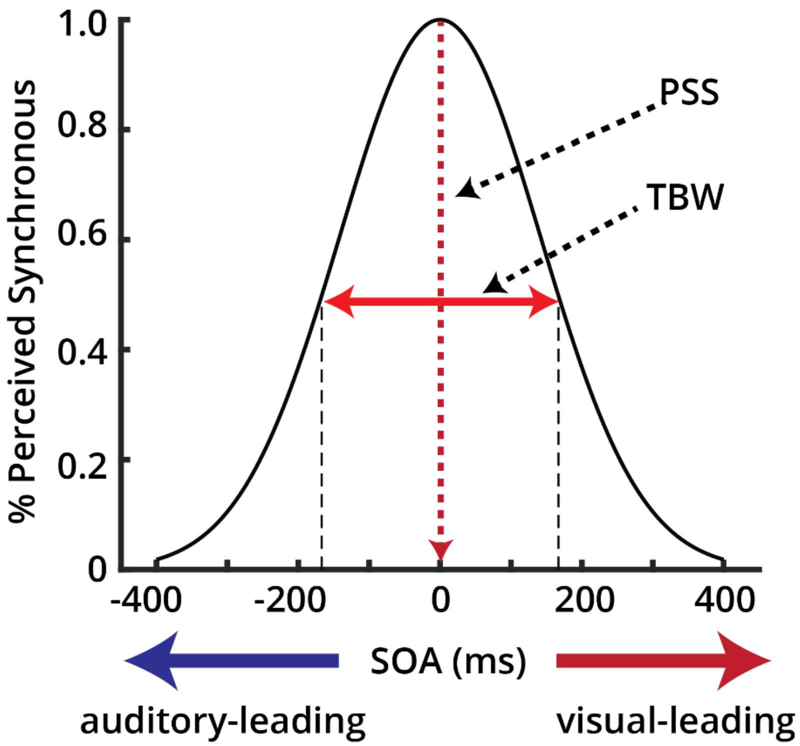

Figure 1. 1 Psychophysical Study of Multisensory Temporal

In a typical simultaneity judgment (SJ) task, participants are presented with paired multisensory stimuli (such as a visual flash and an auditory beep) with varying stimulus-onset asynchronies (SOAs) and are asked to determine whether the stimulus pair was “synchronous” or “asynchronous.” Participants’ reports of synchrony across the various SOAs can be used to create response distributions and allow the derivation of two important measures of multisensory temporal function. The first is the temporal binding window (TBW), defined as the range of stimulus onset asynchronies within which two stimuli are likely to be perceptually bound or integrated, thus serving as a proxy measure for multisensory temporal acuity. The second is the point of subjective simultaneity (PSS) which is defined as the SOA at which perceived simultaneity is maximal

1.3.3 Neural Correlates of Multisensory Temporal Perception

While the psychophysical study of multisensory temporal perception has relied on insightful constructs such as the TBW and the PSS, in the context of understanding the neural correlates of multisensory temporal perception, these constructs are not easily derivable. Instead, several studies have attempted to investigate the neural underpinnings of multisensory temporal perception by comparing the neural responses to synchronous and asynchronous stimuli using techniques as EEG, functional magnetic resonance imaging (fMRI), positron emission tomography (PET) etc.

1.3.3.1 Multisensory Temporal Processing Involves a Network of Brain Regions

Seminal work by Meredith and colleagues demonstrated that the greatest gains in neural activity in the superior colliculus (SC) were observed when the cross-modal stimuli occurred in close temporal proximity (Stein and Meredith, 1993). These findings were among the first to establish the role of temporal correspondence in shaping the outcomes of multisensory processing. More recently, in a study using fMRI recordings in human subjects, Calvert et al. (1999) showed that the blood-oxygenation-level-dependent (BOLD) signal was facilitated for temporally aligned audiovisual stimuli while being depressed for temporally misaligned stimuli. Consistent with Stein and Meredith (1993) work, they found that the SC was heavily involved and produced the most significant cross-modal gains and decrements. Apart from the SC, multisensory interactions were also observed in other polysensory areas such as the superior temporal sulcus (STS), intraparietal sulcus, insula, and several foci in the frontal lobe, including within the superior and ventromedial frontal gyri (Calvert et al., 1999). Other studies have also highlighted the role of unisensory or sensory-specific brains areas in multisensory temporal processing and perception. For instance, Noesselt et al. (2007) presented to participants, streams of auditory tones and visual flashes that were arranged to be either temporally coincident or non-coincident while measuring their BOLD response during the task. Their findings revealed that in the multisensory STS, BOLD responses increased for temporally coincident audiovisual streams and decreased for temporally non-coincident streams (Noesselt et al., 2007). Additionally, the dependence of BOLD response change on the temporal correspondence of the audiovisual stimuli was observed in the primary visual and auditory cortices (Noesselt et al., 2007). Interestingly, connectivity analyses indicated that when the streams were coincident, there was an enhanced influence from the multisensory STS on the primary sensory areas and vice versa for non-coincident streams (Noesselt et al., 2007). Collectively, these findings indicate that dealing with temporal correspondence in the brain involves a network of both sensory-specific and polysensory brain areas in humans.

1.3.3.2 The Role of Neural Oscillations in Multisensory Temporal Perception

Although the studies presented above did a great job in identifying the brain areas involved in the perception of multisensory temporal information, they do not provide the full picture in terms of understanding the neurocognitive processes involved during multisensory temporal perception. Such questions can be answered by the use of techniques such as electroencephalography (EEG) or magnetoencephalography (MEG) which have great temporal resolution allowing for the capturing of on-going brain activity in the order of milliseconds. Moreover, neural activity recorded from EEG and MEG can help elucidate whether underlying neurocognitive processes relate to early, low level sensory processes or late decisional factors.

Recently, several studies have been interested in understanding the role of neural oscillations in sensory processing at both the unisensory and multisensory levels. Generally, neural oscillatory activity recorded by EEG or MEG can be described as the synchronous waxing and waning of summed postsynaptic activity of large populations of neurons (Wang, 2010). Neural oscillations can be described by their frequency, amplitude, and phase. These different aspects of neural oscillations are known to relate to different processes in the brain (Keil and Senkowski, 2018). For instance, neural oscillations can be either evoked or induced. While evoked oscillations are phase locked to the onset of an external event and the summation over trials with identical phase can result in event-related potentials (ERPs), induced oscillations are not strictly stimulus dependent or phase-locked to the onset of the stimulus but can be modulated by cognitive processes as the task unfolds (Tallon-Baudry and Bertrand, 1999). Furthermore, phase-locking of oscillatory responses across several trials can be quantified as inter-trial coherence (Delorme and Makeig, 2004). An increase in ITC that is not accompanied by increased power can reflect the mechanism of phase resetting. While inter-trial coherence defines phase coherence across trials, phase coherence can occur between two brain regions if there is a correlation between the phases of the two signals over time. Importantly, phase coherence between two brain regions can indicate functional connectivity, (Fries, 2015) while a phase delay between oscillatory activities of two brain regions can indicate the direction of information flow between those regions (Maris, Fries, and van Ede, 2016).

Neural oscillations can be separated into different frequency bands with distinct amplitudes and phases namely delta (1-4Hz), theta (4-8Hz), alpha (8-12Hz), beta (12-30Hz) and gamma (>30Hz) (Mitra and Pesaran, 1999). With respect to the role of neural oscillations for cognitive processes, Engel and Fries (2010) suggested that fast oscillations indicate states of high arousal, whereas states of low arousal are primarily characterized by slow oscillations. Additionally, other studies have suggested that the frequency of neural oscillations are related to the direction of information flow that is either feed-forward (i.e., bottom-up information flow) or feedback projections (i.e., top-down information flow). For example, Fontolan et al. (2014) revealed that gamma-band activity was involved in feed-forward projections while alpha-and beta-band activity were involved in feedback projections. Furthermore, using laminar recordings from the macaque visual cortex, Van Kerkoerle et al. (2014) documented the role of gamma activity as a feed-forward rhythm, and the role of alpha-band activity as a feedback rhythm. The bottom-up nature of high frequency oscillations suggests information flow within localized networks whereas the top-down nature of low frequency oscillations is indicative of information flow across distant brains in a global network.

Evidence about the role of neural oscillatory activity in multisensory processing in general shows that neural oscillations in distinct frequency bands reflect different mechanisms of multisensory processing (Keil and Senkowski, 2018). The literature on this topic can be grouped into those looking at the role of pre-stimulus or the role of post-stimulus oscillatory activity in multisensory temporal perception. Concerning pre-stimulus activity, Yuan et al. (2016) investigated how oscillatory power modulated the subsequent discrimination of audiovisual asynchrony and discovered contrasting results for audiovisual stimuli with auditory leading and visual leading. For auditory leading sequences, they discovered that over-scalp high beta (20–28 Hz), parietal-occipital low beta (14–20 Hz), and high gamma oscillations (55–80 Hz) were significantly stronger when participants reported simultaneous rather than non-simultaneous. Conversely, for the visual leading sequences, beta and gamma band activity over a great portion of the electrodes were significantly weaker when participants reported simultaneous versus non-simultaneous. Focusing only on pre-stimulus alpha oscillations, Bastiaansen et al. (2020) discovered that pre-stimulus alpha power was lower when simultaneity judgments were inaccurate while individual alpha peak frequency correlated with making accurate simultaneity judgments. Using a temporal order judgment task, London et al. (2020) discovered that lower alpha power and higher instantaneous frequency were both separately associated with higher temporal sensitivity. These effects were located in highly overlapping clusters over posterior electrodes and were found to be interdependent. In another study, pre-stimulus frontal alpha power was found to correlate with the tendency to respond relative to an own idiosyncratic bias, with stronger a leading to responses matching the bias (Grabot, Kösem, Azizi, and Van Wassenhove, 2017). Comparing these findings establishes an interesting role of alpha power in multisensory temporal perception where parietal alpha power is associated with encoding of sensory information while frontal alpha power reflects mechanisms influencing perceptual decisional processes. Unlike the oscillatory power, the findings on the role of pre-stimulus oscillatory phase on subsequent multisensory temporal perception has been conflicting. For instance, while Bastiaansen et al. (2020) did not find any impact of pre-stimulus alpha phase on simultaneity judgments, Ikumi et al. (2019) found that the phase of pre-stimulus neural oscillations at 13 ± 2 Hz correlated with subjective simultaneity of otherwise identical sound-flash events. The discrepancy between these findings could lie in the differences in the approach used to determine the phase of the neural oscillations and the range of oscillatory frequencies that were explored. In all, further studies are needed to provide a better understanding of the role of pre-stimulus oscillatory phase in multisensory temporal processing.

In terms of post-stimulus neural oscillations, Senkowski et al. (2007) showed that synchronous audiovisual stimulation elicited increased evoked occipital gamma band power compared with asynchronous audiovisual stimulation. Covic et al. (2017) found increased occipital-parietal steady-state power during synchronous compared with asynchronous stimulation. In this same study, they observed an increased inter-trial phase coherence of both visual and auditory SS-EPs for temporally congruent audiovisual condition. Furthermore, Kambe et al. (2015) investigated the role of oscillatory phase for multisensory processing and found increased beta-band ITC in central and occipital electrodes only when subjects perceived audiovisual stimuli as synchronous but not when they perceived them as asynchronous. Moreover, phase reset has been shown to play a prominent role in multisensory processing (Kayser, Petkov, and Logothetis, 2008; Lakatos, Chen, O'Connell, Mills, and Schroeder, 2007). Taken together, these findings show that the power and phase of oscillatory activity in different frequency bands modulate multisensory temporal congruence, thereby establishing the important role of neural oscillation in shaping the multisensory temporal processes.

1.4 Neuroplasticity

Neuroplasticity simply refers to the ability of the nervous system to reorganize (Wong, 2012). Conventional views of plasticity held that it is a major force in the development of the nervous system, shaping structural and functional connectivity in response to environmental factors, but that the capacity for plasticity waned dramatically as maturation was complete (Berardi et al., 2000). Although plasticity is undoubtedly a critical factor in brain development, more contemporary evidence and views suggest that the adult nervous system remains markedly plastic. This view sat at the foundation of historical perspectives and that espoused that damage to the adult visual system was irremediable due to the highly elaborate cortical organization of the visual system and the significantly reduced plasticity in adulthood (Sabel, Henrich-Noack, Fedorov, and Gall, 2011).

1.4.1 Neuroplasticity in the Developing Brain

1.4.1.1 Early Visual Sensory Experience and the Development and Function of the Visual System in Animal and Human Models

The brain's plasticity peaks during a defined period in development, the so-called critical or sensitive period (Pascual-Leone et al., 2005; Berardi et al., 2000). During this period of plasticity, changes in sensory experience can engender morphological and functional changes that have enduring effects on sensory function.

Seminal work by Hubel and Wiesel discovered that early visual deprivation produced several changes in the structural and neurophysiological properties of neurons and circuits in the lateral geniculate nucleus and the visual cortex (Hubel and Wiesel, 1970; Wiesel and Hubel, 1963a, 1963b, 1965a, 1965b). They described that after a few weeks to months of visual deprivation, there was a reduction in the number and overall activity of neurons in these brain regions that responded to the deprived eye, as well as changes in receptive field properties of these cells and a reorganization of ocular dominance columns in favor of the open eye (Hubel and Wiesel, 1970; Wiesel and Hubel, 1963a, 1963b). Importantly, these deprivation-induced changes endured after years of restoring normal sensory experience to the deprived eye (Wiesel and Hubel, 1965b). Furthermore, these changes were found to be more profound in response to monocular vs. binocular deprivation, suggesting that an imbalance of input to the two eyes is a critical factor (Wiesel and Hubel, 1965a).

Even before Hubel and Wiesel conducted these experiments in cats, it was known that early disruption of visual sensory experience occurred in humans as an upshot of conditions such as congenital cataracts, strabismus, and issues of refractive error. The clinical condition that results from this maldevelopment of the visual system due to sensory deprivation is termed as amblyopia. Amblyopia is prevalent in about 2-4% of the population and currently is the leading cause of unilateral vision loss in children (Attebo et al., 1998; Brown et al., 2000; Buch, Vinding, La Cour, and Nielsen, 2001; Preslan and Novak, 1996). Clinically, it presents as a unilateral, or rarely bilateral, reduction in best-corrected visual acuity that cannot be explained solely by a structural eye abnormality. Amblyopia is associated with permanent defects in performance of visual perceptual tasks such as visual acuity, contrast detection, shape, and depth perception (Fine et al., 2002; Fine et al., 2003; Levi et al., 2011; Ostrovsky et al., 2006) and higher level processing such as global shape detection (Hess, Wang, Demanins, Wilkinson, and Wilson, 1999), real-world scene perception (Mirabella, Hay, and Wong, 2011), motion processing (Aaen-Stockdale and Hess, 2008; Simmers, Ledgeway, Hess, and McGraw, 2003), and feature counting (Sharma, Levi, and Klein, 2000). Interestingly, deficits in amblyopia does not only affect the deprived eye but also often extend to the fellow eye (Giaschi, Regan, Kraft, and Hong, 1992; Ho et al., 2005; Kovács, Polat, Pennefather, Chandna, and Norcia, 2000). In younger kids, refractive therapy and patching of the stronger eye is the conventional therapy to recover functional vision in the amblyopic eye (Papageorgiou, Asproudis, Maconachie, Tsironi, and Gottlob, 2019). However, this approach is not very effective in older kids and adults (Scheiman et al., 2005). In fact, there is a whole field of research devoted to the development therapies for the treatment of amblyopia in adults using techniques such as perceptual learning and repetitive transcranial magnetic stimulation (Levi and Li, 2009a, 2009b; Papageorgiou et al., 2019).

1.4.1.2 Early Visual Experience and the Development of Multisensory Temporal Perception in Animal Models

Interestingly, altered visual experience has been shown to influence the maturation of multisensory integration during development in both animal and humans. Using single unit recording of neurons in the superior colliculus (SC) of cats that were dark-reared and those that were reared in normal lighting conditions, Wallace and colleagues (Carriere et al., 2007; Wallace et al., 2004) documented several important findings on the role of visual sensory information in the development of the visual and multisensory systems. First, although no significant difference was found in the percentage of multisensory neurons in the SC of dark-reared and normal cats, the SC of normal cats showed a significantly higher number of unisensory visually response neurons compared to the SC of dark reared cats, implying that visual information is necessary for the development of visual neurons in the SC (Wallace et al., 2004). In terms of the structure and function of neurons, SC multisensory neurons of dark reared cats exhibited large receptive field size and a lack of response enhancement when stimulated with cross-modal stimuli that were weakly effective and spatially and temporally coincident (Wallace et al., 2004). Interestingly, the lack of substantial enhancement was observed not only in visually responsive multisensory neurons but also in non-visually responsive ones indicating the cross-modal nature of the light deprivation effect (Wallace et al., 2004). A similar investigation in the anterior ectosylvian sulcus (AES), a multisensory cortical area, of cats showed that dark rearing had a substantial impact on the integrative capabilities of multisensory AES neurons (Carriere et al., 2007). Specifically, there was a significant increase in the proportion of multisensory neurons that were modulated by, rather than driven by a second sensory modality. More importantly, there was a dramatic shift in the percentage of these modulated neurons in which the pairing of weakly effective, spatially and temporally coincident stimuli resulted in response depressions (Carriere et al., 2007). In normally reared animals, such combinations typically gave rise to robust response enhancements. Taken together, these results imply that while visual sensory experience is not necessary for the SC neurons to respond to both unisensory and multisensory stimuli, it is very necessary for the neurons regarding their ability to synthesize multisensory stimuli in a way that yields enhancement. In other words, visual experience is needed by SC neurons to learn how to integrate complex multisensory stimuli as they create and establish effective connectivity through simple Hebbian-based mechanisms.

1.4.1.3 Early Visual Experience and the Development of Multisensory Temporal Perception in Humans.

Perceptually, the effects of early visual deprivation in humans can impair audiovisual interactions, most notably in the temporal domain (Chen et al., 2017; Putzar et al., 2007; Putzar et al., 2010). For instance, patients who experienced visual deprivation owing to congenital cataracts during the first 0.3 to 29 months of their life showed wider TBWs when compared to age-matched typically developed controls during performance of an audiovisual simultaneity judgment task (Chen et al., 2017). Interestingly, the fellow eye (non-deprived eye; having similar visual acuity to the eyes of normally sighted individuals) had comparable TBW size to the deprived eye indicating that this poor performance cannot be attributed to poor visual acuity in the deprived eye**.** In addition, while monocularly deprived patients exhibited widened TBWs characteristic of an immature audiovisual integration system, binocularly deprived patients showed widened TBW that did not match any stage in typical development (Chen et al., 2017). As mentioned earlier, visual deprivation during development can result from several conditions and not only from cataracts. A study investigated the impact of visual deprivation caused by anisometropia (unequal refractive errors in both eyes) or strabismus (patent lateral deviation of one eye) or both in early life on multisensory temporal perception and found enlarged temporal binding windows (Richards et al., 2017) similar finding as (Chen et al., 2017). Together, these findings indicate that normal visual experience is crucial for the development of both visual and multisensory (i.e., visual-nonvisual) functions.

Besides perceptual evidence, neuroimaging studies have revealed large-scale cross-modal reorganization of the visual cortex in patients who were visually deprived briefly during their early life (Collignon et al., 2015; Guerreiro, Putzar, and Röder, 2015). For instance, when fMRI was used to compare responses of visual and auditory cortical areas to visual, auditory, and audio-visual stimulation, cataract-reversal patients, unlike normally sighted controls, did not exhibit multisensory integration in auditory areas (Guerreiro et al., 2015). Furthermore, cataract-reversal patients, but not normally sighted controls, exhibited lower visual cortical processing within visual cortex during audio-visual stimulation than during visual stimulation (Guerreiro et al., 2015). These results indicate that congenital visual deprivation affects the capability of cortical areas to integrate cross-modal inputs in humans, possibly because visual processing is suppressed during cross-modal stimulation.

While these studies did not investigate multisensory temporal perception directly, the knowledge gleaned from them could help understand the relationship between early visual deprivation and the effects on multisensory temporal perception.

1.4.2 Cross-modal Plasticity after Permanent Visual Sensory Loss

The interaction among sensory modalities and the inherent plasticity in the brain suggest that a complete deprivation of one sensory modality can have striking effects on the structural and functional organization of the brain and on perception and behavior. Intuitively, these changes could reflect adaptive mechanisms occurring as a result of an extreme reliance and effective use of the remaining modalities by sensory deprived individuals (Bavelier and Neville, 2002). In fact, several studies have documented that visually deprived individuals delineate superior abilities than normally sighted individuals during the performance of tasks involving the spared modalities such as auditory discrimination (Gougoux et al., 2004) and localization (Lessard, Paré, Lepore, and Lassonde, 1998; RoÈder et al., 1999; Voss et al., 2004) tasks and tactile discrimination tasks (Alary et al., 2008; Goldreich and Kanics, 2003). This compensatory plasticity is remarkably evident following early-onset blindness (Bedny, Pascual-Leone, Dravida, and Saxe, 2012; Voss, Gougoux, Zatorre, Lassonde, and Lepore, 2008; Wan, Wood, Reutens, and Wilson, 2010) probably owing to the enhanced degree of malleability during the early stages of life.

Whether in response to early or late-onset visual deprivation, cross-modal plasticity changes have been documented in three brain loci, namely: sensory areas representing the intact sensory modalities, multisensory brain areas and the visual cortex (Bavelier and Neville, 2002). For instance, significant cortical reorganization in somatosensory (Sterr et al., 1998) and auditory (Elbert et al., 2002) cortices has been discovered in blind individuals compared to normally sighted subjects using MEG recordings. In addition, Büchel et al. (1998) discovered that individuals with both congenital and adult-onset (after puberty) blindness showed significant activations in the inferior parietal lobe, a multisensory cortical area, during Braille reading. Besides changes in the spared sensory cortices and in multisensory cortical regions, cross-modal plasticity occurs in the visual cortex. For instance, the visual cortex is significantly activated in blind individuals during performance of haptic tasks such as braille reading (C Büchel, 1998; Burton et al., 2002; Held, Freedman, and Harris, 1996), object identification (Pietrini et al., 2004) and electro-tactile stimulation of the tongue (Ptito, Moesgaard, Gjedde, and Kupers, 2005). In addition, auditory tasks including sound-source discrimination (Voss et al., 2008), auditory motion discrimination (Poirier et al., 2006) and speech perception (Röder, Stock, Bien, Neville, and Rösler, 2002) activate visual cortex in blind individuals.

Notably, cross-modal plasticity that involves the visual cortex are functionally relevant to compensatory behaviors observed after visual deprivation. Some studies have provided evidence for a causal relationship between the cross-modal plastic changes in visual cortex and compensatory behaviors using transcranial magnetic stimulation (TMS) to disrupt activity in the visual cortex. For instance, disrupting occipital cortical activity using TMS diminishes performance on braille reading (Cohen et al., 1997; Kupers et al., 2007) and verb-generation (Amedi et al., 2004) and also the use of auditory-based sensory substitution devices (Merabet et al., 2009) in blind individuals. Taken together, the evidence provided under the topics discussed above highlights the extreme interdependence among the different sensory modalities and shows that significant cross-modal plasticity accompanies forms of visual impairment.

1.4.3 Neuroplasticity in the Adult Brain

Despite the reduced plasticity in the brain after the closure of the critical period, the adult brain retains marked plasticity that can influence perception and behavior (Berardi et al., 2000; Sabel, 2008). Plasticity in the adult brain has been shown mostly in two ways: through sensory deprivation and perceptual learning. The following sections explore the perceptual and neural evidence of plasticity following sensory deprivation and perceptual learning in adult humans.

1.4.3.1 Sensory Deprivation

1.4.3.1.1 Influences of Short-term Monocular Deprivation on Visual Perception in the Adult Human Brain: Behavioral Evidence

Occluding one eye with a translucent patch for 150 minutes has been shown to boost the perceptual dominance of the deprived eye during subsequent binocular rivalry and increases the apparent contrast of viewed gratings (Lunghi, Burr, and Morrone, 2011). This enhanced performance in the deprived eye has been demonstrated across different visual tasks (Zhou, Clavagnier, and Hess, 2013). For instance, Zhou et al. (2013) occluded one eye of participants for 150 minutes and discovered that the deprived eye’s contribution to phase combination, global motion coherence, and contrast matching was enhanced. Additionally, they found that 150 minutes of monocular deprivation increased the contrast sensitivity in the deprived eye while reducing the sensitivity in the non-deprived eye (Zhou et al., 2013). This finding is interesting because unlike binocular rivalry where the two eyes direct compete for conscious perception of the target of interest at the same time, in the contrast detection task, one eye was tested at a time and did not involve this direct competition for the visual target.

Moreover, the effect of monocular deprivation on perceptional abilities of the deprived and non-deprived eyes does not depend on the nature of deprivation (Chadnova, Reynaud, Clavagnier, and Hess, 2017; H.-W. Kim, Kim, and Blake, 2017; Zhou et al., 2013). For instance, using either light diffuser, which transmits light but not pattern, or opaque patching, which excludes pattern and mean luminance information produces similar increase in the perceptual abilities of the deprived eye after monocular deprivation (Chadnova et al., 2017; Zhou et al., 2013). In an interesting study, a perceptual form of deprivation yielded comparable plasticity to using the conventional translucent patch or light diffuser (Kim et al., 2017). In their study, they measured binocular rivalry dynamics before and immediately after a short period of continuous flash suppression (Kim et al., 2017). Continuous flash suppression is a more potent form of binocular rivalry where one eye views a continuous stream of complex visual images that dominates perceptual awareness for long periods and the other views a low contrast sinusoidal grating (Tsuchiya and Koch, 2005; Yang and Blake, 2012). Compared to findings after translucent patch, they discovered that CFS produced more robust enhancement in dominance durations of the grating viewed by the deprived eye (Kim et al., 2017).

In an earlier study, the perceptual boost observed in the deprived eye after 150 minutes of monocular deprivation remained at 78% after 15 minutes and was still significant even after 90 minutes after deprivation (Lunghi et al., 2011). This raised an important question concerning whether the magnitude and stability of the effect was dependent on the duration of deprivation. To answer this question, a study patched participants for durations between 15 and 300 minutes (about 5 hours) and discovered an infinitesimal increase in the magnitude of the deprivation-induced effect (Min, Baldwin, Reynaud, and Hess, 2018). Put differently, a 20-fold increase in duration of monocular deprivation only produced a 25% difference in the magnitude of the deprivation-induced effect (Min, Baldwin, Reynaud, and Hess, 2018). In addition, they observed that longer durations of monocular deprivation did not result in a slower recovery of ocular dominance to baseline (Min et al., 2018). Together, these findings imply that monocular deprivation induces a form of homeostatic mechanism that responds instantaneously in an all-or-none fashion.

1.4.3.1.2 Influences of Short-term Monocular Deprivation on Visual Perception in the Adult Human Brain: Neural Mechanisms

Earlier studies investigating the neural mechanisms underlying the effects of monocular deprivation documented that resting GABAergic inhibition played a crucial in triggering ocular dominance plasticity and modulating the onset and offset of the critical period (Fagiolini and Hensch, 2000; Hensch et al., 1998). Other studies also demonstrated that after the critical period, manipulating cortical excitation and inhibition imbalance through GABAergic inhibition in adult animals affected the susceptibility of ocular dominance to monocular deprivation (Harauzov et al., 2010; Morishita, Miwa, Heintz, and Hensch, 2010; Pizzorusso et al., 2002; Vetencourt et al., 2008). For instance, monocular deprivation was able to induce ocular dominance plasticity in adult rats treated with 3-mercaptopropionic acid, a substance that reduces GABA release while no effect of monocular deprivation was observed in saline treated adult rats (Harauzov et al., 2010). Together, these findings highlight the massive role GABAergic inhibition plays in the development, maturation, and continuous functioning of the visual system.

What then are the mechanisms involved in this rapid plasticity induced by a brief period of monocular deprivation in adult humans? According to Lunghi and her colleagues (Lunghi et al., 2011), the short-term monocular deprivation-induced effects they observed reflected a form of rapid homeostatic plasticity to visual deprivation which may be mediated by an upregulation of contrast gain-control mechanisms of the deprived eye to optimize the attenuated or absent visual information. Previously, cortical gain-control mechanisms have been implicated in a number of visual processes including binocular rivalry (Ling and Blake, 2012) and eye dominance during binocular combination (Ding and Sperling, 2006). Some studies have suggested that cortical inhibition may underlie cortical gain-control mechanisms (Katzner, Busse, and Carandini, 2011; Maria Concetta Morrone, Burr, and Speed, 1987; M Concetta Morrone and Burr, 1986). Considering that both cortical inhibition and gain control mechanisms are implicated in phenomena such as binocular rivalry, it is plausible to hypothesize that, cortical inhibition may also be involved in cortical gain-control mechanisms. To provide evidence for the role of GABAergic inhibition in this rapid plasticity induced by monocular deprivation, Lunghi et al. (2015) measured the levels of GABA concentration in the primary visual cortex of adult humans before and after 2.5 hours of monocular deprivation using ultra-high field 7T magnetic resonance spectroscopy. While monocular deprivation did not alter GABA concentration levels in a control parietal area in the brain, in the primary visual cortex, resting GABA concentration decreased after monocular deprivation. Importantly, their findings revealed that the deprivation-induced reduction in GABA concentration was strongly correlated with the deprived eye perceptual boost measured by binocular rivalry implying a common mechanism (Lunghi, Emir, et al., 2015). Consistent with the findings from the animal models, these findings in humans show a critical role of GABAergic inhibition and cortical excitation inhibition imbalance in triggering visual plasticity induced by short-term monocular deprivation.

Considering the contrasting nature of the perceptual effects of short-term monocular deprivation on the deprived and non-deprived eyes, and the implication of cortical excitation inhibition imbalance in monocular deprivation induced plasticity, it is plausible to hypothesize that after monocular deprivation; neural processes potentiate response of the deprived compared to non-deprived eye. However, the nature of the previous investigations could not piece out the relative contributions of the two eyes to the observed effect. Moreover, the narrowed focus of the initial studies on primary visual cortex limited our understanding of the contribution of different stages of visual processing (lower vs. higher visual areas) in this monocular-deprivation effect. Using electrophysiological techniques, Lunghi et al. (2015) recorded visual pattern evoked potentials before and after 150 min of monocular deprivation in adult humans and discovered that the amplitude of the C1 component located on Cz increased for the deprived eye and decreased for the non-deprived eye. Source localization analysis of this amplitude revealed primary cortex as the origin of the effect, ruling out high visual cortical areas (Lunghi, Berchicci, et al., 2015). In addition, findings from frequency analysis revealed an increase in alpha power at Cz for the deprived eye and a decrease for the non-deprived eye. In another study, monocular deprivation boosted the BOLD response to the deprived eye with the boost being strongest in the V1 followed by V2, V3 and V4 but absent in V3a and hMT+ (Binda et al., 2018). Moreover, in that same study, the boost in BOLD signal induced by monocular deprivation primarily occurred for high spatial frequency gratings (Binda et al., 2018). Importantly, these findings highlight a couple of things. First, like the findings from the GABA concentration and the evoked potential studies, the main cortical hub for the effects of monocular deprivation induced plasticity is the primary visual cortex. Secondly, they highlight the fact that ventral or parvocellular pathway (which include area V4) may be more implicated in this plasticity than the dorsal or magnocellular (which include area hMT+) (Binda et al., 2018). Confirming the role of parvocellular pathway in this short-term monocular deprivation-induced effect, Lunghi et al. (2013) demonstrated that the deprivation-induced effect lasted longer for chromatic gratings, information that processed in the parvocellular pathway, compared with achromatic gratings after re-exposure to binocular vision. Collectively, these findings imply that monocular deprivation reduces GABAergic inhibition in the visual areas, which may change the cortical excitation and inhibition imbalance and consequently, increase cortical excitability for the deprived eye and opposite for the non-deprived eye.

1.4.3.2 Perceptual Learning

1.4.3.2.1 Introduction to Perceptual Learning

Perceptual learning is the process through which repeated training results in task improvements. Perceptual learning requires explicit long-term effort to produce an improvement in the trained domain or ability (Li, 2016). Typically, it involves hours of daily training that are repeated for several days to weeks before a desirable result is achieved (Li, 2016). In vision research, perceptual learning has been employed to improve performance in detecting and discriminating stimuli with different levels of contrast, texture, and motion (Ball and Sekuler, 1982; Fiorentini and Berardi, 1980; Karni and Sagi, 1991; Mayer, 1983).

1.4.3.2.2 Multisensory Perceptual Learning

The multisensory nature of the world sets up an advantageous platform for learning. This has led to the view that our brains have evolved to learn and function optimally in the multisensory world (Shams and Seitz, 2008). Although all multisensory learning paradigms employ cross-modal information in an effort to reorganize cortical circuitry, based on the intended perceptual outcome, they can be grouped into two major categories; paradigms that aim to enhance 1) some aspect of cross-modal integration such as temporal integration and 2) perfection in a particular unisensory modality.

The first category is multisensory-based multisensory perceptual learning (MPL). This category of multisensory learning has revealed the malleability of the multisensory system, most notably in the temporal domain. As described earlier, the temporal relationship between the paired inputs of a multisensory stimulus strongly influences how these inputs are integrated. Indeed, this temporal dependency has resulted in a great deal of work oriented around the construct of the multisensory temporal binding window (TBW) (Wallace and Stevenson, 2014). The TBW defines a measure of the range of stimuli-onset asynchronies (SOAs) within which the two cross-modal stimuli can be perceived as constituting a single stimulus event, i.e., can be integrated (Wallace and Stevenson, 2014). Generally, the TBW is wider for visual leading stimuli, likely reflecting the natural statistics of environmental stimuli and in which visual information travels much faster than auditory information (Wallace and Stevenson, 2014). The narrow auditory-leading window may be mediated in part by auditory processes with high temporal resolution while visually driven processes with relatively slower temporal resolution may contribute to the wider visual-leading window (van Wassenhove, 2013). Powers et al. (2009) assessed the plasticity of the TBW by training individuals for an hour a day for five days on a two-alternative forced choice audiovisual simultaneity judgment task with visual feedback. After the training, a significant narrowing of the TBW was observed, with the changes being greatest on the visual leading side of the distribution (Powers et al., 2009). Building off of this result, a recent study suggested that the two sides of the distribution (i.e. the auditory- and visual-leading) may reflect the involvement of differential processes that follow different learning trajectories, with the visual-leading processing having significantly greater flexibility (Cecere, Gross, and Thut, 2016). From that study, training subjects with visual-leading stimuli improved sensitivity to asynchrony for visual leading, but not auditory leading SOAs. Conversely, training with auditory-leading stimuli resulted in no improvements for either the auditory-leading or the visual-leading SOA conditions (Cecere et al., 2016). Taken together, it is plausible to conclude that the flexibility observed for visual-leading conditions may reflect significant plasticity of the visual processing machinery.

The second category of perceptual learning paradigms, which is more relevant to vision rehabilitation, comprises multisensory paradigms that enhance plasticity and facilitate learning in a particular sense (unisensory-based MPL). Remarkably, such multisensory perceptual learning protocols provide significantly improved unisensory learning outcomes than when training is implemented in the unisensory condition. For example, Seitz et al. (2006) trained two groups of individuals on congruent auditory-visual training (AV) and visual only training (V) protocols using a coherent motion detection task. Comparing performance on only visual trials, the AV group achieved a greater detection accuracy within the first session and across all training sessions than the V group (Seitz et al., 2006). In addition, learning using the multisensory paradigm required fewer sessions to reach an asymptotic level (a reduction of ~60%) than V group (Seitz et al., 2006). Interestingly, performance on the hardest-difficulty trials (i.e., when the coherence of the motion stimuli was near random) improved significantly only after audiovisual training, highlighting a guiding principle of multisensory interactions known as inverse effectiveness (Holmes, 2009; Stein and Meredith, 1993). Lastly, training with incongruent audiovisual stimuli did not result in enhanced after-training performance, indicating the importance of stimulus congruence during the learning process (Kim et al., 2008).

1.4.3.2.3 Neural Evidence of Multisensory Perceptual Learning

Shams and Seitz (2008) suggested that multisensory training that facilitates unisensory learning may alter how information is represented in unisensory structures in two ways. First, it is possible that signals from brain areas of the auxiliary modality can modulate neuronal activity in the brain area of the task-relevant modality leading to enhanced plasticity and changes in the neuronal properties of these neurons over the course of learning. Consequently, presenting unisensory stimuli after training will strongly activate the unisensory structures leading to enhanced perception. Another possibility is that the presence of information from the auxiliary sensory modality during training can lead to enhanced connectivity between the unisensory areas or altered processing in the multisensory structures. Here, both mechanisms are likely to lead to enhanced perception stemming from an activation of a wider network of brain areas during subsequent presentation of unisensory stimuli.

Indeed, while these two mechanisms—changes in unisensory structures and changes in multisensory structures—may co-occur to facilitate learning, many of the studies that have investigated the neural mechanisms underlying multisensory-facilitated learning have shown evidence for the latter. For instance, Zilber et al. (2014) investigated the neural mechanisms involved in multisensory learning using a modified version of a visual motion discrimination task and magnetoencephalography (MEG) recordings of various brain regions in human subjects. They developed three training protocols administered in a between-subjects design: use of visual random-dot-kinematograms (V) only, use of auditory textures that share coherence properties and are coupled with the visual random-dot-kinematograms (AV), and use of auditory noise that was uncorrelated with visual random-dot-kinematograms (AVn). After twenty minutes of training, the AV group showed a significant increase in discrimination sensitivity compared to the other groups. Furthermore, comparing pre- and post-training neurometric curves of hMT+ activity revealed an increase in the sensitivity of hMT+ responses to motion of the random-dot-kinematograms in the AV group only. Corroborating Shams and Seitz (2008) later proposal, multisensory perceptual training recruited a network of auditory (auditory cortex) and multisensory areas (middle (mSTS) and posterior (pSTS) superior temporal sulcus) in addition to the visual areas (hMT+, V4, ITC) involved in the visual perceptual training (Zilber et al., 2014). Of particular note, post-training enhancements in visual motion discrimination sensitivity were strongly correlated with post-training enhancement in neural activity in areas mSTS and pSTS for only the AV group, suggesting an instrumental role for these regions (Zilber et al., 2014).

Strengthening of connectivity between these networks of brain regions through Hebbian-like plasticity mechanisms may also be responsible for the benefits observed after multisensory training. For example, through effective connectivity analyses of fMRI data, Powers and colleagues (Powers et al., 2012) showed that training subjects on an audiovisual learning protocol significantly increased the interregional coupling between auditory and visual cortices and pSTS. This is not surprising as short-term concurrent presentation of auditory and visual stimuli can increase effective connectivity between multisensory and unisensory brain areas. For example, pairing sound with low intensity visual targets resulted in a significant increase in effective connectivity between visual and auditory thalamic relay nuclei (i.e., LGN and MGN) and multisensory area STS (Noesselt et al., 2010). Strikingly, the enhanced effective connectivity between these areas correlated strongly with subjects' behavioral performance

1.5 References

Aaen-Stockdale, C., and Hess, R. F. (2008). The amblyopic deficit for global motion is spatial scale invariant. *Vision research, 48*(19), 1965-1971.

Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current biology, 14*(3), 257-262.

Alary, F., Goldstein, R., Duquette, M., Chapman, C. E., Voss, P., and Lepore, F. (2008). Tactile acuity in the blind: a psychophysical study using a two-dimensional angle discrimination task. *Experimental brain research, 187*(4), 587-594.

Amedi, A., Floel, A., Knecht, S., Zohary, E., and Cohen, L. G. (2004). Transcranial magnetic stimulation of the occipital pole interferes with verbal processing in blind subjects. *Nature neuroscience, 7*(11), 1266.

Arnold, P., and Hill, F. (2001). Bisensory augmentation: A speechreading advantage when speech is clearly audible and intact. *British Journal of Psychology, 92*(2), 339-355.

Attebo, K., Mitchell, P., Cumming, R., Smith, W., Jolly, N., and Sparkes, R. (1998). Prevalence and causes of amblyopia in an adult population. *Ophthalmology, 105*(1), 154-159.

Ball, K., and Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science, 218*(4573), 697-698.

Bastiaansen, M., Berberyan, H., Stekelenburg, J. J., Schoffelen, J. M., and Vroomen, J. (2020). Are alpha oscillations instrumental in multisensory synchrony perception? *Brain Research, 1734*, 146744.

Bavelier, D., and Neville, H. J. (2002). Cross-modal plasticity: where and how? *Nature Reviews Neuroscience, 3*(6), 443.

Bedny, M., Pascual-Leone, A., Dravida, S., and Saxe, R. (2012). A sensitive period for language in the visual cortex: distinct patterns of plasticity in congenitally versus late blind adults. *Brain and language, 122*(3), 162-170.

Berardi, N., Pizzorusso, T., and Maffei, L. (2000). Critical periods during sensory development. *Current Opinion in Neurobiology, 10*(1), 138-145. doi: 10.1016/s0959-4388(99)00047-1

Bertelson, P. (1999). Ventriloquism: A case of cross-modal perceptual grouping *Advances in psychology* (Vol. 129, pp. 347-362): Elsevier.

Bertelson, P., and Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception and psychophysics, 29*(6), 578-584.

Binda, P., Kurzawski, J. W., Lunghi, C., Biagi, L., Tosetti, M., and Morrone, M. C. (2018). Response to short-term deprivation of the human adult visual cortex measured with 7T BOLD. *Elife, 7,* e40014.

Blau, V., van Atteveldt, N., Ekkebus, M., Goebel, R., and Blomert, L. (2009). Reduced neural integration of letters and speech sounds links phonological and reading deficits in adult dyslexia. *Current Biology, 19*(6), 503-508.

Boenke, L. T., Deliano, M., and Ohl, F. W. (2009). Stimulus duration influences perceived simultaneity in audiovisual temporal-order judgment. *Experimental brain research, 198*(2-3), 233-244.

Brown, S. A., Weih, L. M., Fu, C. L., Dimitrov, P., Taylor, H. R., and McCarty, C. A. (2000). Prevalence of amblyopia and associated refractive errors in an adult population in Victoria, Australia. *Ophthalmic epidemiology, 7*(4), 249-258.

Buch, H., Vinding, T., La Cour, M., and Nielsen, N. V. (2001). The prevalence and causes of bilateral and unilateral blindness in an elderly urban Danish population. The Copenhagen City Eye Study. *Acta Ophthalmologica Scandinavica, 79*(5), 441-449.

Büchel, C. (1998). Functional neuroimaging studies of Braille reading: cross-modal reorganization and its implications. *Brain: a journal of neurology, 121*(7), 1193-1194.

Büchel, C., Price, C., Frackowiak, R., and Friston, K. (1998). Different activation patterns in the visual cortex of late and congenitally blind subjects. *Brain: a journal of neurology, 121*(3), 409-419.

Burton, H., Snyder, A. Z., Conturo, T. E., Akbudak, E., Ollinger, J. M., and Raichle, M. E. (2002). Adaptive changes in early and late blind: a fMRI study of Braille reading. *Journal of neurophysiology, 87*(1), 589-607.

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

Carriere, B. N., Royal, D. W., Perrault, T. J., Morrison, S. P., Vaughan, J. W., Stein, B. E., and Wallace, M. T. (2007). Visual deprivation alters the development of cortical multisensory integration. *Journal of neurophysiology, 98*(5), 2858-2867.

Cecere, R., Gross, J., and Thut, G. (2016). Behavioural evidence for separate mechanisms of audiovisual temporal binding as a function of leading sensory modality. *European Journal of Neuroscience, 43*(12), 1561-1568.

Chadnova, E., Reynaud, A., Clavagnier, S., and Hess, R. F. (2017). Short-term monocular occlusion produces changes in ocular dominance by a reciprocal modulation of interocular inhibition. *Scientific reports, 7,* 41747.

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., and Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Comput Biol, 5*(7), e1000436.

Chen, Y.-C., Lewis, T. L., Shore, D. I., and Maurer, D. (2017). Early binocular input is critical for development of audiovisual but not visuotactile simultaneity perception. *Current Biology, 27*(4), 583-589.

Cohen, L. G., Celnik, P., Pascual-Leone, A., Corwell, B., Faiz, L., Dambrosia, J., . . . Catala, M. D. (1997). Functional relevance of cross-modal plasticity in blind humans. *Nature, 389*(6647), 180.

Collignon, O., Dormal, G., De Heering, A., Lepore, F., Lewis, T. L., and Maurer, D. (2015). Long-lasting cross-modal cortical reorganization triggered by brief postnatal visual deprivation. *Current Biology, 25*(18), 2379-2383.

Congdon, N., O'Colmain, B., Klaver, C., Klein, R., Muñoz, B., Friedman, D. S., . . . Mitchell, P. (2004). Causes and prevalence of visual impairment among adults in the United States. *Archives of Ophthalmology (Chicago, Ill.: 1960), 122*(4), 477-485.

Covic, A., Keitel, C., Porcu, E., Schröger, E., and Müller, M. M. (2017). Audio-visual synchrony and spatial attention enhance processing of dynamic visual stimulation independently and in parallel: a frequency-tagging study. *Neuroimage, 161*, 32-42.

De Winkel, K. N., Katliar, M., and Bülthoff, H. H. (2017). Causal inference in multisensory heading estimation. *PloS one, 12*(1), e0169676.

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods, 134*(1), 9-21.

Diederich, A., and Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Perception and psychophysics, 66*(8), 1388-1404.

Ding, J., and Sperling, G. (2006). A gain-control theory of binocular combination. *Proceedings of the National Academy of Sciences, 103*(4), 1141-1146.

Elbert, T., Sterr, A., Rockstroh, B., Pantev, C., Müller, M. M., and Taub, E. (2002). Expansion of the tonotopic area in the auditory cortex of the blind. *Journal of Neuroscience, 22*(22), 9941-9944.

Engel, A. K., and Fries, P. (2010). Beta-band oscillations—signalling the status quo? *Current opinion in neurobiology, 20*(2), 156-165.

Erber, N. P. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of speech and hearing research, 12*(2), 423-425.

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

Ernst, M. O., and Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in cognitive sciences, 8*(4), 162-169.

Fagiolini, M., and Hensch, T. K. (2000). Inhibitory threshold for critical-period activation in primary visual cortex. *Nature, 404*(6774), 183-186.

Fine, I., Smallman, H. S., Doyle, P., and MacLeod, D. I. (2002). Visual function before and after the removal of bilateral congenital cataracts in adulthood. *Vision Research, 42*(2), 191-210.

Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., . . . MacLeod, D. I. (2003). Long-term deprivation affects visual perception and cortex. *Nature Neuroscience*, *6*(9), 915.

Fiorentini, A., and Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency.

Fister, J. K., Stevenson, R. A., Nidiffer, A. R., Barnett, Z. P., and Wallace, M. T. (2016). Stimulus intensity modulates multisensory temporal processing. *Neuropsychologia, 88*, 92-100.

Fontolan, L., Morillon, B., Liegeois-Chauvel, C., and Giraud, A.-L. (2014). The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nature communications, 5*(1), 1-10.

Foxe, J. J., and Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *Neuroreport, 16*(5), 419-423.

Frassinetti, F., Bolognini, N., and Làdavas, E. (2002). Enhancement of visual perception by cross-modal visuo-auditory interaction. *Experimental brain research, 147*(3), 332-343.

Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron, 88*(1), 220-235.

Ghazanfar, A. A., and Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in cognitive sciences, 10*(6), 278-285.

Giaschi, D., Regan, D., Kraft, S., and Hong, X. (1992). Defective processing of motion-defined form in the fellow eye of patients with unilateral amblyopia. *Investigative ophthalmology and visual science, 33*(8), 2483-2489.

Goldreich, D., and Kanics, I. M. (2003). Tactile acuity is enhanced in blindness. *Journal of Neuroscience, 23*(8), 3439-3445.

Gougoux, F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R. J., and Belin, P. (2004). Neuropsychology: pitch discrimination in the early blind. *Nature, 430*(6997), 309.

Grabot, L., Kösem, A., Azizi, L., and Van Wassenhove, V. (2017). Prestimulus alpha oscillations and the temporal sequencing of audiovisual events. *Journal of cognitive neuroscience, 29*(9), 1566-1582.

Grant, K. W., and Seitz, P.-F. (2000). The use of visible speech cues for improving auditory detection of spoken sentences. *The journal of the acoustical society of america, 108*(3), 1197-1208.

Guerreiro, M. J., Putzar, L., and Röder, B. (2015). The effect of early visual deprivation on the neural bases of multisensory processing. *Brain, 138*(6), 1499-1504.

Hairston, W. D., Laurienti, P. J., Mishra, G., Burdette, J. H., and Wallace, M. T. (2003). Multisensory enhancement of localization under conditions of induced myopia. *Experimental brain research, 152*(3), 404-408.

Harauzov, A., Spolidoro, M., DiCristo, G., De Pasquale, R., Cancedda, L., Pizzorusso, T., . . . Maffei, L. (2010). Reducing intracortical inhibition in the adult visual cortex promotes ocular dominance plasticity. *Journal of Neuroscience, 30*(1), 361-371.

Held, R., Freedman, S., and Harris, C. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature, 380*, 11.

Hensch, T. K., Fagiolini, M., Mataga, N., Stryker, M. P., Baekkeskov, S., and Kash, S. F. (1998). Local GABA circuit control of experience-dependent plasticity in developing visual cortex. *Science, 282*(5393), 1504-1508.

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *Journal of Experimental Psychology, 63*(3), 289.

Hess, R. F., Wang, Y.-Z., Demanins, R., Wilkinson, F., and Wilson, H. R. (1999). A deficit in strabismic amblyopia for global shape detection. *Vision research, 39*(5), 901-914.

Hillock-Dunn, A., and Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental science, 15*(5), 688-696.

Hillock, A. R., Powers, A. R., and Wallace, M. T. (2011). Binding of sights and sounds: age-related changes in multisensory temporal processing. *Neuropsychologia, 49*(3), 461-467.

Ho, C. S., Giaschi, D. E., Boden, C., Dougherty, R., Cline, R., and Lyons, C. (2005). Deficient motion perception in the fellow eye of amblyopic children. *Vision research, 45*(12), 1615-1627.

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

Hubel, D. H., and Wiesel, T. N. (1970). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *The Journal of physiology, 206*(2), 419-436.

Ikumi, N., Torralba, M., Ruzzoli, M., and Soto-Faraco, S. (2019). The phase of pre-stimulus brain oscillations correlates with cross-modal synchrony perception. *European Journal of Neuroscience, 49*(2), 150-164.

Jaśkowski, P. (1999). Reaction time and temporal-order judgment as measures of perceptual latency: The problem of dissociations *Advances in psychology* (Vol. 129, pp. 265-282): Elsevier.

Kambe, J., Kakimoto, Y., and Araki, O. (2015). Phase reset affects auditory-visual simultaneity judgment. *Cognitive neurodynamics, 9*(5), 487-493.

Karni, A., and Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences, 88*(11), 4966-4970.

Katzner, S., Busse, L., and Carandini, M. (2011). GABAA inhibition controls response gain in visual cortex. *Journal of Neuroscience, 31*(16), 5931-5941.

Kayser, C., Petkov, C. I., and Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex, 18*(7), 1560-1574.

Keil, J., and Senkowski, D. (2018). Neural oscillations orchestrate multisensory processing. *The Neuroscientist, 24*(6), 609-626.

Kim, H.-W., Kim, C.-Y., and Blake, R. (2017). Monocular perceptual deprivation from interocular suppression temporarily imbalances ocular dominance. *Current Biology, 27*(6), 884-889.

Kim, R. S., Seitz, A. R., and Shams, L. (2008). Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS One, 3*(1), e1532.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., and Shams, L. (2007). Causal inference in multisensory perception. *PLoS One, 2*(9), e943.

Kovács, I., Polat, U., Pennefather, P. M., Chandna, A., and Norcia, A. M. (2000). A new test of contour integration deficits in patients with a history of disrupted binocular experience during visual development. *Vision research, 40*(13), 1775-1783.

Kupers, R., Pappens, M., de Noordhout, A. M., Schoenen, J., Ptito, M., and Fumal, A. (2007). rTMS of the occipital cortex abolishes Braille reading and repetition priming in blind subjects. *Neurology, 68*(9), 691-693.

Kwakye, L. D., Foss-Feig, J. H., Cascio, C. J., Stone, W. L., and Wallace, M. T. (2010). Altered auditory and multisensory temporal processing in autism spectrum disorders. *Frontiers in integrative neuroscience, 4*.

Lakatos, P., Chen, C.-M., O'Connell, M. N., Mills, A., and Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron, 53*(2), 279-292.

Lessard, N., Paré, M., Lepore, F., and Lassonde, M. (1998). Early-blind human subjects localize sound sources better than sighted subjects. *Nature, 395*(6699), 278.

Levi, D. M., and Li, R. W. (2009a). Improving the performance of the amblyopic visual system. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1515), 399-407.

Levi, D. M., and Li, R. W. (2009b). Perceptual learning as a potential treatment for amblyopia: a mini-review. *Vision research, 49*(21), 2535-2549.

Levi, D. M., McKee, S. P., and Movshon, J. A. (2011). Visual deficits in anisometropia. *Vision Research, 51*(1), 48-57.

Li, W. (2016). Perceptual Learning: Use-Dependent Cortical Plasticity. *Annual Review of Vision Science, 2*, 109-130.

Ling, S., and Blake, R. (2012). Normalization regulates competition for visual awareness. *Neuron, 75*(3), 531-540.

London, R. E., Benwell, C. S., Cecere, R., Quak, M., Thut, G., and Talsma, D. (2020). EEG Alpha Power Predicts the Temporal Sensitivity of Multisensory Integration. *bioRxiv*.

Lovelace, C. T., Stein, B. E., and Wallace, M. T. (2003). An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Cognitive brain research, 17*(2), 447-453.

Lunghi, C., and Alais, D. (2013). Touch interacts with vision during binocular rivalry with a tight orientation tuning. *PLoS One, 8*(3), e58754.

Lunghi, C., Berchicci, M., Morrone, M. C., and Di Russo, F. (2015). Short-term monocular deprivation alters early components of visual evoked potentials. *The Journal of physiology, 593*(19), 4361-4372.

Lunghi, C., Burr, D. C., and Morrone, C. (2011). Brief periods of monocular deprivation disrupt ocular balance in human adult visual cortex. *Current Biology, 21*(14), R538-R539.

Lunghi, C., Burr, D. C., and Morrone, M. C. (2013). Long-term effects of monocular deprivation revealed with binocular rivalry gratings modulated in luminance and in color. *Journal of vision, 13*(6), 1-1.

Lunghi, C., Emir, U. E., Morrone, M. C., and Bridge, H. (2015). Short-term monocular deprivation alters GABA in the adult human visual cortex. *Current Biology, 25*(11), 1496-1501.

MacDonald, J., and McGurk, H. (1978). Visual influences on speech perception processes. *Perception and psychophysics, 24*(3), 253-257.

Magnotti, J. F., Ma, W. J., and Beauchamp, M. S. (2013). Causal inference of asynchronous audiovisual speech. *Frontiers in psychology, 4*, 798.

Mallick, D. B., Magnotti, J. F., and Beauchamp, M. S. (2015). Variability and stability in the McGurk effect: contributions of participants, stimuli, time, and response type. *Psychonomic bulletin and review, 22*(5), 1299-1307.

Maris, E., Fries, P., and van Ede, F. (2016). Diverse phase relations among neuronal rhythms and their potential function. *Trends in neurosciences, 39*(2), 86-99.

Mayer, M. J. (1983). Practice improves adults' sensitivity to diagonals. *Vision research, 23*(5), 547-550.

McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746-748.

Merabet, L. B., Battelli, L., Obretenova, S., Maguire, S., Meijer, P., and Pascual-Leone, A. (2009). Functional recruitment of visual cortex for sound encoded object identification in the blind. *Neuroreport, 20*(2), 132.

Min, S. H., Baldwin, A. S., Reynaud, A., and Hess, R. F. (2018). The shift in ocular dominance from short-term monocular deprivation exhibits no dependence on duration of deprivation. *Scientific reports, 8*(1), 1-9.

Mirabella, G., Hay, S., and Wong, A. M. (2011). Deficits in perception of images of real-world scenes in patients with a history of amblyopia. *Archives of Ophthalmology, 129*(2), 176-183.

Mitra, P. P., and Pesaran, B. (1999). Analysis of dynamic brain imaging data. *Biophysical journal, 76*(2), 691-708.

Møller, C., Højlund, A., Bærentsen, K. B., Hansen, N. C., Skewes, J. C., and Vuust, P. (2018). Visually induced gains in pitch discrimination: Linking audio-visual processing with auditory abilities. *Attention, Perception, and Psychophysics, 80*(4), 999-1010.

Morishita, H., Miwa, J. M., Heintz, N., and Hensch, T. K. (2010). Lynx1, a cholinergic brake, limits plasticity in adult visual cortex. *Science, 330*(6008), 1238-1240.

Morrone, M. C., Burr, D., and Speed, H. (1987). Cross-orientation inhibition in cat is GABA mediated. *Experimental brain research, 67*(3), 635-644.

Morrone, M. C., and Burr, D. C. (1986). Evidence for the existence and development of visual inhibition in humans. *Nature, 321*(6067), 235-237.

Murai, Y., and Yotsumoto, Y. (2018). Optimal multisensory integration leads to optimal time estimation. *Scientific reports, 8*(1), 1-11.

Murray, M. M., and Wallace, M. T. (2011). *The neural bases of multisensory processes*: CRC Press.

Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H.-J., and Driver, J. (2007). Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *Journal of Neuroscience*, *27*(42), 11431-11441.

Noesselt, T., Tyll, S., Boehler, C. N., Budinger, E., Heinze, H.-J., and Driver, J. (2010). Sound-induced enhancement of low-intensity vision: multisensory influences on human sensory-specific cortices and thalamic bodies relate to perceptual enhancement of visual detection sensitivity. *Journal of Neuroscience, 30*(41), 13609-13623.

Odegaard, B., and Shams, L. (2016). The brain’s tendency to bind audiovisual signals is stable but not general. *Psychological science, 27*(4), 583-591.

Odegaard, B., Wozny, D. R., and Shams, L. (2017). A simple and efficient method to enhance audiovisual binding tendencies. *PeerJ, 5*, e3143.

Ostrovsky, Y., Andalman, A., and Sinha, P. (2006). Vision following extended congenital blindness. *Psychological Science, 17*(12), 1009-1014.

Papageorgiou, E., Asproudis, I., Maconachie, G., Tsironi, E. E., and Gottlob, I. (2019). The treatment of amblyopia: current practice and emerging trends. *Graefe's Archive for Clinical and Experimental Ophthalmology, 257*(6), 1061-1078.

Parise, C. V., Knorre, K., and Ernst, M. O. (2014). Natural auditory scene statistics shapes human spatial hearing. *Proceedings of the National Academy of Sciences, 111*(16), 6104-6108.

Parise, C. V., Spence, C., and Deroy, O. (2016). Understanding the correspondences: Introduction to the special issue on cross-modal correspondences. *Multisensory Research, 29*(1-3), 1-6.

Pascual-Leone, A., Amedi, A., Fregni, F., and Merabet, L. B. (2005). The plastic human brain cortex. *Annu. Rev. Neurosci., 28*, 377-401.

Philippi, T. G., van Erp, J. B., and Werkhoven, P. J. (2008). Multisensory temporal numerosity judgment. *Brain research, 1242*, 116-125.

Pietrini, P., Furey, M. L., Ricciardi, E., Gobbini, M. I., Wu, W.-H. C., Cohen, L., . . . Haxby, J. V. (2004). Beyond sensory images: Object-based representation in the human ventral pathway. *Proceedings of the National academy of Sciences of the United States of America, 101*(15), 5658-5663.

Pizzorusso, T., Medini, P., Berardi, N., Chierzi, S., Fawcett, J. W., and Maffei, L. (2002). Reactivation of ocular dominance plasticity in the adult visual cortex. *Science, 298*(5596), 1248-1251.

Poirier, C., Collignon, O., Scheiber, C., Renier, L., Vanlierde, A., Tranduy, D., . . . De Volder, A. (2006). Auditory motion perception activates visual motion areas in early blind subjects. *Neuroimage, 31*(1), 279-285.

Pöppel, E., and Artin, T. T. (1988). *Mindworks: Time and conscious experience*: Harcourt Brace Jovanovich.

Pöppel, E., Schill, K., and von Steinbüchel, N. (1990). Sensory integration within temporally neutral systems states: a hypothesis. *Naturwissenschaften, 77*(2), 89-91.

Powers, A. R., Hevey, M. A., and Wallace, M. T. (2012). Neural correlates of multisensory perceptual learning. *Journal of Neuroscience, 32*(18), 6263-6274.

Powers, A. R., Hillock, A. R., and Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *The Journal of neuroscience, 29*(39), 12265-12274.

Preslan, M. W., and Novak, A. (1996). Baltimore vision screening project. *Ophthalmology, 103*(1), 105-109.

Ptito, M., Moesgaard, S. M., Gjedde, A., and Kupers, R. (2005). Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind. *Brain, 128*(3), 606-614.

Putzar, L., Goerendt, I., Lange, K., Rösler, F., and Röder, B. (2007). Early visual deprivation impairs multisensory interactions in humans. *Nature Neuroscience, 10*(10), 1243.

Putzar, L., Hötting, K., and Röder, B. (2010). Early visual deprivation affects the development of face recognition and of audio-visual speech perception. *Restorative neurology and neuroscience, 28*(2), 251-257.

Reisberg, D., Mclean, J., and Goldfield, A. (1987). Easy to hear but hard to understand: a lip-reading advantage with intact auditory stimuli.

Richards, M. D., Goltz, H. C., and Wong, A. M. (2017). Alterations in audiovisual simultaneity perception in amblyopia. *PloS one, 12*(6), e0179516.

Röder, B., Stock, O., Bien, S., Neville, H., and Rösler, F. (2002). Speech processing activates visual cortex in congenitally blind humans. *European Journal of Neuroscience, 16*(5), 930-936.

RoÈder, B., Teder-SaÈlejaÈrvi, W., Sterr, A., RoÈsler, F., Hillyard, S. A., and Neville, H. J. (1999). Improved auditory spatial tuning in blind humans. *Nature, 400*(6740), 162.

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., and Foxe, J. J. (2006). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex, 17*(5), 1147-1153.

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

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Molholm, S., Javitt, D. C., and Foxe, J. J. (2007). Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophrenia research, 97*(1), 173-183.

Sabel, B. A. (2008). Plasticity and restoration of vision after visual system damage: an update. *Restorative neurology and neuroscience, 26*(4, 5), 243-247.

Sabel, B. A., Henrich-Noack, P., Fedorov, A., and Gall, C. (2011). Vision restoration after brain and retina damage: the "residual vision activation theory". *Prog Brain Res, 192*(1), 199-262.

Sanford, A. (1971). Effects of changes in the intensity of white noise on simultaneity judgements and simple reaction time. *The Quarterly Journal of Experimental Psychology, 23*(3), 296-303.

Scheiman, M. M., Hertle, R. W., Beck, R. W., Edwards, A. R., Birch, E., Cotter, S. A., . . . Donahue, S. (2005). Randomized trial of treatment of amblyopia in children aged 7 to 17 years. *Archives of ophthalmology (Chicago, Ill.: 1960), 123*(4), 437-447.

Schneider, K. A., and Bavelier, D. (2003). Components of visual prior entry. *Cognitive psychology, 47*(4), 333-366.

Seitz, A. R., Kim, R., and Shams, L. (2006). Sound facilitates visual learning. *Current Biology, 16*(14), 1422-1427.

Senkowski, D., Talsma, D., Grigutsch, M., Herrmann, C. S., and Woldorff, M. G. (2007). Good times for multisensory integration: effects of the precision of temporal synchrony as revealed by gamma-band oscillations. *Neuropsychologia, 45*(3), 561-571.

Shams, L., Kamitani, Y., and Shimojo, S. (2000). What you see is what you hear. *Nature, 408*(6814), 788-788.

Shams, L., and Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in cognitive sciences, 12*(11), 411-417.

Sharma, V., Levi, D. M., and Klein, S. A. (2000). Undercounting features and missing features: evidence for a high-level deficit in strabismic amblyopia. *Nature neuroscience, 3*(5), 496-501.

Simmers, A. J., Ledgeway, T., Hess, R. F., and McGraw, P. V. (2003). Deficits to global motion processing in human amblyopia. *Vision research, 43*(6), 729-738.

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

Stein, B. E., and Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience, 9*(4), 255-266.

Stelmach, L. B., and Herdman, C. M. (1991). Directed attention and perception of temporal order. *Journal of Experimental Psychology: Human Perception and Performance, 17*(2), 539.

Sterr, A., Müller, M. M., Elbert, T., Rockstroh, B., Pantev, C., and Taub, E. (1998). Perceptual correlates of changes in cortical representation of fingers in blind multifinger Braille readers. *Journal of Neuroscience, 18*(11), 4417-4423.

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., and Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *Journal of Neuroscience, 34*(3), 691-697.

Stevenson, R. A., and Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Experimental brain research, 227*(2), 249-261.

Su, Y.-H. (2014). Visual enhancement of auditory beat perception across auditory interference levels. *Brain and Cognition, 90*, 19-31.

Sumby, W. H., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The journal of the acoustical society of america, 26*(2), 212-215.

Szycik, G. R., Münte, T. F., Dillo, W., Mohammadi, B., Samii, A., Emrich, H., and Dietrich, D. (2009). Audiovisual integration of speech is disturbed in schizophrenia: an fMRI study. *Schizophrenia research, 110*(1-3), 111-118.

Tallon-Baudry, C., and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in cognitive sciences, 3*(4), 151-162.

Tsuchiya, N., and Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature neuroscience, 8*(8), 1096-1101.

Van der Burg, E., Cass, J., Olivers, C. N., Theeuwes, J., and Alais, D. (2010). Efficient visual search from synchronized auditory signals requires transient audiovisual events. *PloS one, 5*(5), e10664.

Van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., Van Der Togt, C., and Roelfsema, P. R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences, 111*(40), 14332-14341.

van Wassenhove, V. (2013). Speech through ears and eyes: interfacing the senses with the supramodal brain.

Vetencourt, J. F. M., Sale, A., Viegi, A., Baroncelli, L., De Pasquale, R., O'Leary, O. F., . . . Maffei, L. (2008). The antidepressant fluoxetine restores plasticity in the adult visual cortex. *Science*.

Voss, P., Gougoux, F., Zatorre, R. J., Lassonde, M., and Lepore, F. (2008). Differential occipital responses in early-and late-blind individuals during a sound-source discrimination task. *Neuroimage*, *40*(2), 746-758.

Voss, P., Lassonde, M., Gougoux, F., Fortin, M., Guillemot, J.-P., and Lepore, F. (2004). Early-and late-onset blind individuals show supra-normal auditory abilities in far-space. *Current Biology, 14*(19), 1734-1738.

Wallace, M. T., Perrault, T. J., Hairston, W. D., and Stein, B. E. (2004). Visual experience is necessary for the development of multisensory integration. *Journal of Neuroscience, 24*(43), 9580-9584.

Wallace, M. T., and Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia, 64*, 105-123.

Wallace, M. T., and Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia, 64*, 105-123. doi: 10.1016/j.neuropsychologia.2014.08.005

Wan, C. Y., Wood, A. G., Reutens, D. C., and Wilson, S. J. (2010). Early but not late-blindness leads to enhanced auditory perception. *Neuropsychologia*, *48*(1), 344-348.

Wang, X.-J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological reviews, 90*(3), 1195-1268.

Wen, P., Opoku-Baah, C., Park, M., and Blake, R. (2020). Judging Relative Onsets and Offsets of Audiovisual Events. *Vision*, *4*(1), 17.

Wiesel, T. N., and Hubel, D. H. (1963a). Effects of visual deprivation on morphology and physiology of cells in the cat's lateral geniculate body. *Journal of neurophysiology*, *26*(6), 978-993.

Wiesel, T. N., and Hubel, D. H. (1963b). Single-cell responses in striate cortex of kittens deprived of vision in one eye. *Journal of neurophysiology*, *26*(6), 1003-1017.

Wiesel, T. N., and Hubel, D. H. (1965a). Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. *Journal of neurophysiology*, *28*(6), 1029-1040.

Wiesel, T. N., and Hubel, D. H. (1965b). Extent of recovery from the effects of visual deprivation in kittens. *Journal of neurophysiology*, *28*(6), 1060-1072.

Wong, A. M. (2012). New concepts concerning the neural mechanisms of amblyopia and their clinical implications. *Canadian Journal of Ophthalmology/Journal Canadien d'Ophtalmologie*, *47*(5), 399-409.

Yang, E., and Blake, R. (2012). Deconstructing continuous flash suppression. *Journal of vision*, *12*(3), 8-8.

Yuan, X., Li, H., Liu, P., Yuan, H., and Huang, X. (2016). Pre-stimulus beta and gamma oscillatory power predicts perceived audiovisual simultaneity. *International Journal of Psychophysiology*, *107*, 29-36.

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

Zampini, M., Shore, D. I., and Spence, C. (2003). Audiovisual temporal order judgments. *Experimental brain research*, *152*(2), 198-210.

Zampini, M., Shore, D. I., and Spence, C. (2005). Audiovisual prior entry. *Neuroscience letters*, *381*(3), 217-222.

Zhou, J., Clavagnier, S., and Hess, R. F. (2013). Short-term monocular deprivation strengthens the patched eye's contribution to binocular combination. *Journal of vision*, *13*(5), 12-12.

Zilber, N., Ciuciu, P., Gramfort, A., Azizi, L., and Van Wassenhove, V. (2014). Supramodal processing optimizes visual perceptual learning and plasticity. *Neuroimage*, *93*, 32-46.

Zou, H., Müller, H. J., and Shi, Z. (2012). Non-spatial sounds regulate eye movements and enhance visual search. *Journal of vision*, *12*(5), 2-2.

Chapter 2

Brief Period of Monocular Deprivation Drives Changes in Audiovisual Temporal Perception

The contents of this chapter are adapted from

Opoku-Baah, C. and Wallace, M.T., 2020. Brief period of monocular deprivation drives changes in audiovisual temporal perception. Journal of Vision, 20(8), pp.8-8.

2.1 Abstract

The human brain retains a striking degree of plasticity into adulthood. Recent studies have demonstrated that a short period of altered visual experience (via monocular deprivation) can change the dynamics of binocular rivalry in favor of the deprived eye, a compensatory action thought to be mediated by an upregulation of cortical gain control mechanisms. Here, we sought to better understand the impact of monocular deprivation on multisensory abilities; specifically examining audiovisual temporal perception. Using an audiovisual simultaneity judgment task, we discovered that 90 minutes of monocular deprivation produced opposing effects on the temporal binding window (TBW) depending upon the eye used in the task. Thus, in those who performed the task with their deprived eye there was a narrowing of the TBW, whereas in those performing the task with their non-deprived eye there was a widening of the TBW. The effect was short-lived, being observed only in the first 10 minutes of post-deprivation testing. These findings indicate that changes in

visual experience in the adult can rapidly impact multisensory perceptual processes, a finding that has important clinical implications for those patients with adult-onset visual deprivation and for therapies founded on monocular deprivation.

2.2 Introduction

The brain possesses a remarkable degree of plasticity that enables it to reorganize to meet the demands of the changing environment (Pascual-Leone, Amedi, Fregni, and Merabet, 2005). This intrinsic property of the nervous system peaks during a defined period in development, the so-called critical or sensitive period (Berardi, Pizzorusso, and Maffei, 2000). During this period of plasticity, changes in sensory experience can engender morphological and functional changes that have enduring effects on sensory function.

Seminal work by Hubel and Wiesel discovered that early visual deprivation produced several changes in the structural and neurophysiological properties of neurons and circuits in the lateral geniculate nucleus and the visual cortex (Hubel and Wiesel, 1970; Wiesel and Hubel, 1963, 1965a, 1965b). They described that after a few weeks to months of visual deprivation, there was a reduction in the number and overall activity of neurons in these brain regions that responded to the deprived eye, as well as changes in receptive field properties of these cells and a reorganization of ocular dominance columns in favor of the open eye (Hubel and Wiesel, 1970; Wiesel and Hubel, 1963). Importantly, these deprivation-induced changes endured after years of restoring normal sensory experience to the deprived eye (Wiesel and Hubel, 1965b). Furthermore, these changes were found to be more profound in response to monocular vs. binocular deprivation, suggesting that an imbalance of input to the two eyes is a critical factor (Wiesel and Hubel, 1965a). In humans, early visual deprivation can occur through conditions such as untreated congenital cataracts and uncorrected refractive errors and is associated with permanent defects in performance of visual perceptual tasks such as visual acuity, contrast detection, shape, and depth perception (Fine, Smallman, Doyle, and MacLeod, 2002; Fine et al., 2003; Levi, McKee, and Movshon, 2011; Ostrovsky, Andalman, and Sinha, 2006).

In recent years, the study of multisensory processing has received a great deal of attention. A growing body of compelling evidence supports the interaction and/or integration of information from different sensory modalities at various brain levels, especially in areas, which were once considered to be classically “unisensory” (Foxe and Schroeder, 2005; Ghazanfar and Schroeder, 2006; Stein and Stanford, 2008). Integrating information from multiple senses has several benefits including striking response gains at the neural level (Stein and Meredith, 1993; Stein and Stanford, 2008), as well as enhanced detection (Frassinetti, Bolognini, and Làdavas, 2002; Lovelace, Stein, and Wallace, 2003), discrimination (Ernst and Banks, 2002) and localization (Zou, Müller, and Shi, 2012), and speeding response times (Diederich and Colonius, 2004) at the perceptual level. These neural and perceptual benefits depend on the physical characteristics of the combined stimuli, with the temporal proximity between the different cues that make up a multisensory stimulus being one of the important factors. Intuitively, natural events that originate from a common source are likely to arrive at the brain centers in close temporal proximity and thus, integrating them may enhance perceptual saliency and performance. Indeed, the largest neural and perceptual gains are achieved when the two inputs making up a multisensory stimulus are close in time (Meredith, Nemitz, and Stein, 1987).

Extending this work into the psychophysical arena, a number of studies have focused on characterizing the window of stimulus onset asynchronies (SOAs) within which two stimuli are likely to be integrated or perceptually "bound" (Wallace and Stevenson, 2014). This window, termed the temporal binding window (TBW), is essentially a measure of multisensory temporal acuity with a larger window indicating poorer acuity and an increased likelihood that the brain will bind events that do not originate from a common source. Over the years, several studies have revealed some notable characteristics of the TBW. For example, the size of the TBW varies significantly across subjects (Stevenson, Zemtsov, and Wallace, 2012). This individual variability has been linked with multisensory integrative abilities such as individual susceptibility to the McGurk effect, scored as an illusory perception of /da/ or /ta/ when an observer fuses an auditory syllable /ba/ and a visual syllable /ga/ (Stevenson et al., 2012). In addition, the TBW size has been shown to be larger in clinical populations such as those with autism and schizophrenia (Foss-Feig et al., 2010; Stevenson et al., 2014; Stevenson et al., 2017; Noel, Stevenson, and Wallace, 2018). As with the adult plasticity seen in the visual system, multisensory temporal function has been shown to be highly malleable. For example, a short interval of perceptual training can significantly narrow the TBW (Powers, Hevey, and Wallace, 2012; Powers, Hillock, and Wallace, 2009).

Interestingly, altered visual experience has been shown to influence the maturation of multisensory integration during development. For example, work in animal models has shown that early visual deprivation impairs the ability of neurons to synthesize multisensory information (Carriere et al., 2007; Wallace, Perrault, Hairston, and Stein, 2004). In humans, imaging studies have revealed large-scale cross-modal reorganization of the visual cortex in patients who were visually deprived briefly during their early life (Collignon et al., 2015; Guerreiro, Putzar, and Röder, 2015). Perceptually, the effects of early visual deprivation in humans can also impair audiovisual interactions, most notably in the temporal domain (Chen, Lewis, Shore, and Maurer, 2017; Putzar, Goerendt, Lange, Rösler, and Röder, 2007; Putzar, Hötting, and Röder, 2010). For instance, patients who experienced visual deprivation owing to congenital cataracts showed wider TBWs when compared to age-matched typically developed controls during performance of an audiovisual simultaneity judgment task (Chen et al., 2017). In addition, while monocularly deprived patients exhibited widened TBWs characteristic of an immature audiovisual integration system, binocularly deprived patients showed widened TBW that did not match any stage in typical development (Chen et al., 2017). Together, these findings indicate that normal visual experience is crucial for the development of both visual and multisensory (i.e., visual-nonvisual) functions.

While the impact of atypical early visual experience on visual and multisensory functions have been well-established, the question of whether a short-term change in visual experience in adults can drive plastic changes in multisensory temporal function remains unanswered. Recent studies have shown that in spite of the reduced plasticity in the brain after the closure of the critical period, the adult brain still retains marked plasticity that can impact perception and behavior (Berardi et al., 2000; Sabel, 2008). One example of such adult plasticity is the perceptual effect of short-term monocular deprivation. Occluding one eye with a translucent patch for 150 minutes boosts the perceptual dominance of the deprived eye during subsequent binocular rivalry and increases the apparent contrast of viewed gratings (Lunghi, Burr, and Morrone, 2011). According to Lunghi and her colleagues (Lunghi et al., 2011), this finding reflects a form of rapid homeostatic plasticity to visual deprivation and is mediated by an upregulation of contrast gain-control mechanisms of the deprived eye in an attempt to optimize the attenuated or absent visual information. This enhanced performance in the deprived eye has been demonstrated across different visual tasks such as phase combination, global motion coherence, and contrast matching and detection (Zhou, Clavagnier, and Hess, 2013) and is independent of the nature of deprivation (Chadnova, Reynaud, Clavagnier, and Hess, 2017; Kim, Kim, and Blake, 2017; Zhou et al., 2013). Given the evidence of the malleability of the TBW to effects of visual deprivation-induced plasticity during development, we hypothesized that a brief period of monocular deprivation will impact the TBW in adult humans through homeostatic plasticity. Such a result would reinforce the importance of adult visual experience in maintaining not only visual representations, but also in maintaining multisensory (e.g., audiovisual) representations.

2.3 Methods

2.3.1 Participants

Twenty Vanderbilt undergraduates were recruited to perform an audiovisual simultaneity judgment task using either their deprived or non-deprived eye before and after 90 minutes of monocular deprivation. Each subject was randomly assigned to either the deprived eye group (N = 10, 3 males, mean age (in years) = 19.1, SD = 3.0) or the non-deprived eye group (N = 10, 4 males, mean age = 18.9 (in years), SD = 1.2). All students had normal or corrected-to-normal vision and normal hearing. Normal vision was confirmed using a visual acuity task. Participants gave informed consent before being allowed to participate and received course credits for their participation. All recruitment and experimental procedures were approved by the Vanderbilt University Institutional Review Board (IRB). This work was carried out in accordance with the Declaration of Helsinki.

2.3.2 Stimulus and Apparatus

The experimental stimuli were generated and presented using MATLAB (Math Works Inc., Natick, MA) software with the Psychophysics Toolbox Version 3 (Brainard and Vision, 1997; Pelli, 1997). The visual stimulus was displayed on a gamma-corrected monitor (21-inch Asus LCD) with 120-Hz refresh rate while the auditory stimulus was presented binaurally through headphones (Sennheiser HD559). The visual stimulus was a gabor patch (sigma = 0.5 degrees), which had a spatial frequency of 4 cycles per degree, subtended an angle of 6 degrees and was presented at the center of fixation. The gabor patch was displayed at 20% contrast on a uniform gray background with a luminance of 10cd/m2. The auditory stimulus comprised a white noise burst at ~75dB and with a sampling frequency of 41kHz. Both stimuli were presented for 4 frame periods equaling 33ms. While the intensity of the auditory stimulus was linearly ramped up and down, each for 25% of the stimulus duration, a similar effect was achieved with the visual stimulus by presenting the first and last frames at half the actual contrast of the stimulus. A Minolta Chroma Meter CS-100 and a sound level meter were used to verify the luminance and sound intensity levels respectively. The durations of all visual and auditory stimuli, as well as the SOAs, were confirmed using a Hameg 507 oscilloscope with a photovoltaic cell and microphone.

2.3.3 Procedure

Participants performed the experiment inside a WhisperRoom^{TM} (SE 2000 Series) with their forehead and chin placed comfortably against a HeadSpot^{TM} (University of Houston Optometry). The whisper room was only lit by the background luminance of the screen. Before the experiment began, each subject was randomly assigned to either the deprived eye group or the non-deprived eye group. The eye to-be deprived was randomly chosen for each subject and counterbalanced across all subjects. The deprived eye group performed the task with their deprived eye while the non-deprived eye group performed the task with their non-deprived eye before and after 90 minutes of monocular deprivation with an opaque eye patch (Figure 2.1A). Both the pre and the post deprivation sessions comprised two blocks of an unspeeded audiovisual simultaneity judgment task. A block took approximately 10 minutes to complete. During each block, the subject viewed a gabor patch and heard an auditory white noise burst after a brief fixation period lasting between 600 and 1000ms (Figure 2.1B). The white circular dot that marked the fixation period was always prese5nt and subjects were instructed to always fixate at it. The onsets of the two stimuli were separated by 13 parametrically varied SOAs (in ms): -500, -350, -200, -150, -100, -50, 0, 50, 100, 150, 200, 350, 500. Negative and positive SOA values corresponded to auditory-preceding-vision and vision-preceding-auditory SOAs respectively. Each SOA was presented 20 times (totaling 260 trials) in randomized fashion. After each presentation, participants accurately delivered their response by pressing ‘1’ on the keyboard if the pair of audiovisual stimuli was synchronous or by pressing ‘2’, if the pair was asynchronous. Each subject was given two brief initial practice sessions, each consisting of 9 SOAs and 10 trials per SOA. The rationale for this initial practice was to first, exclude subjects who could not perform adequately well on the task and second, reduce familiarization effects during the main experiment. No feedback was provided on the correctness of their responses during the main experiment.

2.3.4 Analysis

Each participant completed 2 pre and 2 post deprivation blocks of the audiovisual simultaneity judgment task. Proportions of synchrony reports as a function of SOA were computed for each block and each participant. We then pooled (i.e., averaged) the data from the 2 pre-deprivation blocks into 1 combined pre-deprivation block and likewise, the 2 post-deprivation blocks into 1 post-deprivation block. These data were then fitted separately with a single-term Gaussian distribution model (through MATLAB fit.m) whose amplitude, mean and standard deviation (SD) were free to vary (Noel et al., 2018; Simon, Noel, and Wallace, 2017). Unlike the mean and the SD parameters, the range of possible amplitude values was bound between 0 and 1. The mean and the SD of the best fitting model were then taken as the point of subjective simultaneity (PSS) and the temporal window of simultaneity (TBW) respectively. The averaged r2 values across participants in both groups showed reasonable fits to the data for the combined pre-deprivation block (0.92±0.01) and the combined post-deprivation block (0.91±0.02). In order to determine whether the deprivation effect differed between the first and second post-deprivation blocks, the Gaussian distribution described above was fitted to the data for the two post-deprivation blocks separately and their PSS and TBW measures were derived. Again, the averaged r2 values across participants for the two post-deprivation blocks were satisfactory: post-block 1 (0.89±0.03) and post-block 2 (0.87±0.04). Next, we computed the effect of deprivation on the performance measures for each participant by normalizing the post-deprivation measure (either combined or divided) to the baseline pre-deprivation measure. For the TBW, the relative change calculation ((post-pre)/pre * 100%) was used while for the PSS, an absolute change (post – pre) computation was used because of the inclusion of negative numbers.

The following statistical analyses were then carried out for each performance measure using the JASP software version 0.11 (JASP Team, 2018). First, an independent samples t-test was done to determine whether participants in both groups were comparable in their pre-deprivation performance. Second, a one-sample t-test against zero (baseline) was used to determine whether the computed deprivation effect (i.e., absolute or relative change) was statistically significant for each eye group. Third, we conducted an independent samples t-test to determine whether the absolute change in PSS differed between the deprived and non-deprived eye groups. For the TBW, a one-way analysis of covariance test (ANCOVA) was done to determine whether the relative change in TBW differed between the deprived and non-deprived eye groups while adjusting for the individual variability in baseline TBW. For the divided blocks, we conducted multiple t-tests (for 4 comparisons in total) to determine 1) whether the deprivation effect differed between the first and second post-deprivation blocks for each group and 2) whether the deprivation effect differed between the two groups for each post-deprivation block. Note that in the latter scenario, a one-way ANCOVA test was conducted to adjust for the between-subject differences in pre-deprivation TBW. Correlation analyses were done using Pearson's correlation. All statistical analyses were two-tailed and significance level, alpha, was set to 0.05. In the case of multiple tests, alpha was adjusted with a Bonferroni-Holm correction. Furthermore, the statistical results were reported in both frequentist (i.e., p values) and Bayesian (i.e., Bayes factors) frameworks of hypothesis testing.

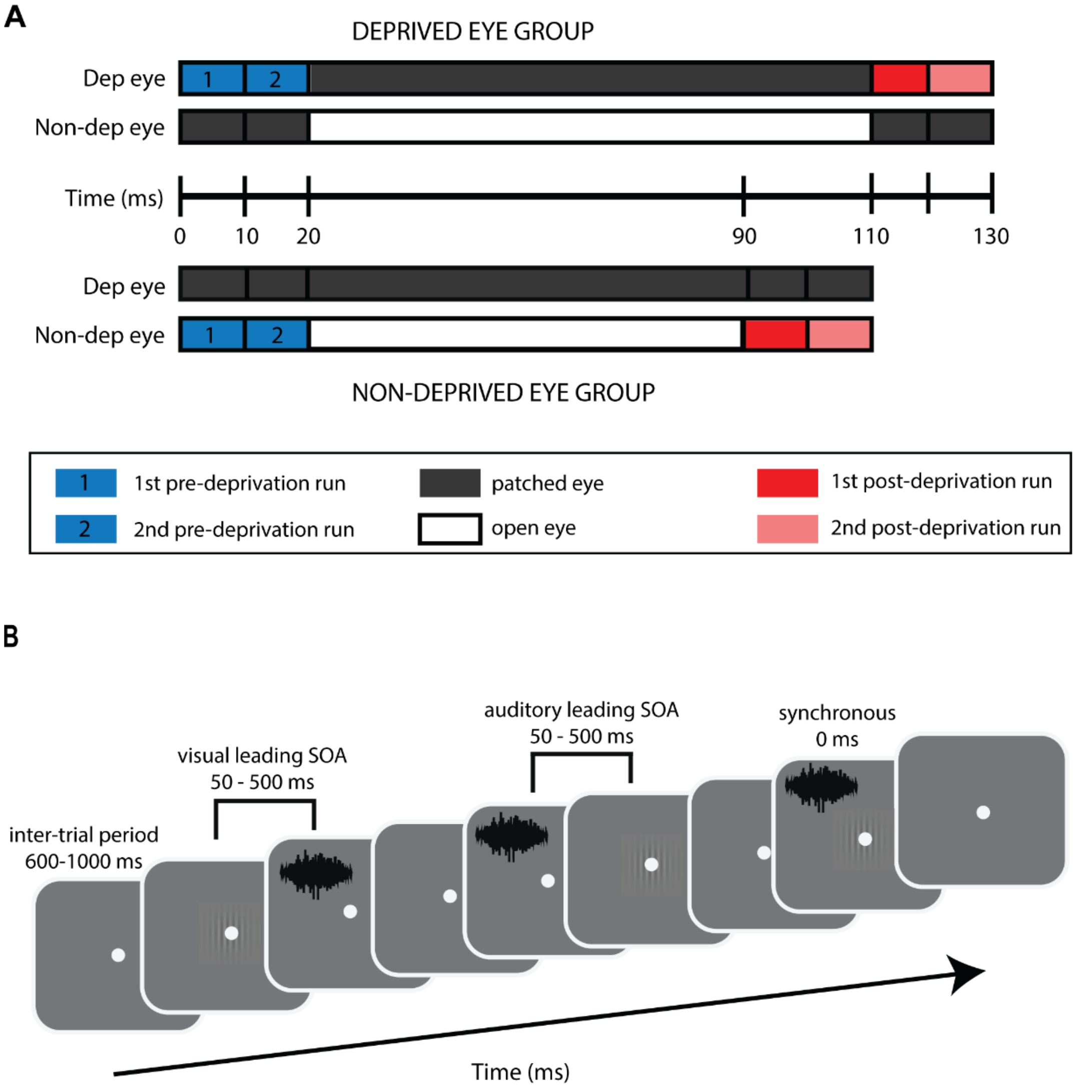

Figure 2. 1Experimental procedure.

(A) Participants were randomly assigned to either the deprived eye or non-deprived eye group. Participants in the deprived eye group performed the audiovisual temporal simultaneity judgment task with their deprived eye before and after 90 minutes of deprivation. For the non-deprived eye group, the deprivation phase started with the first pre-deprivation block of the non-deprived eye. The post-deprivation blocks were then performed using the non-deprived eye after 90 minutes from the start of the first pre-deprivation block. (B) On each block, participants judged the simultaneity of a gabor patch (33ms) and an auditory white noise burst (33ms). The onsets of the two stimuli were separated by a range of pre-selected stimulus onset asynchronies (SOAs) between -500ms and 500ms. Negative and positive SOAs signified auditory leading and visual leading respectively. On each trial, there was a brief fixation period (600-1000ms), followed by the stimulus presentation. Participants were then asked to respond by pressing the keyboard after which the next trial began automatically.

2.4 Results

The proportions of synchrony reports averaged across participants are plotted as a function of SOA for the deprived eye group (red) and the non-deprived eye group (blue) in figure 2.2. While figure 2.2A displays the results between the averaged pre-deprivation blocks (dashed line) and the first post-deprivation block (solid line), figure 2.2B shows results for the averaged pre-deprivation blocks (dashed line) and the second post-deprivation block (solid line). For each performance measure (i.e. PSS and TBW), an independent samples t-test on the combined pre-deprivation blocks showed no statistically significant difference between the two eye groups, indicating that both groups were comparable at baseline (PSS: $t(18) = 0.52$, $p = 0.611$, $BF_{10} = 0.437$; TBW: $t(18) = -1.55$, $p = 0.138$, $BF_{10} = 0.904$).

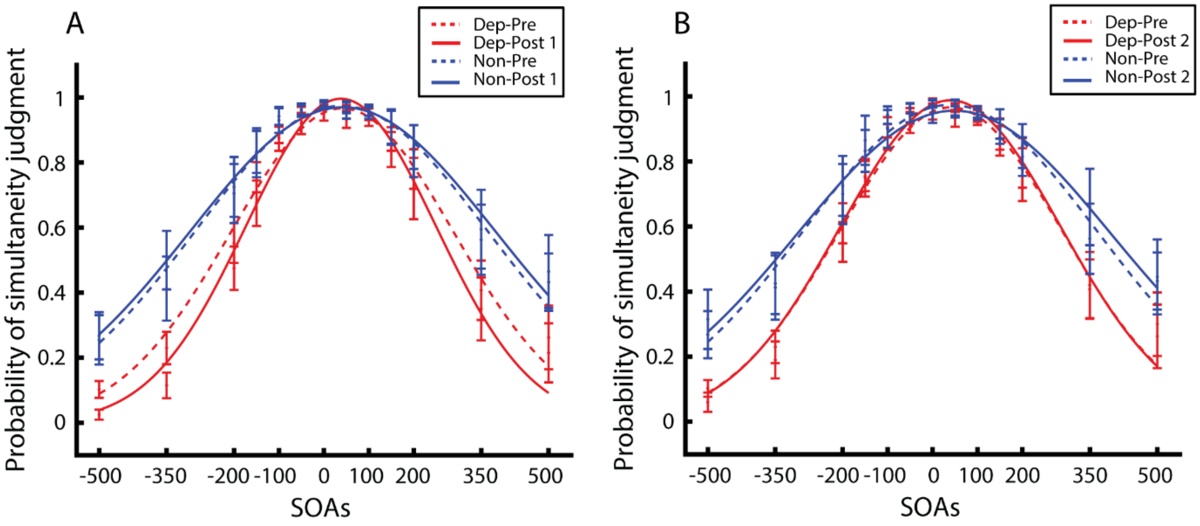

Figure 2. 2 Mean proportion of synchrony reports

Proportion of synchrony reports averaged across participants in the deprived eye group (red) and the non-deprived eye group (blue) is plotted as a function of SOA for (A) the combined average of the pre-deprivation blocks (dashed line) and the first post-deprivation block (solid line) and for (B) the combined average of the pre-deprivation blocks (dashed line) and the first post-deprivation block (solid line).

2.4.1 No effect of deprivation on PSS

The PSS is the mean of the TBW and is defined as the SOA at which maximum perceived simultaneity is perceived. Essentially, the PSS measures which modality is given more weight in determining audiovisual temporal simultaneity. Interestingly, the PSS is not usually zero and may be shifted towards the auditory-leading (i.e., more visual weight or bias) or visual-leading (i.e., more auditory weight or bias) side depending on bottom-up stimulus related factors (such as stimulus duration and intensity) and top-down attention related factors (such as being asked to attend to one modality). In order to determine the effect of monocular deprivation on the PSS, we computed the absolute change in the PSS pre- and post-deprivation. Based on this computation, a negative change signified an increase in visual bias or weighting while a positive change signified a decrease in visual bias.

The effect of 90 minutes of monocular deprivation on the PSS is displayed in Figure 2.3. Our results showed that the absolute change in PSS was not statistically significant for either the deprived eye (DG: t(9)= -0.45, Bonferroni-Holm adjusted alpha = 0.05, p= 0.663, Cohen d = -0.142, BF_{10} = 0.337) or non-deprived eye (NG: t(9)= 1.68, Bonferroni-Holm adjusted alpha = 0.025, p = 0.128, Cohen d = -0.53, BF_{10} = 0.884) groups when data were pooled across the entire post-deprivation phase. Comparing the mean absolute change in PSS between the deprived and non-deprived eye groups showed no significant difference (t(18)= -1.59, p=0.128, Cohen d = -0.713, BF_{10} = 0.943). Next, we determined whether the absolute change in PSS differed between the groups for each post deprivation phase as well as between the post deprivation phases for each group. Our results revealed that for both the first and second post deprivation phases, there was no statistically significant difference between the mean absolute change in PSS for the deprived group and that of the non-deprived group (First: t(18)= -1.28, Bonferroni-Holm adjusted alpha = 0.0167, p=0.215, Cohen d = -0.574, BF_{10} = 0.702; Second: t(18)= -1.64, Bonferroni-Holm adjusted alpha = 0.0125, p=0.118, Cohen d = -0.734, BF_{10} = 0.992). Also, the mean absolute change in PSS did not differ significantly between the first and second post-deprivation phases for the deprived group (t(9)= -0.71, Bonferroni-Holm adjusted alpha = 0.05, p=0.499, Cohen d = -0.223, BF_{10} = 0.381) and for the non-deprived group (t(9)= -0.98, Bonferroni-Holm adjusted alpha = 0.025, p=0.353, Cohen d = -0.309, BF_{10} = 0.457).

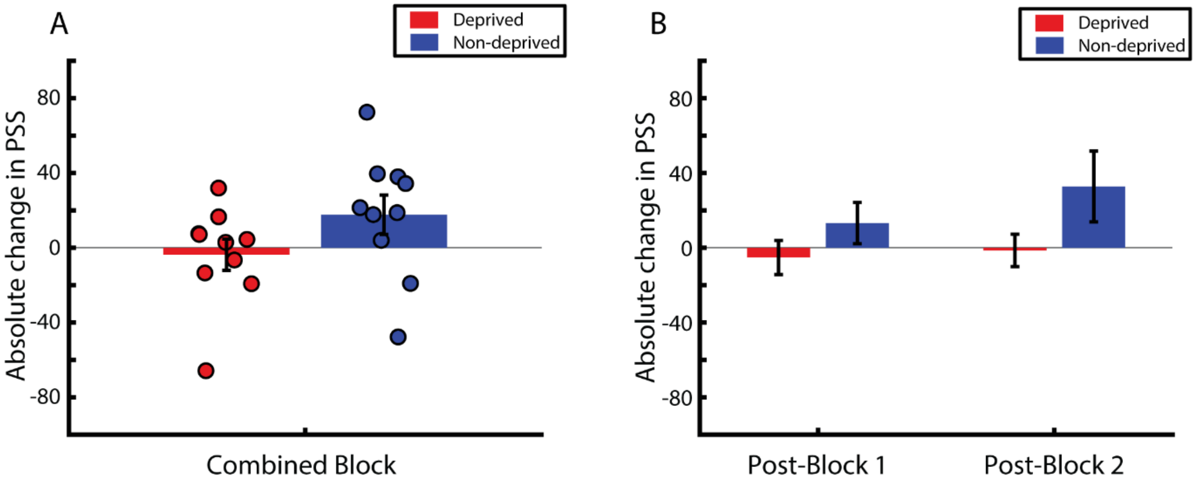

Figure 2. 3 Effects of deprivation on the point of subjective simultaneity (PSS).

(A) The mean absolute change in PSS after deprivation is plotted for the deprived (red) and non-deprived (blue) eye groups for data pooled across the entire post deprivation phase. Absolute change was computed as post – pre. (B) The mean absolute change in PSS after deprivation is plotted for the deprived (red) and non-deprived eye (blue) groups for the first and the second post-deprivation blocks. Circular dots represent individual subject data point within each eye group. The error bars represent ±SEM. ns $p>0.05$, * $p <0.05$, ** $p<0.01$, *** $p<0.001$

2.4.2 Deprivation has contrasting effects on the TBW when indexed via the deprived and non-deprived eyes

As mentioned earlier, the TBW is an index for multisensory temporal acuity with a narrow window signifying greater acuity and vice versa. In order to assess the impact of monocular deprivation on the size of the TBW, we computed the relative change in TBW between the post-deprivation and pre-deprivation phases when the data was pooled across the entire post-deprivation phase. In contrast to the lack of changes in PSS, deprivation resulted in a number of significant changes in the TBW (Figure 2.4). First, results from the combined post-deprivation block showed a significant decrease (i.e. negative relative change) in the TBW for the deprived eye group (DG: $t(9)= -4.99$, Bonferroni-Holm adjusted alpha = 0.025, $p< 0.001$, Cohen $d = -1.578$, $BF_{10} = 52.012$) and a significant increase (i.e. positive relative change) in the TBW for the non-deprived eye group (NG: $t(9)= 2.53$, Bonferroni-Holm adjusted alpha = 0.05, $p= 0.032$, Cohen $d = 0.801$, $BF_{10} = 2.477$) after deprivation (Figure 2.4A). To determine whether the two groups differed in their mean relative change in TBW, we conducted a one-way ANCOVA test controlling for the individual differences in pre-deprivation TBW. Our analysis revealed a significant difference between the mean relative change in TBW for the deprived eye group versus that for the non-deprived eye ($F(1, 17)= 16.47$, $p<0.001$, $\eta^2_p = 0.492$, $BF_{10} = 107.06$) while adjusting for pre-deprivation TBW.

Comparing the mean relative change in TBW between the deprived eye and non-deprived eye groups while adjusting for pre-deprivation TBW showed that the mean difference between the groups was significant during the first post-deprivation phase ($F(1, 17)= 24.31$, Bonferroni-Holm adjusted alpha = 0.0125, $p<0.001$, $\eta^2_p = 0.588$, $BF_{10} = 619.3$) and non-significant during the second post-deprivation phase (F(1, 17)= 2.42, Bonferroni-Holm adjusted alpha = 0.025, p= 0.138, $\eta^2_p = 0.125$, $BF_{10} = 2.167$) (Figure 2.4B). On the other hand, the mean relative change in TBW differed significantly between the first post-deprivation phase and second post-deprivation phase for the deprived eye group (DG: t(9)= -3.38, Bonferroni-Holm adjusted alpha = 0.0167, p= 0.008, Cohen d = -1.07, $BF_{10} = 7.338$). However, for the non-deprived eye group, there was no significant difference between the mean relative change in TBW for the first post-deprivation and second-post deprivation phases (NG: t(9)= -0.71, Bonferroni-Holm adjusted alpha = 0.05, p= 0.499, Cohen d = -0.223, $BF_{10} = 0.381$) (Figure 2.4B). These results emphasize two points: first, the differential impact of deprivation on the TBW of the deprived eye versus the non-deprived eye occurred only in the first deprivation phase and second, the lack of significant difference observed during the second deprivation phase is attributable to the waning of the deprivation-induced effect observed in the deprived eye.

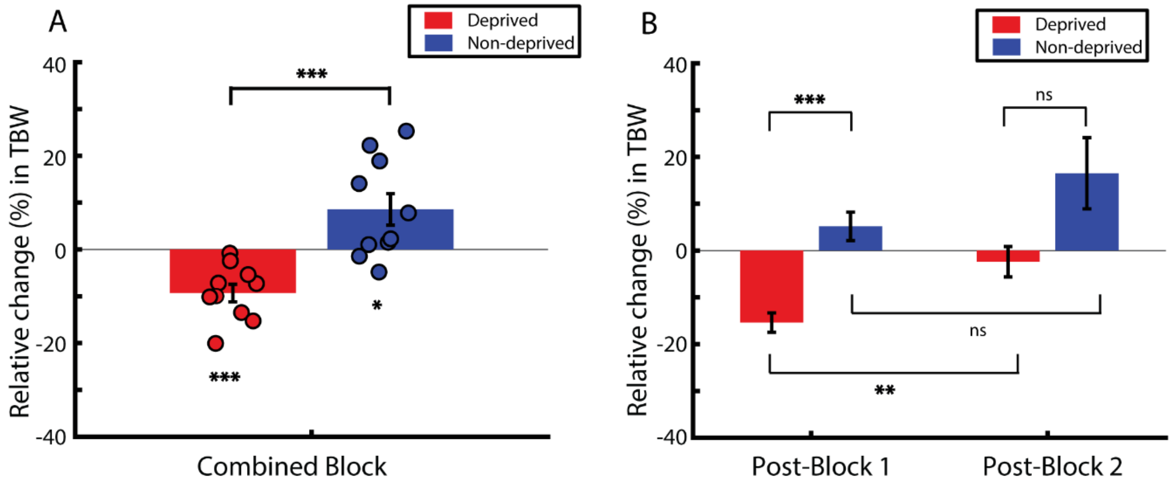

Figure 2. 4 Effects of deprivation on the temporal binding window (TBW).

(A) The mean relative change in TBW after deprivation is plotted for the deprived (red) and non-deprived (blue) eye groups for data pooled across the entire post deprivation phase. Relative change was computed as (post – pre)/pre * 100%. (B) The mean relative change in TBW after deprivation is plotted for the deprived (red) and non-deprived eye (blue) groups for the first and the second post-deprivation blocks. Circular dots represent individual subject data point within each eye group. The error bars represent ±SEM. ns $p>0.05$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

2.4.3 Relative change in TBW strongly correlates with the absolute change in PSS

Finally, we conducted a Pearson's correlation between the absolute change in PSS values pooled across all participants and post-deprivation blocks and the corresponding relative change in TBW values (Figure 2.5). Our analysis revealed a strong positive correlation between the absolute change in PSS values and the relative change in TBW values ($r = 0.7$, $p < 0.001$, $\log(BF_{10}) = 10.607$). This finding indicates that an enhancement in sensitivity to audiovisual temporal asynchrony (i.e., narrowing of the TBW), which was observed mostly for the deprived eye group, is associated with PSS shifts toward the auditory leading side (i.e., an increase in visual bias or weighting). On the contrary, a reduction in temporal acuity or a widening of the TBW, which was observed mostly in the non-deprived eye group, is associated with PSS shifts toward the visual leading side (i.e., a decrease in visual bias or weighting).

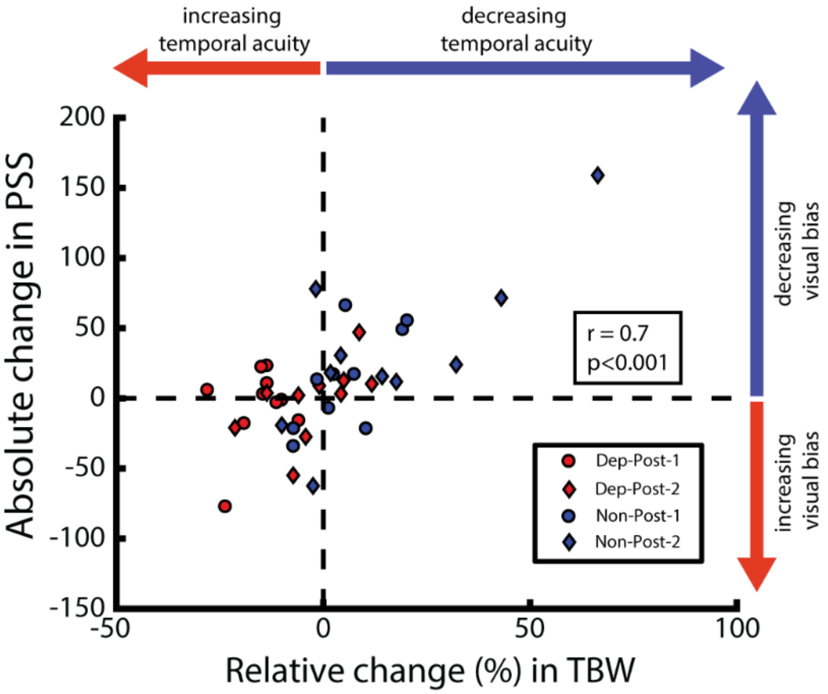

Figure 2.5 Correlation between the absolute change in point of subjective simultaneity (PSS) and the relative change in temporal window of simultaneity (TWS).

Plotted on the x-axis is relative change in TWS data pooled from participants across both eye groups (deprived eye: red; non-deprived eye group: blue) and the two post-deprivation blocks (first: circle; second: diamond) and plotted on the y-axis is the corresponding absolute change in PSS data. Positive and negative data points on the x-axis signify enhancement (narrowing of TWS) and depreciation (widening of TWS) of sensitivity to audiovisual temporal asynchrony while positive and negative values on the y-axis signify more (negative shift in PSS) and less (positive shift in PSS) dominance (dom.) of the visual modality respectively. Our results revealed a strong positive correlation between relative change in TWS and the absolute change in PSS ($r = 0.7$, $p < 0.001$).

2.5 Discussion

Even after the closure of the critical period, the adult brain retains a marked degree of plasticity that can impact perception and behavior and moreover, support recovery and rehabilitation after injury (Berardi et al., 2000; Sabel, 2008). Much of the evidence supporting this claim for the visual system has been provided through studies of perceptual learning (Goldstone, 1998; Huxlin, 2008) and recently, through brief periods of monocular deprivation (Lunghi et al., 2011; Zhou et al., 2013). Adding to this body of knowledge, our study provides the first clear evidence that short-term monocular deprivation can induce changes in multisensory (i.e., audiovisual) temporal perception. Using an audiovisual simultaneity judgment task, we discovered that 90 minutes of monocular deprivation produced contrasting effects on the TBW (a measure of audiovisual temporal acuity), narrowing the TBW of participants who performed the task with their deprived eye and widening the TBW of those who used their non-deprived eye. Moreover, we observed that this effect was short-lived, occurring only in the first phase of post-deprivation testing and which took place within 10 minutes of the removal of deprivation.

During visual deprivation, the loss of visual information also impacts any visual-based multisensory processing (Carriere et al., 2007). In fact, the absence (often through the removal of visual information) or alteration (by manipulating the spatial and temporal properties of the cross-modal cues) of multisensory experience can influence normal multisensory function both before and after the critical period (Carriere et al., 2007; Odegaard, Wozny, and Shams, 2017; Stein, Stanford, and Rowland, 2014; Wallace et al., 2004). While the deprivation effect observed in our study could be a result of compensation to the reduction in visually based multisensory information, we believe this is highly unlikely for two reasons. First, the monocular deprivation paradigm likely has very little impact on visually based multisensory experience, given the remaining information coming from the non-deprived eye. Second, the contrasting nature of the effects observed in the deprived and non-deprived eyes in our study are consistent with results that have been observed in purely visual studies of monocular deprivation (Lunghi et al., 2011; Lunghi et al., 2013; Zhou et al., 2013) and hence, points to the effects being driven in large measure by visual, as opposed to multisensory, based compensatory mechanisms.

Monocular deprivation represents a powerful tool to study plasticity in the adult visual system. Indeed, evidence from past studies points to more detrimental effects of monocular as opposed to binocular deprivation on visual cortical organization and behavioral performance (Lewis, Maurer, and Brent, 1995; Wiesel and Hubel, 1965a). Mechanistically, monocular deprivation triggers various forms of homeostatic plasticity in an effort to restore the loss of visual drive and maintain the average levels of neural activity (Maffei, Nelson, and Turrigiano, 2004; Maffei and Turrigiano, 2008; Wang, Fontanini, and Maffei, 2012). These forms of homeostatic plasticity induced by monocular deprivation involve alterations in the balance of cortical excitation and inhibition (E/I) via a complex interplay of excitation, inhibition, and intrinsic excitability (Maffei et al., 2004; Maffei and Turrigiano, 2008; Wang et al., 2012). In adult humans, short-term monocular deprivation has been associated with the downregulation of GABA concentration in the primary visual cortex (Lunghi, Emir, Morrone, and Bridge, 2015). Using EEG recordings, other studies have also demonstrated that short-term monocular deprivation increases the amplitude of visual evoked potentials and neural oscillations believed to originate from the primary visual cortex when the deprived eye is stimulated and decreases these measures when the non-deprived eye is stimulated (Lunghi et al., 2015; Zhou, Baker, Simard, Saint-Amour, and Hess, 2015). While these results emphasize an increase in overall neural activity after monocular deprivation, a recent 7T fMRI study has shown that besides an increase in the BOLD signal in V1 for the deprived eye, monocular deprivation also shifts ocular dominance distributions in favor of the deprived eye (Binda et al., 2018). Collectively, these findings indicate that monocular deprivation induces plastic mechanisms geared at enhancing the visual information coming from the deprived eye.

Extending these findings to our current results, we believe that the changes observed in the TBW could arise from two main causes which can occur separately or in tandem. Using the causal inference model, these two factors can be expressed in terms of changes in visual unisensory encoding and changes in the tendency to bind the audiovisual information (Magnotti, Ma, and Beauchamp, 2013; Körding et al., 2007; Beierholm, Quartz, and Shams, 2009). First, changes in the TBW as a result of changes in visual encoding after monocular deprivation is consistent with previous visual studies showing that, monocular deprivation enhances the perception of the deprived eye while depressing that of the non-deprived eye (Lunghi et al., 2011; Zhou et al., 2013). In the multisensory domain, several studies have provided evidence that support the hypothesis that manipulating stimulus reliability, which occurs in the context of monocular deprivation, can influence multisensory interactions. For instance, Fister, Stevenson, Nidiffer, Barnett, and Wallace (2016) demonstrated that the proportion of perceived simultaneity of the paired audiovisual stimuli across SOA decreased faster for high than low saliency conditions. Although the above study did not compute TBWs, their finding indicates that increasing stimulus effectiveness reduces the tolerance to audiovisual temporal asynchrony, which is comparable to narrowing of the TBW. Furthermore, Magnotti et al. (2013) demonstrated that decreasing the reliability (via blurring) of the visual stimulus during the performance of an SJ task widened the TBW. Studies using causal inference models have demonstrated that such effects emerge from changes in sensory noise influencing the process of visual encoding (Magnotti et al., 2013; Beierholm et al., 2009).

Second, it is plausible that the changes observed in the TBW after monocular deprivation could arise from genuine cross-modal plasticity reflecting changes in the tendency to bind the audiovisual stimulus pair. In a very interesting study, Lo Verde, Morrone, and Lunghi (2017) discovered using a binocular rivalry paradigm that after monocular deprivation, the effect of visuo-haptic interaction on perceptual dominance disappeared for the deprived eye, which was potentiated in the visual domain, but was not affected for the non-deprived eye, which was weakened. In a separate control experiment where post-deprivation effect was simulated by increasing the intensity of the visual stimulus for one eye and decreasing the intensity for the other eye, this unequal effect of monocular deprivation on cross-modal interaction for the deprived and non-deprived eyes disappeared, indicating that the effects on monocular deprivation were due to mechanisms that exceeded changes in contrast-gain of the deprived eye (Lo Verde et al., 2017). The idea that changes in the TBW can result from changes in binding tendency is consistent with studies that have demonstrated a relationship between TBW and intracortical E/I balance – the underlying mechanism for homeostatic plasticity induced by monocular deprivation. For example, a recent study demonstrated that the size of the audio-tactile TBW measured across participants was associated with cortical E/I balance indexed as glutamate/ GABA concentrations in the left primary auditory cortex (Ferri et al., 2017), a region noted for integration of auditory and tactile information (Schürmann, Caetano, Hlushchuk, Jousmäki, and Hari, 2006; Kayser, Petkov, Augath, and Logothetis, 2005). Specifically, they showed that moderate E/I balance was associated with the narrowest TBW sizes, while extreme (i.e., relatively high or low) E/I balance was associated with wider TBW sizes (Ferri et al., 2017). Additionally, the link between changes in cortical E/I balance mediating deprivation-induced homeostatic plasticity and the TBW can be observed in clinical conditions or in perceptual mechanisms where E/I balance is affected. For example, individuals with obesity, known to exhibit impaired homeostatic plasticity (suggestive of the presence of an altered E/I balance) due to a stronger GABAergic inhibition (Lunghi et al., 2019) also show widened TBW compared to healthy-weight controls (Scarpina et al., 2016). In addition, the TBW has been shown to be highly malleable to perceptual learning (Powers et al., 2012; Powers et al., 2009), which is known mechanistically to involve Hebbian plasticity, a form of experience dependent plasticity (similar to homeostatic plasticity induced by monocular deprivation) that depend on the net intracortical E/I balance (see Binda et al., 2018 for discussion). Together, these studies provide evidence that suggests that the TBW can be modulated to some extent by a direct influence of changes in cortical E/I levels and or through changes in visual representations occurring as a result of changes in E/I balance after monocular deprivation. In order to determine the relative contributions of these two mechanisms to the changes in TBW, further studies using neuroimaging techniques will be necessary to unravel the brain areas involved, the time course of these effects and the changes in connectivity patterns induced by these effects.

Apart from changes in the TBW, we found that depriving participants monocularly for 90 minutes did not reveal any significant changes in the PSS within and between the eye conditions. The PSS is a central tendency of the TBW and represents a measure of perceptual weighting between the two modalities whose simultaneity is being judged. Mechanistically, it is known to stem from a combination of 1) the difference between the arrival time latencies and 2) the difference between the processing time of the two modalities (García-Pérez and Alcalá-Quintana, 2012). The lack of deprivation effect on the PSS seen in our study can be explained by prior EEG work that showed that monocular deprivation did not affect the latencies of C1, P1 and P2 components of the visual evoked potentials (Lunghi, Berchicci, Morrone, and Di Russo, 2015). While the change in PSS was non-significant, we did observe a strong positive correlation between the absolute change in PSS and the relative change in the TBW when the data were pooled across the two post-deprivation blocks and participants in both eye groups. This finding implies that, the two measures i.e., PSS and TBW were driven by similar mechanisms underlying monocular deprivation, albeit less strongly for the PSS than for the TBW.

In conclusion, we have shown that a short perturbation in normal visual experience in adults through monocular deprivation results in significant, albeit transient changes in multisensory temporal perception. These findings have implications on how the adult brain may react to visual impairment and how the clinical use of monocular deprivation as therapy may impact the restoration of balanced binocular function in adult patients with amblyopia. First, changes in visual experience in the adult can influence the way the visual system interacts with other sensory modalities, and thus such changes in multisensory perception may contribute to the overall compensatory behaviors observed in patients with adult-onset visual deprivation. Second, several studies have reported intact homeostatic plasticity among clinical populations such as amblyopia (Zhou, Thompson, and Hess, 2013; Lunghi, Morrone, Secci, and Caputo, 2016) and consequently, have employed monocular deprivation as therapy for improving visual perception in the amblyopic eye (Zhou et al., 2019; Lunghi et al., 2019). Our findings imply that besides improving visual function in the amblyopic eye, monocular deprivation therapy may also ameliorate the impaired multisensory temporal processing observed in adult patients with amblyopia (Richards, Goltz, and Wong, 2017).

2.6 References

Alais, D., and Blake, R. (Eds.). (2005). *Binocular rivalry*. MIT press.

Beierholm, U. R., Quartz, S. R., and Shams, L. (2009). Bayesian priors are encoded independently from likelihoods in human multisensory perception. *Journal of Vision*, *9*(5), 23-23.

Berardi, N., Pizzorusso, T., and Maffei, L. (2000). Critical periods during sensory development. *Current Opinion in Neurobiology, 10*(1), 138-145.

Binda, P., Kurzawski, J. W., Lunghi, C., Biagi, L., Tosetti, M., and Morrone, M. C. (2018). Response to short-term deprivation of the human adult visual cortex measured with 7T BOLD. *Elife*, *7*, e40014.

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

Carriere, B. N., Royal, D. W., Perrault, T. J., Morrison, S. P., Vaughan, J. W., Stein, B. E., and Wallace, M. T. (2007). Visual deprivation alters the development of cortical multisensory integration. *Journal of Neurophysiology*, *98*(5), 2858-2867.

Chadnova, E., Reynaud, A., Clavagnier, S., and Hess, R. F. (2017). Short-term monocular occlusion produces changes in ocular dominance by a reciprocal modulation of interocular inhibition. *Scientific Reports, 7*, 41747.

Chen, Y.-C., Lewis, T. L., Shore, D. I., and Maurer, D. (2017). Early binocular input is critical for development of audiovisual but not visuotactile simultaneity perception. *Current Biology, 27*(4), 583-589.

Collignon, O., Dormal, G., De Heering, A., Lepore, F., Lewis, T. L., and Maurer, D. (2015). Long-lasting cross-modal cortical reorganization triggered by brief postnatal visual deprivation. *Current Biology, 25*(18), 2379-2383.

Diederich, A., and Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Perception and Psychophysics, 66*(8), 1388-1404.

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

Ferri, F., Nikolova, Y. S., Perrucci, M. G., Costantini, M., Ferretti, A., Gatta, V., ... and Sibille, E. (2017). A neural “tuning curve” for multisensory experience and cognitive-perceptual schizotypy. *Schizophrenia Bulletin, 43*(4), 801-813.

Fine, I., Smallman, H. S., Doyle, P., and MacLeod, D. I. (2002). Visual function before and after the removal of bilateral congenital cataracts in adulthood. *Vision Research, 42*(2), 191-210.

Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., . . . MacLeod, D. I. (2003). Long-term deprivation affects visual perception and cortex. *Nature Neuroscience*, *6*(9), 915.

Fister, J. K., Stevenson, R. A., Nidiffer, A. R., Barnett, Z. P., and Wallace, M. T. (2016). Stimulus intensity modulates multisensory temporal processing. *Neuropsychologia, 88*, 92-100.

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., and Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research*, *203*(2), 381-389.

Foxe, J. J., and Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *Neuroreport, 16*(5), 419-423.

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

García-Pérez, M. A., and Alcalá-Quintana, R. (2012). On the discrepant results in synchrony judgment and temporal-order judgment tasks: a quantitative model. *Psychonomic Bulletin and Review, 19*(5), 820-846.

Ghazanfar, A. A., and Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences, 10*(6), 278-285.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*(1), 585-612.

Guerreiro, M. J., Putzar, L., and Röder, B. (2015). The effect of early visual deprivation on the neural bases of multisensory processing. *Brain, 138*(6), 1499-1504.

Hillock-Dunn, A., and Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental Science, 15*(5), 688-696.

Hillock, A. R., Powers, A. R., and Wallace, M. T. (2011). Binding of sights and sounds: age-related changes in multisensory temporal processing. *Neuropsychologia, 49*(3), 461-467.

Hubel, D. H., and Wiesel, T. N. (1970). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *The Journal of Physiology, 206*(2), 419-436.

Huxlin, K. R. (2008). Perceptual plasticity in damaged adult visual systems. *Vision Research, 48*(20), 2154-2166.

JASP Team. (2018). JASP [Computer software, Version 0.11. 0.1].

Kayser, C., Petkov, C. I., Augath, M., and Logothetis, N. K. (2005). Integration of touch and sound in auditory cortex. *Neuron, 48*(2), 373-384.

Kim, H.-W., Kim, C.-Y., and Blake, R. (2017). Monocular perceptual deprivation from interocular suppression temporarily imbalances ocular dominance. *Current Biology, 27*(6), 884-889.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., and Shams, L. (2007). Causal inference in multisensory perception. *PLoS One, 2*(9), 943.

Levi, D. M., McKee, S. P., and Movshon, J. A. (2011). Visual deficits in anisometropia. *Vision Research, 51*(1), 48-57.

Lewis, T. L., Maurer, D., and Brent, H. P. (1995). Development of grating acuity in children treated for unilateral or bilateral congenital cataract. *Investigative Ophthalmology and Visual Science, 36*(10), 2080-2095.

Lovelace, C. T., Stein, B. E., and Wallace, M. T. (2003). An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Cognitive Brain Research, 17*(2), 447-453.

Lo Verde, L., Morrone, M. C., and Lunghi, C. (2017). Early cross-modal plasticity in adults. *Journal of Cognitive Neuroscience*, *29*(3), 520-529.

Lunghi, C., Berchicci, M., Morrone, M. C., and Di Russo, F. (2015). Short-term monocular deprivation alters early components of visual evoked potentials. *The Journal of Physiology, 593*(19), 4361-4372.

Lunghi, C., Burr, D. C., and Morrone, C. (2011). Brief periods of monocular deprivation disrupt ocular balance in human adult visual cortex. *Current Biology, 21*(14), R538-R539.

Lunghi, C., Burr, D. C., and Morrone, M. C. (2013). Long-term effects of monocular deprivation revealed with binocular rivalry gratings modulated in luminance and in color. *Journal of Vision, 13*(6), 1-1.

Lunghi, C., Daniele, G., Binda, P., Dardano, A., Ceccarini, G., Santini, F., ... and Morrone, M. C. (2019). Altered Visual Plasticity in Morbidly Obese Subjects. *iScience*, *22*, 206-213.

Lunghi, C., Emir, U. E., Morrone, M. C., and Bridge, H. (2015). Short-term monocular deprivation alters GABA in the adult human visual cortex. *Current Biology*, *25*(11), 1496-1501.

Lunghi, C., Morrone, M. C., Secci, J., and Caputo, R. (2016). Binocular rivalry measured 2 hours after occlusion therapy predicts the recovery rate of the amblyopic eye in anisometropic children. *Investigative Ophthalmology and Visual Science*, *57*(4), 1537-1546.

Lunghi, C., Sframeli, A. T., Lepri, A., Lepri, M., Lisi, D., Sale, A., and Morrone, M. C. (2019). A new counterintuitive training for adult amblyopia. *Annals of Clinical and Translational Neurology*, *6*(2), 274-284.

Maffei, A., Nelson, S. B., and Turrigiano, G. G. (2004). Selective reconfiguration of layer 4 visual cortical circuitry by visual deprivation. *Nature Neuroscience*, *7*(12), 1353-1359.

Maffei, A., and Turrigiano, G. G. (2008). Multiple modes of network homeostasis in visual cortical layer 2/3. *Journal of Neuroscience*, *28*(17), 4377-4384.

Magnotti, J. F., Ma, W. J., and Beauchamp, M. S. (2013). Causal inference of asynchronous audiovisual speech. *Frontiers in Psychology*, *4*, 798.

Meredith, M. A., Nemitz, J. W., and Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience*, *7*(10), 3215-3229.

Noel, J.-P., De Niear, M., Van der Burg, E., and Wallace, M. T. (2016). Audiovisual simultaneity judgment and rapid recalibration throughout the lifespan. *PLoS One, 11*(8), e0161698.

Noel, J. P., Stevenson, R. A., and Wallace, M. T. (2018). Atypical audiovisual temporal function in autism and schizophrenia: similar phenotype, different cause. *European Journal of Neuroscience*, *47*(10), 1230-1241.

Odegaard, B., Wozny, D. R., and Shams, L. (2017). A simple and efficient method to enhance audiovisual binding tendencies. *PeerJ, 5*, e3143.

Ostrovsky, Y., Andalman, A., and Sinha, P. (2006). Vision following extended congenital blindness. *Psychological Science, 17*(12), 1009-1014.

Pascual-Leone, A., Amedi, A., Fregni, F., and Merabet, L. B. (2005). The plastic human brain cortex. *Annu. Rev. Neurosci.*, *28*, 377-401.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437-442.

Powers, A. R., Hevey, M. A., and Wallace, M. T. (2012). Neural correlates of multisensory perceptual learning. *Journal of Neuroscience, 32*(18), 6263-6274.

Powers, A. R., Hillock, A. R., and Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *Journal of Neuroscience*, *29*(39), 12265-12274.

Putzar, L., Goerendt, I., Lange, K., Rösler, F., and Röder, B. (2007). Early visual deprivation impairs multisensory interactions in humans. *Nature Neuroscience, 10*(10), 1243.

Putzar, L., Hötting, K., and Röder, B. (2010). Early visual deprivation affects the development of face recognition and of audio-visual speech perception. *Restorative Neurology and Neuroscience, 28*(2), 251-257.

Richards, M. D., Goltz, H. C., and Wong, A. M. (2017). Alterations in audiovisual simultaneity perception in amblyopia. *PloS One, 12*(6), e0179516.

Sabel, B. A. (2008). Plasticity and restoration of vision after visual system damage: an update. *Restorative Neurology and Neuroscience, 26*(4, 5), 243-247.

Scarpina, F., Migliorati, D., Marzullo, P., Mauro, A., Scacchi, M., and Costantini, M. (2016). Altered multisensory temporal integration in obesity. *Scientific Reports, 6*, 28382.

Schürmann, M., Caetano, G., Hlushchuk, Y., Jousmäki, V., and Hari, R. (2006). Touch activates human auditory cortex. *Neuroimage*, *30*(4), 1325-1331.

Simon, D. M., Noel, J. P., and Wallace, M. T. (2017). Event related potentials index rapid recalibration to audiovisual temporal asynchrony. *Frontiers in Integrative Neuroscience*, *11*, 8.

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

Stein, B. E., and Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience, 9*(4), 255.

Stein, B. E., Stanford, T. R., and Rowland, B. A. (2014). Development of multisensory integration from the perspective of the individual neuron. *Nature Reviews Neuroscience*, *15*(8), 520-535.

Stevenson, R. A., Park, S., Cochran, C., McIntosh, L. G., Noel, J. P., Barense, M. D., ... and Wallace, M. T. (2017). The associations between multisensory temporal processing and symptoms of schizophrenia. *Schizophrenia Research*, *179*, 97-103.

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., and Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *Journal of Neuroscience*, *34*(3), 691-697.

Stevenson, R. A., Zemtsov, R. K., and Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental Psychology: Human Perception and Performance, 38*(6), 1517.

Tong, F., Meng, M., and Blake, R. (2006). Neural bases of binocular rivalry. *Trends in Cognitive Sciences*, *10*(11), 502-511.

Wallace, M. T., Perrault, T. J., Hairston, W. D., and Stein, B. E. (2004). Visual experience is necessary for the development of multisensory integration. *Journal of Neuroscience*, *24*(43), 9580-9584.

Wallace, M. T., and Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia*, *64*, 105-123.

Wang, L., Fontanini, A., and Maffei, A. (2012). Experience-dependent switch in sign and mechanisms for plasticity in layer 4 of primary visual cortex. *Journal of Neuroscience*, *32*(31), 10562-10573.

Wiesel, T. N., and Hubel, D. H. (1963). Effects of visual deprivation on morphology and physiology of cells in the cat's lateral geniculate body. *Journal of Neurophysiology, 26*(6), 978-993.

Wiesel, T. N., and Hubel, D. H. (1965a). Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. *Journal of Neurophysiology, 28*(6), 1029-1040.

Wiesel, T. N., and Hubel, D. H. (1965b). Extent of recovery from the effects of visual deprivation in kittens. *Journal of Neurophysiology, 28*(6), 1060-1072.

Zhou, J., Baker, D. H., Simard, M., Saint-Amour, D., and Hess, R. F. (2015). Short-term monocular patching boosts the patched eye’s response in visual cortex. *Restorative Neurology and Neuroscience, 33*(3), 381-387.

Zhou, J., Clavagnier, S., and Hess, R. F. (2013). Short-term monocular deprivation strengthens the patched eye's contribution to binocular combination. *Journal of Vision, 13*(5), 12-12.

Zhou, J., He, Z., Wu, Y., Chen, Y., Chen, X., Liang, Y., ... and Hess, R. F. (2019). Inverse occlusion: a binocularly motivated treatment for amblyopia. *Neural Plasticity, 2019.*

Zhou, J., Thompson, B., and Hess, R. F. (2013). A new form of rapid binocular plasticity in adult with amblyopia. *Scientific Reports, 3*, 2638.

Zou, H., Müller, H. J., and Shi, Z. (2012). Non-spatial sounds regulate eye movements and enhance visual search. *Journal of Vision, 12*(5), 2-2.

Chapter 3

Neural Correlates of Short-Term Monocular Deprivation Effects on Multisensory Temporal Perception

3.1 Abstract

In adult humans, short-term monocular deprivation can induce changes in multisensory temporal perception specifically by narrowing the temporal binding window (TBW) in the deprived eye and widening the window in the non-deprived eye. However, the neural mechanisms underlying this intriguing plasticity phenomenon have not yet been explored. Here, using EEG recordings in human participants, we investigated the role of neural oscillatory activity in the short-term monocular deprivation effects on multisensory temporal perception. Behaviorally, we discovered that 120 minutes of monocular deprivation had significantly different impact on the TBW for the deprived eye and the non-deprived eye. However, unlike our previous findings, monocular deprivation significantly decreased the TBW for the deprived eye and while there was a trend of increased TBW for the non-deprived eye, this trend was not statistically significant. Furthermore, monocular deprivation significantly reduced the maximum probability of simultaneity only for the deprived eye group indicating a reduced tendency to bind audiovisual stimuli. Our neural analysis revealed a couple of interesting findings. First, we found that monocular deprivation increased the oscillatory power and inter-trial phase coherence in the alpha and low beta band for the deprived eye compared with the non-

deprived eye during stimulus conditions where the visual stimulus preceded. Surprisingly, our result showed that auditory alpha oscillatory activity in the occipital-parietal region increased in power for the deprived eye after monocular deprivation implying cross-modal mechanisms. Lastly, uniquely for the audiovisual conditions, we found increased high beta power for the deprived eye compared with the non-deprived eye after monocular deprivation. Taken together, our findings indicate that both alpha and beta oscillations are involved in the short-term monocular deprivation effects on multisensory temporal perception.

3.2 Introduction

Most of the events we encounter in the world are multisensory in nature, thus the brain is constantly bombarded with information that simultaneously stimulate our different sense organs (Driver and Noesselt, 2008; Barry E Stein and Meredith, 1993; Barry E. Stein and Stanford, 2008). Integrating such information from multiple senses has striking neural (Barry E Stein and Meredith, 1993; Barry E. Stein and Stanford, 2008) and perceptual benefits (Diederich and Colonius, 2015; Ernst and Banks, 2002; Frassinetti, Bolognini, and Làdavas, 2002; Hairston, Laurienti, Mishra, Burdette, and Wallace, 2003; Lovelace, Stein, and Wallace, 2003; Zou, Müller, and Shi, 2012). A vital component of multisensory integration involves determining which signals originate from a single source and hence, should be integrated in order to construct a unified representation of the world (Körding et al., 2007; Shams and Beierholm, 2010; Barry E Stein and Meredith, 1993). The brain achieves this by relying on the amodal and physical characteristics of the stimuli that are to be combined such as their spatial and temporal relationship (Barry E Stein and Meredith, 1993). Intuitively, the closer two stimuli are in space and in time, the more likely they belong to the same event. In recent years, much attention has been given to how the brain uses the temporal relationship between sensory cues to make inferences about their sources. Out of this work has emerged an important construct known as the temporal binding window, which refers to a window of time within which the multisensory stimuli are highly likely to be perceptually bound or integrated, thus, serving as a proxy measure for multisensory temporal acuity (Wallace and Stevenson, 2014). The size of the TBW is influenced by factors such stimulus complexity (Stevenson and Wallace, 2013) and stimulus effectiveness (Nidiffer, Stevenson, Krueger Fister, Barnett, and Wallace, 2016) and is malleable to the effects of perceptual learning (Powers, Hevey, and Wallace, 2012; Powers, Hillock, and Wallace, 2009).

Recently, several studies have been interested in understanding the role of neural oscillatory activity in multisensory temporal processing especially in the alpha band. For instance, while lower pre-stimulus parietal alpha power is associated with temporal sensitivity indicating encoding of sensory information (London et al., 2020), pre-stimulus frontal alpha power is found to correlate with the tendency to respond relative to one's idiosyncratic bias (Grabot, Kösem, Azizi, and Van Wassenhove, 2017) reflecting mechanisms influencing perceptual decisional processes. Furthermore, alpha-frequency is associated with the accuracy of making simultaneity (Bastiaansen, Berberyan, Stekelenburg, Schoffelen, and Vroomen, 2020) and temporal order judgments (London et al., 2020). Besides, the alpha band, other studies have demonstrated the role of pre-stimulus beta oscillatory power in multisensory temporal perception. For instance, decreased high beta power (20-30Hz) is predictive of correct response on an audiovisual temporal order judgment task (Grabot et al., 2017). Furthermore, increased beta-band inter-trial phase coherence (ITC) in central and occipital electrodes is associated with the perception of synchronous but not asynchronous audiovisual stimuli (Kambe, Kakimoto, and Araki, 2015). Taken together, these findings show that the power and phase of oscillatory activity in alpha and beta frequency bands modulate multisensory temporal congruence, thereby establishing the important role of neural oscillations in shaping the multisensory temporal processes.

During development, the TBW has been shown to be broader in children and in adolescence before it begins to mature (i.e., narrow) in adults (Hillock-Dunn and Wallace, 2012; Hillock, Powers, and Wallace, 2011). Because of its remarkable plasticity, the development and maturation of the TBW can be influenced by conditions that impair normal sensory development even in a single sensory modality. For instance, patients who experienced visual deprivation owing to congenital cataracts during the first 0.3 to 29 months of their life showed wider TBWs when compared to age-matched typically developed controls during performance of an audiovisual simultaneity judgment task (Chen, Lewis, Shore, and Maurer, 2017), indicating an immature audiovisual integration system. Additionally, similar evidence of widened TBW has been found in patients with early visual deprivation caused by anisometropia (unequal refractive errors in both eyes) or strabismus (patent lateral deviation of one eye) or both in early life (Richards, Goltz, and Wong, 2017).

Surprisingly, the effects of visual deprivation on the multisensory temporal perception can occur and in a very rapid fashion in the adult brain where plasticity is known to reduce after development. In Chapter 2, we showed that 90 minutes of monocular deprivation produced opposing effects on the temporal binding window depending on the eye used in task; narrowing the temporal binding window for the deprived eye whereas widening it for the non-deprived eye. Prior to this finding, a similar contrasting effect after monocular deprivation was demonstrated using pure visual tasks such as binocular rivalry (Lunghi, Burr, and Morrone, 2011; Lunghi, Burr, and Morrone, 2013) and contrast detection etc. (Zhou, Clavagnier, and Hess, 2013). In the visual domain, the perceptual effect observed after short-term monocular deprivation is thought to reflect a form of rapid homeostatic plasticity to visual deprivation, which may be mediated by an upregulation of contrast gain-control mechanisms of the deprived eye to optimize the attenuated or absent visual information (Lunghi et al., 2011). Follow-up studies have demonstrated the role of alpha oscillations in the primary visual cortex in mediating this short-term plasticity (Lunghi, Berchicci, Morrone, and Di Russo, 2015). Specifically, short-term monocular deprivation increased the power of post-stimulus alpha oscillations for the deprived eye and decreased the power for the non-deprived eye (Lunghi et al., 2015).

In order to understand the neural mechanisms underlying the effect of short-term monocular deprivation on multisensory temporal perception, we recorded EEG signals from participants performing an audiovisual simultaneity judgement task using their deprived eye or non-deprived eye before and after 120 minutes of deprivation. Given that, oscillatory activity in the alpha band has been implicated in both short-term monocular deprivation and multisensory temporal perception, we focused our investigation primarily on the role of this oscillatory band in the deprivation-induced effects on multisensory temporal perfection. In addition, we explored oscillatory activity in other frequency bands in the low beta and high beta regions to determine their involvement in this observed perceptual effect. In short, our findings showed that both alpha and beta oscillations are involved in the short-term monocular deprivation effects on multisensory temporal perception.

3.3 Method

We recruited sixteen participants to perform an audiovisual simultaneity judgment task. Before the start of the experiment, participants were randomly assigned to either the deprived eye group (N = 8, 3 males, mean age (in years) = 19.1, SD = 3.0) or the non-deprived eye group (N = 8, 4 males, mean age = 18.9 (in years), SD = 1.2). All participants reported normal or corrected-to-normal vision and normal hearing and normal vision was confirmed using a Snellen visual acuity task. This study was approved by the Vanderbilt University Institutional Review Board (IRB) and each participant gave informed consent before being allowed to participate. Compensations for participation were in the form of gift cards or course credits. This work was carried out in accordance with the Declaration of Helsinki.

3.3.1 Stimulus and Apparatus

All stimuli used in the study were generated and presented using MATLAB (Math Works Inc., Natick, MA) software with the Psychophysics Toolbox Version 3 (Brainard and Vision, 1997; Pelli, 1997). The visual stimulus was a Gabor patch (sigma = 0.5 degrees) displayed on a gamma-corrected monitor (21-inch Asus LCD) with 120-Hz refresh rate. The Gabor patch had a spatial frequency of 4 cycles per degree and a size of 6 degrees and was displayed at 20% contrast at the center of fixation on a uniform gray background with a luminance of 10cd/m^2. The auditory stimulus on the other hand was a white noise burst with a sampling frequency of 48 kHz displayed at ~75dB through external speakers placed on the sides of the monitor. Both stimuli were presented for four monitor frame periods equaling 33ms. To avoid stimulus transients, the intensity of the auditory stimulus was linearly ramped up and down, each for 25% of the stimulus duration. For the visual stimulus, a similar effect was achieved by presenting the first and last frames at half the actual contrast of the stimulus. A Minolta Chroma Meter CS-100 and a sound level meter were used to verify the luminance and sound intensity levels respectively. The durations of all visual and auditory stimuli, as well as the SOAs, were confirmed using a Hameg 507 oscilloscope with a photovoltaic cell and microphone.

3.3.2 Procedure

Participants performed an audiovisual simultaneity judgment task where they were required to judge whether a Gabor patch and an auditory white noise burst occurred simultaneously or not. These experiments were conducted inside a WhisperRoomTM (SE 2000 Series) while participants rested their forehead and chin comfortably against a HeadSpotTM (University of Houston Optometry). The whisper room was dimly lit by the background luminance of the monitor.

Before the experiment began, each participant was randomly assigned to either the deprived eye group (i.e., performed the test with their deprived eye) or the non-deprived eye group (i.e., performed the test with their non-deprived eye). The eye to-be deprived was randomly chosen for each participant and counterbalanced across all participants. Deprivation was performed monocularly for 120 minutes with an opaque eye patch (Figure 3.1A). Both the pre and the post deprivation sessions comprised five blocks of an unspeeded audiovisual simultaneity judgment task including EEG recordings.

Each trial begun with a brief fixation period, which was marked by the presence of white circular dot and lasting between 1200 and 1500ms. After the fixation period, the Gabor patch and the auditory white noise burst were presented with some specified stimulus onset asynchrony (Figure 3.1B). After each presentation, participants accurately delivered their response by pressing ‘1’ on the keyboard if the pair of audiovisual stimuli was synchronous or by pressing ‘2’, if the pair was asynchronous. No feedback was provided on the correctness of their responses during the main experiment.

The SOAs for the pre- and post-deprivation simultaneity tasks were determined for each participant in an initial pre-testing phase using a method of constant stimuli with SOAs ranging between -500ms and 500ms and for 20 trials each. Negative and positive SOA values corresponded to auditory-preceding-vision and vision-preceding-auditory SOAs respectively. After the initial testing phase, the SOAs for the following probability of simultaneity judgments were selected. SOAs corresponding to 1%, 50% and the maximum probability of simultaneity judgements were considered as major SOAs (i.e., included in the EEG analysis) and presented for 100 trials (i.e., 20 trials in each block), while the SOAs corresponding to 25% and 75% probability were considered as minor (i.e., used in fitting Gaussian distribution but were not included in the EEG analysis) and presented for 25 trials (i.e., 5 trials in each block). The SOAs corresponding to 1% simultaneity response on both the auditory and visual leading sides were considered as unisensory trials in our analysis and hence, to ensure that there was enough duration after the first stimulus onset, these SOAs were set to at least 500ms and -500ms for visual and auditory leading conditions respectively.

3.3.3 Behavioral Analysis

Each participant completed 5 pre and 5 post deprivation blocks of the audiovisual simultaneity judgment task with EEG. Proportions of synchrony reports as a function of SOA were computed for each block and each participant and pooled across blocks for each session. Single-term Gaussian distribution models (through MATLAB fit.m) whose amplitude, mean and standard deviation (SD) were free to vary (Noel, De Niear, Van der Burg, and Wallace, 2016) were fitted to the data. Unlike the mean and the SD parameters, the range of possible amplitude values was bound between 0 and 1. The mean and the SD of the best fitting model were then taken as the point of subjective simultaneity (PSS) and the temporal window of simultaneity (TBW) respectively. The amplitude (AMP) of the distribution represented the maximum probability of simultaneity perception occurring at the PSS. The averaged r2 values across participants in both groups showed reasonable fits to the data for the combined pre-deprivation block (0.93±0.012) and the combined post-deprivation block (0.95±0.013).

The following statistical analyses were then carried out for each psychophysical measure using the JASP software version 0.11 (JASP Team, 2018). First, an independent samples t-test was carried out to determine whether participants in both groups were comparable in their pre-deprivation performance. Second, a two-way mixed ANOVA with Greenhouse-Geisser correction of sphericity was conducted to determine the interaction effect of group (deprived vs non-deprived) and session (pre vs post deprivation) on each of the psychophysical measures. A Bonferroni corrected pre-planned simple effect analysis of session with group as the moderating factor was conducted to determine whether deprivation induced any significant changes in the psychophysical measures for each group separately. All statistical analyses were two-tailed and significance level, alpha, was set to 0.05.

3.3.4 EEG Acquisition and Preprocessing

A Net Amps 400 amplifier and Hydrocel GSN 128 EEG cap (EGI Systems Inc., Eugene, OR) were used to record continuous EEG signal from 128 electrodes referenced to the vertex (Cz). EEG data were acquired at a sampling rate of 1000Hz with NetStation 5.3. Off-line preprocessing was performed with MATLAB using EEGLAB (Delorme and Makeig, 2004). Continuous EEG data were high pass filtered at 1Hz and line noise was removed using CleanLine(). The clean_rawdata() was applied to reject bad channels, correct continuous data using Artifact subspace reconstruction (ASR) and subsequently remove data that still contained artefacts after ASR. Additionally, channels that were located at the end of the nets (17 of them) were also removed. All channels that were removed were reconstructed using a spherical spine interpolation. The raw data were then epoched from -1s before the first stimulus onset to 1s after. We performed an independent component analysis (ICA) using runica() and manually rejected ICA components that reflected eye movements, muscle artifacts, channel noise and heart noise. Further trial rejection were conducted upon visual inspection. Finally, the EEG signals were re-referenced against the common average reference.

3.3.5 Time-Frequency Analysis

Preprocessed EEG data were analyzed in MATLAB using Fieldtrip toolbox (Oostenveld, Fries, Maris, and Schoffelen, 2011). Time-frequency decomposition of single trial EEG data was performed using convolution with Morlet wavelets that varied in frequencies from

4 to 80 Hz in 0.5 Hz steps and had a temporal resolution of 20ms. The number of cycles of the Morlet wavelets at the lowest and highest frequencies were 3 and 15 respectively. The power of single time frequency representations (TFRs) averaged across trials and inter-trial phase coherence (ITC) measures were determined for each participant (deprived/non-deprived), session (pre/post) and condition (visual-only, auditory-only, AV at 50%, VA at 50%, PSS). ITC indicates the phase concentration among all of the trials ranging from 0 to 1 with a value of 1 indicating perfect phase locking, while 0 indicates random phase distribution. To determine the effect of monocular deprivation on neural oscillatory activity, we focused our analysis on post-stimulus time periods. The power of the averaged TRFs and ITCs were baseline normalized to the -500 to -200ms pre-stimulus period by decibel and relative transformation respectively. For each participant and condition, deprivation induced change in the measures was computed as the difference between the baseline normalized data of the post and pre deprivation sessions.

For the visual only and auditory only condition, the computed relative (i.e., post-pre) data were averaged over time points 0 to 300ms to avoid any contamination from the onset of the second stimulus and over different frequency bands (alpha:8-12, low beta: 13-20, high beta: 21-30Hz). For the audiovisual conditions which are AV at 50% and VA at 50%, the computed relative data were averaged over time points 0 to 500ms and the over the different frequency bands.

3.3.6 Statistical Analysis

We tested whether the deprivation-induced (post-pre) change in power and inter-trial phase coherence differed between the two groups for each condition using a nonparametric cluster-based permutation procedure based on independent-sample t tests using the following parameters: two-sided t test, a level for thresholding individual points at p = 0.05, minimal number of neighbors in a cluster of three, t statistics performed on the maximal sum across cluster, 4000 randomizations.

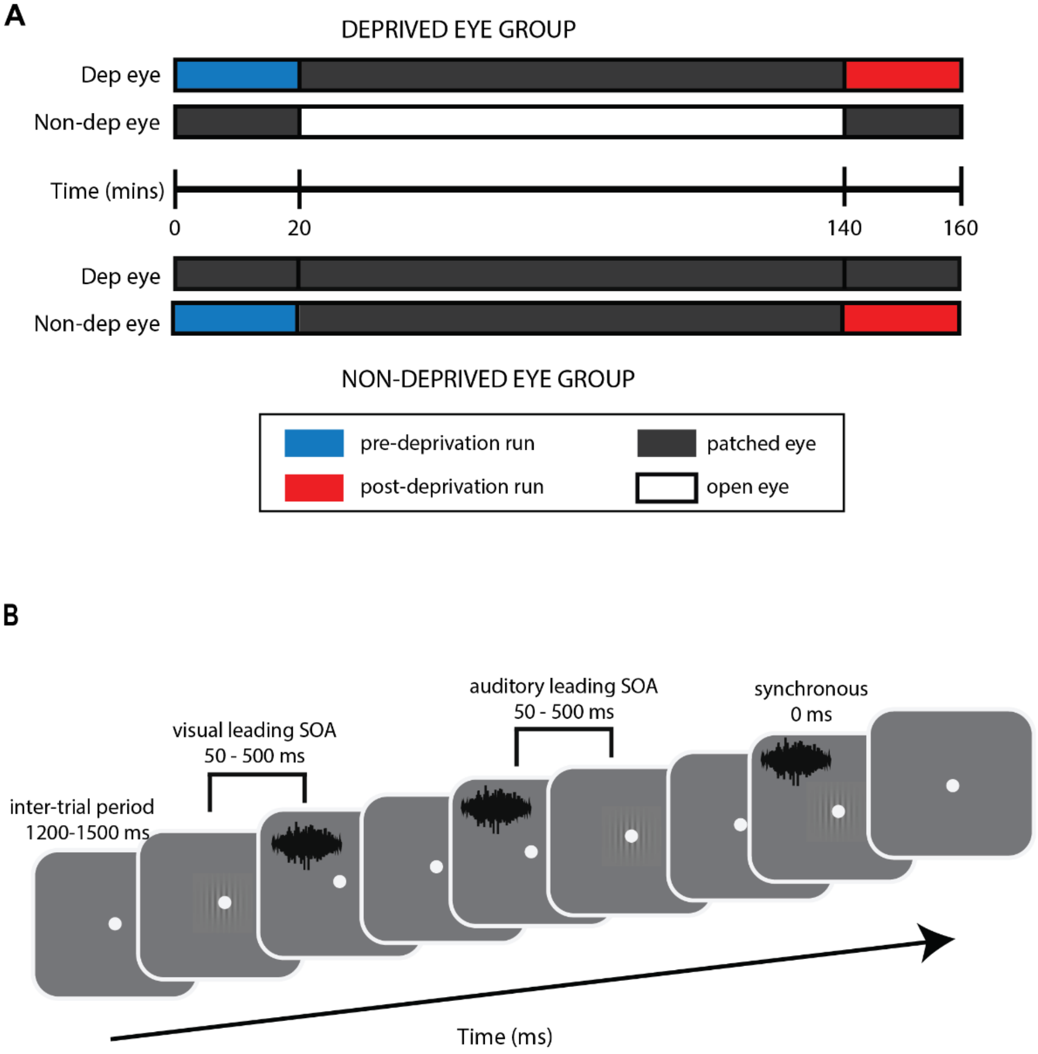

Figure 3. 1 Experimental procedure.

(A) Participants were randomly assigned to either the deprived eye or non-deprived eye group. Participants in the deprived eye group performed the audiovisual temporal simultaneity judgment task with their deprived eye before and after 120 minutes of deprivation. For the non-deprived eye group, the deprivation phase started with the first pre-deprivation block of the non-deprived eye. The post-deprivation blocks were then performed using the non-deprived eye after 90 minutes from the start of the first pre-deprivation block. (B) On each block, participants judged the simultaneity of a gabor patch (33ms) and an auditory white noise burst (33ms). The onsets of the two stimuli were separated by a range of pre-selected stimulus onset asynchronies (SOAs) between -500ms and 500ms. Negative and positive SOAs signified auditory leading and visual leading respectively. On each trial, there was a brief fixation period (1200-1500ms), followed by the stimulus presentation. Participants were then asked to respond by pressing the keyboard after which the next trial began automatically

3.4 Results

3.4.1 Behavioral Results

We determined psychophysical measures of audiovisual temporal asynchrony perception for the deprived eye and non-deprived eye groups before and after 120 minutes of monocular deprivation. At baseline (i.e., during the pre-deprivation test), our results showed that the two groups did not differ in any of the measures i.e., AMP ($t(14) = 0.088$, $p = 0.931$), PSS ($t(14) = -0.329$, $p = 0.747$) and the TBW ($t(14) = -0.488$, $p = 0.633$).

Next, we investigated whether the effect of 120 minutes of monocular deprivation on the psychophysical measures differed for the deprived eye and non-deprived eye groups. For the amplitude measure defined as the maximum probability of simultaneity perception, a 2 * 2 mixed ANOVA with group (deprived eye, non-deprived eye) as the between subject factor and session (pre, post) as the within subject factor revealed a significant main effect of session ($F(1, 14) = 5.948$, $p = 0.0286$, $\eta^2 = 0.0593$) but non-significant main effect of group ($F(1, 14) = 0.523$, $p = 0.481$, $\eta^2 = 0.036$) and interaction effect ($F(1, 14) = 3.332$, $p = 0.0893$, $\eta^2 = 0.033$; Figure 3.2A). The significant main effect of session indicates an overall reduction in the probability of simultaneity across the groups and thus, a possible reduced tendency to bind audiovisual stimuli after deprivation. Interestingly, a simple effect test with Bonferroni correction revealed that while amplitude of simultaneity perception decreased significantly after deprivation for the deprived eye group ($F(1) = 5.86$, $p = 0.046$), for the non-deprived eye group, the difference did not reach statistical significance ($F(1) = 0.42$, $p = 0.538$; Figure 3.2A). Combining both results highlights that the observed finding of a likely reduced tendency to bind audiovisual stimuli may apply more to the deprived eye group than the non-deprived eye group.

Furthermore, we conducted a 2 * 2 mixed ANOVA to determine the main effect of group (deprived eye, non-deprived eye), session (pre, post) and their interaction on PSS. The PSS is the mean of the synchrony probability distribution and defined as the SOA at maximum perceived simultaneity. Consistent with results from our previous study, there was no significant main effect of group (F(1, 14) = 2.247, p = 0.156, η^2 = 0.138), session (F(1, 14) = 1.05, p = 0.323, η^2 = 0.022) or interaction between group and session (F(1, 14) = 2.207, p = 0.160, η^2 = 0.046). A pre-planned simple effect analysis revealed no significant difference between pre and post deprivation PSS for the deprived eye group (F(1) = 2.615, p = 0.15) and non-deprived eye group (F(1) = 0.134, p = 0.726; Figure 3.2B).

However, for the TBW, a 2 * 2 mixed ANOVA with group as the between subject factor and session as the within subject factor revealed a significant interaction (F(1, 14) = 4.986, p = 0.0424, η^2 = 0.0041) indicating that the deprivation induced effect on the TBW differed between the two groups. Conversely, there was no significant main effect of group (F(1, 14) = 0.506, p = 0.488, η^2 = 0.035) or session (F(1, 14) = 0.094, p = 0764, η^2 = 7.665e-5). Next, we wanted to understand whether deprivation induced changes in TBW for each group separately. A simple effect test conducted on the data revealed a significant reduction in TBW for the deprived eye group (F(1) = 10.00, p = 0.0159, Figure 2C). For the non-deprived, there was a trend of an increase in the TBW after deprivation mimicking findings from our previous study, however, this trend was not statistically significant (F(1) = 1.106, p = 0.328, Figure 3.2C).

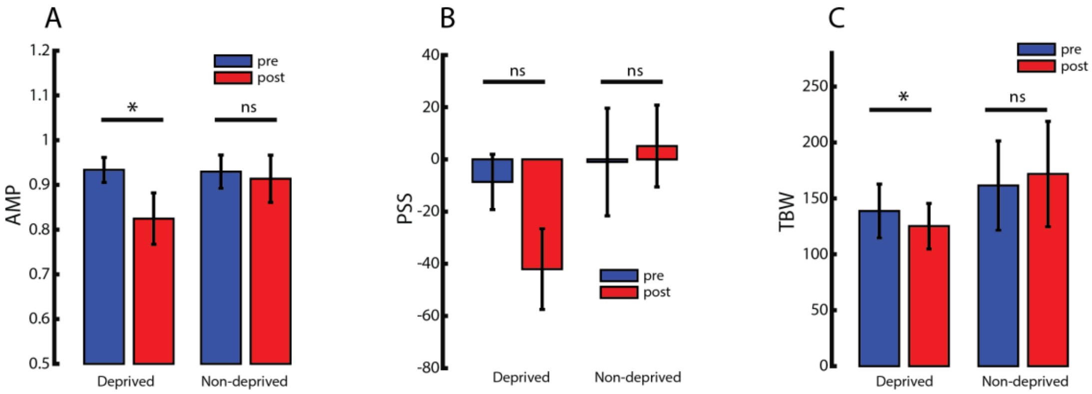

Figure 3. 2 Effects of monocular deprivation on the psychophysical measures of audiovisual temporal perception.

(A) Mean maximum probability of simultaneity perception (AMP) after deprivation is plotted for the pre (blue) and post (red) deprivation sessions for the deprive eye and non-deprived eye groups. (B) Mean point of subjective simultaneity (PSS) in milliseconds after deprivation is plotted for the pre (blue) and post (red) deprivation sessions for the deprive eye and non-deprived eye groups. (C) Mean width of the temporal binding window (TBW) in milliseconds after deprivation is plotted for the pre (blue) and post (red) deprivation sessions for the deprive eye and non-deprived eye groups. The error bars represent ±SEM. ns $p>0.05$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

3.4.2 Effects of Monocular Deprivation on Post-Stimulus Oscillatory Power

In the visual domain, short-term monocular deprivation has been shown to increase the power of post-stimulus alpha oscillations for the deprived eye and decrease the power for the non-deprived eye. Based on this evidence, we investigated the differential impact of short-term monocular deprivation on post-stimulus alpha band oscillatory activity for the deprived eye and non-deprived eye groups across both unisensory and multisensory stimulus conditions. Our analysis revealed a couple of interesting findings. First, we discovered a significant positive cluster over the parietal occipital sensors for all stimulus conditions (i.e., auditory only (cluster-value = 53.30, p = 0.022), visual only (cluster-value = 119.88, p = 0.0075), VA50% (cluster-value = 46.28, p = 0.011) and PSS (cluster-value = 59.25, p = 0.016) except the AV50% condition (cluster-value = 22.07, p = 0.057) where the cluster was near significant. Additionally, for the visual only condition, we observed significant positive clusters over the central (cluster-value = 119.88, p = 0.0075) and frontal sensors (cluster-value = 52.11, p = 0.022). Turning to the other oscillatory bands, we observed significant positive clusters in low and high beta band for some of the stimulus conditions. For the low beta band, the significant positive clusters were located at central parietal electrodes during the visual only condition (cluster-value = 99.16, p = 0.0047) and parietal-occipital electrodes for both the VA50% (cluster-value = 32.29, p = 0.024) and PSS conditions (cluster-value = 61.664, p = 0.013). For the high beta band, the clusters were identified parietal occipital regions during the AV50% (cluster-value = 17.29, p = 0.0355) and VA50% conditions (cluster-value = 24.98, p = 0.0312).

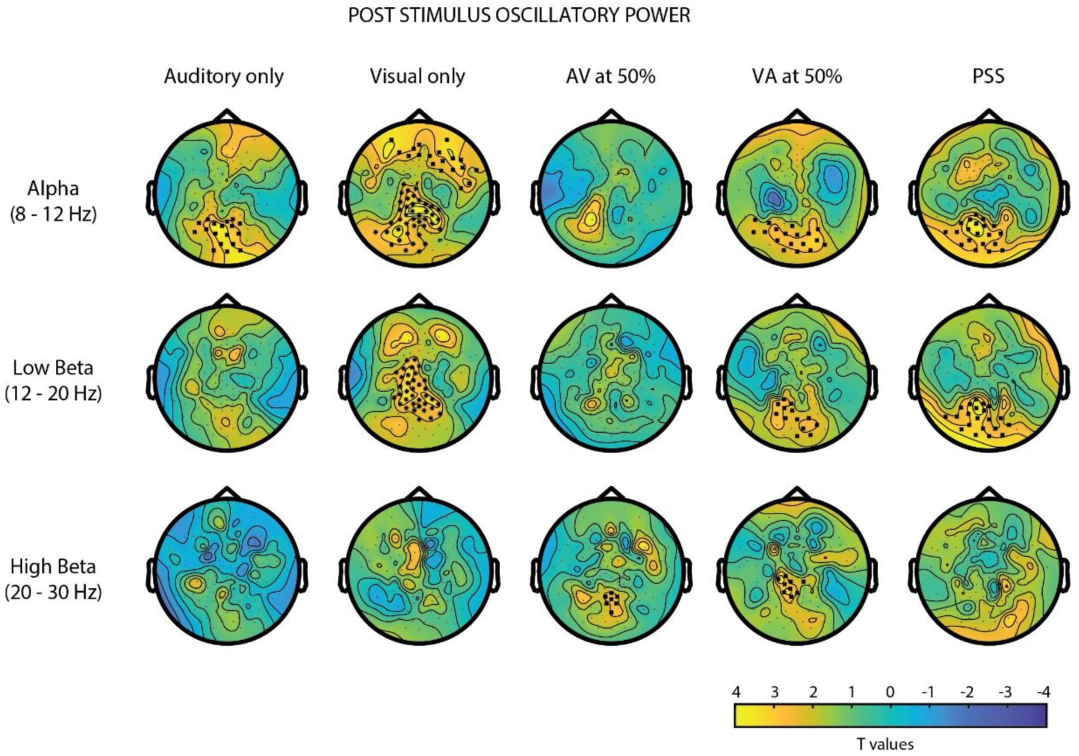

Figure 3. 3 Effects of monocular deprivation on alpha (8-12), low beta (12-20) and high beta (20-30) post-stimulus oscillatory power.

Statistical maps showing t-values from independent samples t-tests between the deprivation-induced (post-pre) change in oscillatory power for the deprived eye and non-deprived eye groups averaged across post-stimulus time points for the different stimulus conditions (0-300ms for visual and auditory only, and 0-500ms for the audiovisual conditions). AV – auditory leading, VA – visual leading. The x symbol (**^{x}) represent** channel clusters with $p < 0.05$ while the asterisk (*) represent channel clusters with $p < 0.01$

3.4.3 Effects of Monocular Deprivation on Post-Stimulus Oscillatory Inter-Trial Phase Coherence

Next, we investigated whether monocular deprivation differentially influenced post-stimulus alpha band ITC for the deprived eye and non-deprived eye groups. Our results show that across all stimulus conditions, only the visual condition showed significant positive clusters. These clusters were located over frontal (cluster-value = 62.33, p = 0.011) and parietal-occipital (cluster-value = 25.65, p = 0.045) sensors. Besides the alpha band, we observed significant clusters in the low and high beta bands. For the low beta band, we found two significant positive clusters located over frontal and parieto-occipital sensors during the visual only (frontal: cluster-value = 54.91, p = 0.005; parieto-occipital: cluster-value = 24.81, p = 0.027) and the VA50% (frontal: cluster-value = 33.96, p = 0.0085; parieto-occipital: cluster-value = 22.95, p = 0.023) conditions. For the high beta band, one significant negative cluster over the frontal electrodes was found during the auditory only condition (cluster-value = -19.99, p = 0.018) and one significant positive cluster over the right parietal electrodes (cluster-value = 22.34, p = 0.034) was found during the visual only condition.

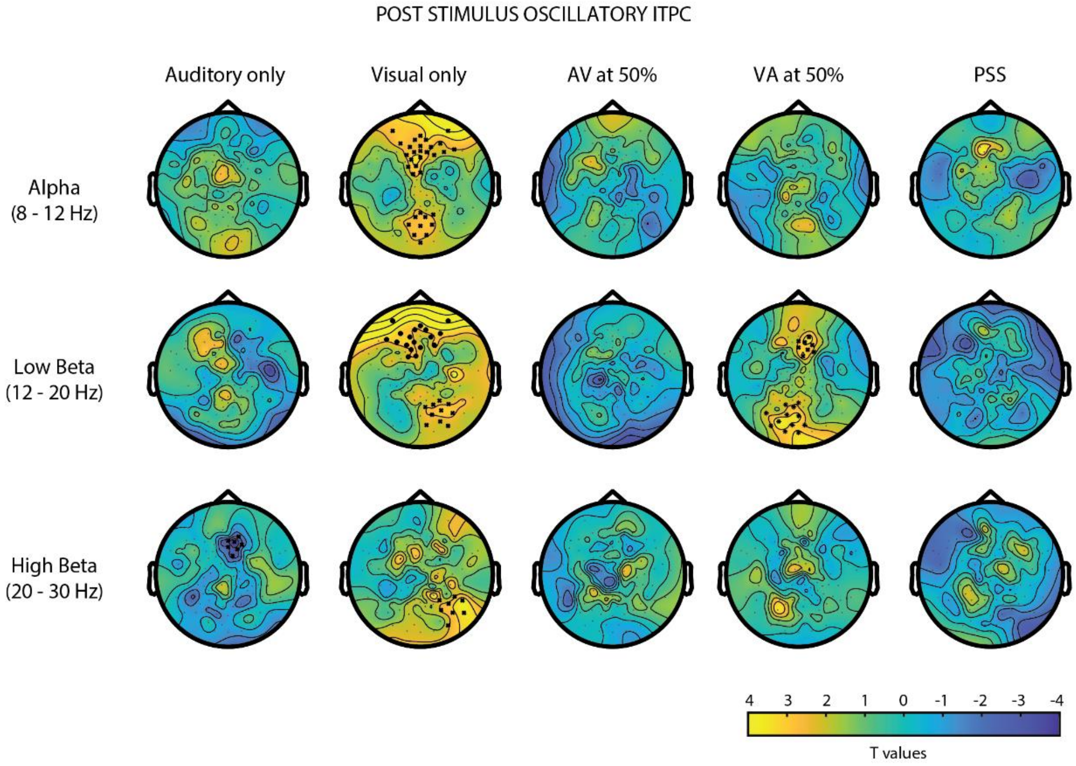

Figure 3. 4 Effects of monocular deprivation on alpha (8-12), low beta (12-20) and high beta (20-30) post-stimulus oscillatory inter-trial phase coherence (itpc).

Statistical maps showing t-values from independent samples t-tests between the deprivation-induced (post-pre) change in oscillatory itpc for the deprived eye and non-deprived eye groups averaged across post-stimulus time points for the different stimulus conditions (0-300ms for visual and auditory only, and 0-500ms for the audiovisual conditions). AV – auditory leading, VA – visual leading. The x symbol (^{x}) **represent** channel clusters with $p < 0.05$ while the asterisk (*) represent channel clusters with $p < 0.01$

3.5 Discussion

In adult humans, short-term monocular deprivation can induce changes in multisensory temporal perception specifically by narrowing the temporal binding window in the deprived eye and widening the window in the non-deprived eye (see Chapter 2). However, the neural mechanisms underlying this intriguing plasticity phenomenon have not yet been explored. Here, using EEG recordings in human participants, we show that both alpha and beta oscillations are involved in the short-term monocular deprivation effects on multisensory temporal perception.

In general, our behavioral findings are consistent with findings from our previous study (see Chapter 2) and those of previous visual studies that revealed opposing effects of monocular deprivation on binocular rivalry and contrast detection for the deprived and non-deprived eyes (Lunghi et al., 2011; Lunghi et al., 2013; Zhou et al., 2013). Nevertheless, we found a non-significant trend of increased TBW for non-deprived eye group, which does not align with our previous finding. Examining this further, in this study, we discovered that the percentage of participants in the non-deprived eye group that showed increased TBW was 50%, compared to 80% in our previous study (see Supplementary Figure 3.5). While almost all participants in the deprived eye group in our previous study showed the trend of reduced TBW after deprivation, in the present study, one participant did not show reduced TBW (see Supplementary Figure 3.5). This seeming variability in our results indicate that different persons may vary in their susceptibility to the effects of monocular deprivation on multisensory temporal perception. Furthermore, the effect monocular deprivation has on multisensory temporal perception may be robust when the deprived eye is tested compared with the non-deprived eye. Indeed, this explanation is consistent with findings from a study showing that after monocular deprivation, the effect of visuo-haptic interaction on perceptual dominance disappeared for the deprived eye, which was potentiated in the visual domain, but was not affected for the non-deprived eye, which was weakened (Lo Verde, Morrone, and Lunghi, 2017). Nevertheless, the fact that in this present study, the percentage of participants who experienced decreased TBW was more in the deprived eye group than the non-deprived eye group and those who experienced increased TBW for the non-deprived eye group were more than the deprived eye group hints at the presence of the opposing effect of monocular deprivation on the TBW. To remedy this, more data need to be collected to increase the statistical power in this study. While the mechanisms underlying short-term monocular deprivation effects on TBW can be attributed to changes in sensory encoding and or changes in decisional mechanisms (Magnotti, Ma, and Beauchamp, 2013), interestingly, this finding of decreased maximum probability of simultaneity judgments for the deprived eye supports the role of cognitive decisional factors and suggests that the effects of monocular deprivation on cross-modal processes may be dependent on the eye used.

Previous studies have established that the power and phase of neural oscillations play important roles in the multisensory processing, perception, and cognition (Keil and Senkowski, 2018). The presence of changes in oscillatory power and or phase may suggest different underlying neural activity and mechanisms. For instance, increases in oscillatory inter-trial phase coherence accompanied by increases in power may reflect evoked type responses (i.e., related to stimulus) while unaccompanied changes in inter-trial phase coherence may suggest phase resetting, that is, reorganizing of the phase of ongoing neural oscillations (Lakatos, Chen, O'Connell, Mills, and Schroeder, 2007; Lakatos et al., 2009).

Evidence of changes in oscillatory power without phase changes may indicate neural activity that is not strictly phase-locked to any stimulus but can be modulated by cognitive processes (Keil and Senkowski, 2018; Tallon-Baudry and Bertrand, 1999).

In this study, we found evidence of both increases in oscillatory power and inter-trial phase coherence in the alpha band for the deprived eye during conditions where the visual stimulus was presented first, which are the visual only stimulation and the VA at 50% condition. According to the framework suggested by Lakatos et al. (2009), these findings represent an increase in evoked alpha oscillations when the visual stimulus was presented to the deprived eye after monocular deprivation. Such an explanation agrees with Lunghi et al. (2015) findings that short-term monocular deprivation increased the amplitude of evoked alpha oscillations for the deprived eye and decreased the amplitude for the non-deprived eye during visual stimulation. However, unlike Lunghi et al. (2015) findings where changes in alpha oscillations were observed in a central scalp location, in our study, we found changes in alpha oscillatory power in central-occipital-parietal areas as well as frontal areas especially during the visual-only condition. Interestingly, previous studies have suggested that alpha activity may play different roles in sensory processing depending on the scalp configuration (Noonan et al., 2016; Sadaghiani and Kleinschmidt, 2016; Wöstmann, Alavash, and Obleser, 2019). While parieto-occipital alpha is associated with spatial attention and perceptual sensitivity indicating mechanisms of sensory encoding and overall cortical excitability (Romei, Gross, and Thut, 2010; Sauseng et al., 2005; Thut, Nietzel, Brandt, and Pascual-Leone, 2006), frontal alpha is associated with cognitive control and task-relevant information selection indicating decision-making mechanisms independent of sensory evidence (Sadaghiani and Kleinschmidt, 2016; Sauseng et al., 2005; Wöstmann et al., 2019). Thus, our findings imply the presence of both sensory encoding and decisional mechanisms underlying the effects of monocular deprivation.

Considering the fact that deprivation occurred in the visual domain, we were not expecting any changes in oscillatory activity during auditory-only stimulation after deprivation. However, surprisingly, monocular deprivation produced significant increases in auditory occipital alpha oscillations for the deprived eye compared with the non-deprived eye. Interestingly, this finding may support behavioral findings that short-term light deprivation causes improvement in performance in auditory tasks such as spatial localization (Lewald, 2007) and harmony perception (Landry, Shiller, and Champoux, 2013). Taken together, these findings suggest that monocular deprivation induces supramodal compensatory mechanisms that seek to enhance processing of cross-modal information following the absence of visual stimulation during deprivation.

Similar to the changes in alpha oscillatory activity, we found that monocular deprivation increased low beta power and inter-trial phase coherence for the visual condition and VA at 50% conditions. While these findings suggest evoked activity to the onset of the visual stimulus, increases in power unaccompanied by changes in phase-locking activity for the deprived eye was identified in the high beta ranged for only the audiovisual conditions (i.e., AV at 50% and VA at 50%). Consistent with our findings, narrowing of the TBW after perceptual learning has been associated with increased beta-band activity over parietal and occipital sensors (Theves, Chan, Naumer, and Kaiser, 2020). In contrast, decreased high beta power was predictive of correct responses during an audiovisual simultaneity judgment task (Grabot et al., 2017). Furthermore, decreased beta power was related to better sensory encoding (Griffiths et al., 2019) and conversely, increase beta power to mechanisms underlying the perception of illusory phenomena (Kaiser, Senkowski, Busch, Balz, and Keil, 2019). While these findings are equivocal, they indicate that beta activity play a major role in the changes in multisensory temporal perception associated with short-term monocular deprivation.

Although, our findings implicate both alpha and beta oscillations in the multisensory temporal perceptual changes following short-term monocular deprivation, they do not directly provide evidence to support the brain regions involved. Moreover, we are not able to determine whether the perceptual or oscillatory changes observed reflect sensory encoding and or decisional processes. In the future, Bayesian casual inference model could be employed to provide a mechanistic understanding of whether these changes in multisensory temporal perception after monocular deprivation can be expressed in parameters that reflect sensory encoding processes and or those that reflect changes in decisional priors. In addition, from a neuroimaging standpoint, source localization analysis of EEG signals can be used to probe which areas in the brain are involved in the changes observed after monocular deprivation. Furthermore, it will be interesting to know whether there are strengthening or weakening of functional relationships between various areas such early sensory and high association areas following short-term monocular deprivation. Lastly, it will be interesting to determine whether correlations exist between parameters of Bayesian causal inference models and source localization measures obtained from EEG source localization analysis.

3.6 References

Bastiaansen, M., Berberyan, H., Stekelenburg, J. J., Schoffelen, J. M., and Vroomen, J. (2020). Are alpha oscillations instrumental in multisensory synchrony perception? *Brain research, 1734*, 146744.

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

Chen, Y.-C., Lewis, T. L., Shore, D. I., and Maurer, D. (2017). Early binocular input is critical for development of audiovisual but not visuotactile simultaneity perception. *Current Biology, 27*(4), 583-589.

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods, 134*(1), 9-21.

Diederich, A., and Colonius, H. (2015). The time window of multisensory integration: Relating reaction times and judgments of temporal order. *Psychological review, 122*(2), 232.

Driver, J., and Noesselt, T. (2008). Multisensory interplay reveals cross-modal influences on 'sensory-specific'brain regions, neural responses, and judgments. *Neuron, 57*(1), 11-23.

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

Frassinetti, F., Bolognini, N., and Làdavas, E. (2002). Enhancement of visual perception by cross-modal visuo-auditory interaction. *Experimental brain research, 147*(3), 332-343.

Grabot, L., Kösem, A., Azizi, L., and Van Wassenhove, V. (2017). Prestimulus alpha oscillations and the temporal sequencing of audiovisual events. *Journal of cognitive neuroscience, 29*(9), 1566-1582.

Griffiths, B. J., Mayhew, S. D., Mullinger, K. J., Jorge, J., Charest, I., Wimber, M., and Hanslmayr, S. (2019). Alpha/beta power decreases track the fidelity of stimulus-specific information. *Elife, 8*, e49562.

Hairston, W. D., Laurienti, P. J., Mishra, G., Burdette, J. H., and Wallace, M. T. (2003). Multisensory enhancement of localization under conditions of induced myopia. *Experimental brain research, 152*(3), 404-408.

Hillock-Dunn, A., and Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental science, 15*(5), 688-696.

Hillock, A. R., Powers, A. R., and Wallace, M. T. (2011). Binding of sights and sounds: age-related changes in multisensory temporal processing. *Neuropsychologia, 49*(3), 461-467.

JASP Team. (2018). JASP (Version 0.11)[Computer software].

Kaiser, M., Senkowski, D., Busch, N. A., Balz, J., and Keil, J. (2019). Single trial prestimulus oscillations predict perception of the sound-induced flash illusion. *Scientific reports, 9*(1), 1-8.

Kambe, J., Kakimoto, Y., and Araki, O. (2015). Phase reset affects auditory-visual simultaneity judgment. *Cognitive neurodynamics, 9*(5), 487-493.

Keil, J., and Senkowski, D. (2018). Neural oscillations orchestrate multisensory processing. *The Neuroscientist, 24*(6), 609-626.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., and Shams, L. (2007). Causal inference in multisensory perception. *PloS one, 2*(9). doi: 10.1371/journal.pone.0000943

Lakatos, P., Chen, C.-M., O'Connell, M. N., Mills, A., and Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron, 53*(2), 279-292.

Lakatos, P., O'Connell, M. N., Barczak, A., Mills, A., Javitt, D. C., and Schroeder, C. E. (2009). The leading sense: supramodal control of neurophysiological context by attention. *Neuron, 64*(3), 419-430.

Landry, S. P., Shiller, D. M., and Champoux, F. (2013). Short-term visual deprivation improves the perception of harmonicity. *Journal of Experimental Psychology: Human Perception and Performance, 39*(6), 1503.

Lewald, J. (2007). More accurate sound localization induced by short-term light deprivation. *Neuropsychologia, 45*(6), 1215-1222.

Lo Verde, L., Morrone, M. C., and Lunghi, C. (2017). Early cross-modal plasticity in adults. *Journal of Cognitive Neuroscience, 29*(3), 520-529.

London, R. E., Benwell, C. S., Cecere, R., Quak, M., Thut, G., and Talsma, D. (2020). EEG Alpha Power Predicts the Temporal Sensitivity of Multisensory Integration. *bioRxiv*.

Lovelace, C. T., Stein, B. E., and Wallace, M. T. (2003). An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Cognitive brain research, 17*(2), 447-453.

Lunghi, C., Berchicci, M., Morrone, M. C., and Di Russo, F. (2015). Short-term monocular deprivation alters early components of visual evoked potentials. *The Journal of physiology, 593*(19), 4361-4372.

Lunghi, C., Burr, D. C., and Morrone, C. (2011). Brief periods of monocular deprivation disrupt ocular balance in human adult visual cortex. *Current Biology, 21*(14), R538-R539.

Lunghi, C., Burr, D. C., and Morrone, M. C. (2013). Long-term effects of monocular deprivation revealed with binocular rivalry gratings modulated in luminance and in color. *Journal of Vision, 13*(6), 1-1.

Magnotti, J. F., Ma, W. J., and Beauchamp, M. S. (2013). Causal inference of asynchronous audiovisual speech. *Frontiers in psychology, 4*, 798.

Nidiffer, A. R., Stevenson, R. A., Krueger Fister, J., Barnett, Z. P., and Wallace, M. T. (2016). Interactions between space and effectiveness in human multisensory performance. *Special Issue: Synaesthesia and Multisensory Processes, 88*, 83-91. doi: 10.1016/j.neuropsychologia.2016.01.031

Noel, J.-P., De Niear, M., Van der Burg, E., and Wallace, M. T. (2016). Audiovisual simultaneity judgment and rapid recalibration throughout the lifespan. *PLoS One, 11*(8), e0161698.

Noonan, M. P., Adamian, N., Pike, A., Printzlau, F., Crittenden, B. M., and Stokes, M. G. (2016). Distinct mechanisms for distractor suppression and target facilitation. *Journal of Neuroscience, 36*(6), 1797-1807.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience, 2011*.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision, 10*(4), 437-442.

Powers, A. R., Hevey, M. A., and Wallace, M. T. (2012). Neural correlates of multisensory perceptual learning. *Journal of Neuroscience, 32*(18), 6263-6274.

Powers, A. R., Hillock, A. R., and Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *The Journal of neuroscience, 29*(39), 12265-12274.

Richards, M. D., Goltz, H. C., and Wong, A. M. (2017). Alterations in audiovisual simultaneity perception in amblyopia. *PloS one, 12*(6), e0179516.

Romei, V., Gross, J., and Thut, G. (2010). On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: correlation or causation? *Journal of Neuroscience, 30*(25), 8692-8697.

Sadaghiani, S., and Kleinschmidt, A. (2016). Brain networks and α-oscillations: structural and functional foundations of cognitive control. *Trends in Cognitive Sciences, 20*(11), 805-817.

Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., . . . Birbaumer, N. (2005). A shift of visual spatial attention is selectively associated with human EEG alpha activity. *European journal of neuroscience, 22*(11), 2917-2926.

Shams, L., and Beierholm, U. R. (2010). Causal inference in perception. *Trends in cognitive sciences, 14*(9), 425-432. doi: 10.1016/j.tics.2010.07.001

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

Stein, B. E., and Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature reviews neuroscience, 9*(4), 255-266. doi: 10.1038/nrn2331

Stevenson, R. A., and Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Experimental brain research, 227*(2), 249-261.

Tallon-Baudry, C., and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in cognitive sciences, 3*(4), 151-162.

Theves, S., Chan, J. S., Naumer, M. J., and Kaiser, J. (2020). Improving audio-visual temporal perception through training enhances beta-band activity. *Neuroimage, 206*, 116312.

Thut, G., Nietzel, A., Brandt, S. A., and Pascual-Leone, A. (2006). α-Band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *Journal of Neuroscience, 26*(37), 9494-9502.

Wallace, M. T., and Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia, 64*, 105-123. doi: 10.1016/j.neuropsychologia.2014.08.005

Wöstmann, M., Alavash, M., and Obleser, J. (2019). Alpha oscillations in the human brain implement distractor suppression independent of target selection. *Journal of Neuroscience, 39*(49), 9797-9805.

Zhou, J., Clavagnier, S., and Hess, R. F. (2013). Short-term monocular deprivation strengthens the patched eye's contribution to binocular combination. *Journal of Vision, 13*(5), 12-12.

Zou, H., Müller, H. J., and Shi, Z. (2012). Non-spatial sounds regulate eye movements and enhance visual search. *Journal of Vision, 12*(5), 2-2.

3.7 Supplementary Figures

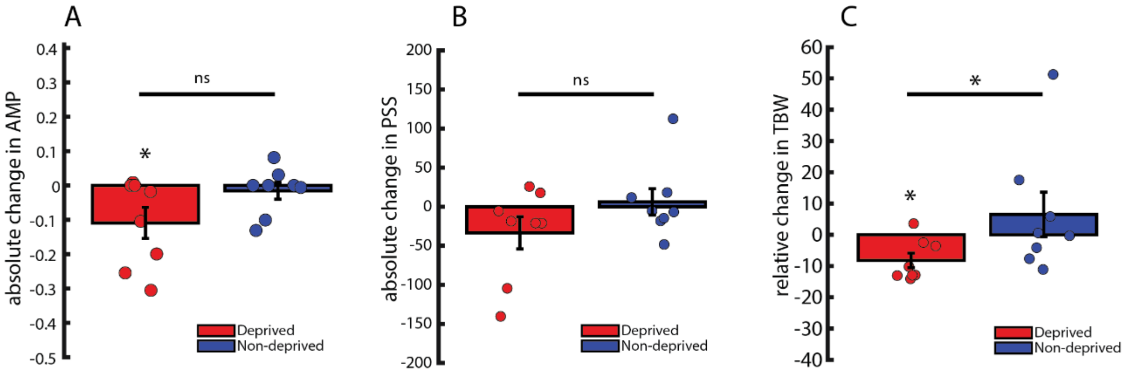

Figure 3.5 Plots of deprivation-induced (post-pre) changes for the psychophysical measures of audiovisual temporal perception.

(A) Mean absolute change in maximum probability of simultaneity perception (AMP) after deprivation is plotted for the deprived (red) and non-deprived (blue) eye groups. Absolute change was computed as post – pre. (B) Mean absolute change in point of subjective simultaneity (PSS) after deprivation is plotted for the deprived (red) and non-deprived (blue) eye groups. (C) Mean relative change in TBW after deprivation is plotted for the deprived (red) and non-deprived (blue) eye groups. Relative change was computed as (post – pre)/pre * 100%. Circular dots represent individual subject data point within each eye group. The error bars represent ±SEM. ns $p>0.05$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Chapter 4

A Temporally Correlated Sound Facilitates Visual Learning of Contrast Detection

4.1 Abstract

Perceptual learning paradigms that engage multiple sensory modalities can improve visual sensitivity more than visual-only paradigms. However, the visual stimulus features (e.g., contrast) or tasks (e.g., detection) for which visual learning may benefit from multisensory facilitation have not been quite explored. This study tested the hypothesis that learning to detect visual stimulus contrast will benefit from multisensory facilitation by a temporally correlated auditory stimulus. Two groups (N=10 each) of participants were recruited and trained to detect a Gabor patch undergoing counterphase flickering (~21Hz) in the presence of varying levels of external noise. Both groups performed a 2AFC contrast detection task without feedback in the pre (day 1-2) and post-training (day 9-10) sessions. During the training sessions (day 3-8), the visual-only training (VOT) group performed the task with feedback while the audiovisual training (AVT) group performed the task with feedback and an auditory white noise. The auditory stimulus was amplitude modulated with the same frequency and duration as the visual stimulus in order to enhance temporal binding. Contrast detection thresholds for each external noise level were measured using an adaptive staircase procedure. The perceptual template model (PTM) was adopted to investigate whether different mechanisms underlie the two perceptual training paradigms. While both groups improved with practice, the AVT group showed significantly more reduction in

threshold than the VOT group during the training sessions when sound was present for AVT and after training. In addition, AVT produced performance enhancements in conditions where the visual stimulus was weakly effective, highlighting the principle of inverse effectiveness in multisensory interactions. Analysis from the PTM model revealed that while external noise reduction and changes in transducer nonlinearities underlie performance enhancement under AVT training, for VOT, training-induced performance enhancement was accounted for by only external noise reduction mechanism. Compared to visual PL, our results suggest that multisensory PL paradigms are more effective for visual learning, may engage separate and or additional mechanisms and thus, will provide a powerful new set of rehabilitative tools in the quest to improve visual function in patients with low vision.

4.2 Introduction

Repeatedly training adult humans can improve their ability to detect and discriminate visual stimuli with different levels of contrast, texture, and motion (Ball and Sekuler, 1982; Fiorentini and Berardi, 1980; Karni and Sagi, 1991; Mayer, 1983). This phenomenon known as perceptual learning demonstrates that the adult brain retains a remarkable degree of cortical plasticity that can support the acquisition of new skills and more importantly, recovery and rehabilitation after injury occurring in conditions such as amblyopia (Levi and Li, 2009; Polat, Ma-Naim, Belkin, and Sagi, 2004; Polat, Ma-Naim, and Spierer, 2009), hemianopia (Kasten, Poggel, and Sabel, 2000; Sahraie et al., 2010), glaucoma and age-related macular degeneration (Camilleri, Pavan, Ghin, Battaglini, and Campana, 2014; Durrie and McMinn, 2007; Polat et al., 2012).

Given that our brains have evolved to function optimally in a multisensory world (Shams and Seitz, 2008), leveraging the integrative function among the senses presents a powerful platform for learning, particularly when information from one sense is reduced or compromised. Indeed, perceptual learning protocols that train with cross-modal stimuli provide significantly improved learning outcomes on unisensory tasks when compared with unisensory training protocols. For example, training with a congruent audiovisual stimulus on a coherent motion detection task produced faster and more improved detection performance on visual only trials compared with training using only the visual stimulus (Seitz, Kim, and Shams, 2006). Interestingly, performance on the hardest-difficulty trials (i.e., when the coherence of the motion stimuli was near random) improved significantly only after audiovisual training, highlighting a guiding principle of multisensory interactions known as inverse effectiveness (Holmes, 2009; Stein and Meredith, 1993). Mechanistically, improved learning outcomes via multisensory perceptual training may stem from the activation and strengthening of a wider network of brain areas consisting of both unisensory and multisensory regions (Powers, Hevey, and Wallace, 2012; Zilber, Ciuciu, Gramfort, Azizi, and Van Wassenhove, 2014).

Although integrating information from different senses may yield more effective, more efficient, and perhaps more generalizable visual learning outcomes, it is still unclear as to the visual stimulus features (e.g., contrast vs motion) and or tasks (e.g., detection vs discrimination) for which visual learning may benefit from this multisensory facilitation. For instance, arbitrary pairing of a sinusoidal grating of a given orientation with a particular task-uninformative auditory tone only improved detection performance of the grating in the presence but not in the absence of the auditory tone (Shams, Wozny, Kim, and Seitz, 2011; Wozny, Seitz, and Shams, 2008). Consequently, Shams et al. (2011) proposed that multisensory facilitation of unisensory learning may only occur for ecologically associated auditory and visual features (such as motion, or lip movements and voice, etc.) where ecologically valid associations are characterized by their hardwired connectivity in the brain or synaptic structures that can only be learned during the critical period. Interestingly, these "ecologically valid" cross-modal features usually share very complex and robust spatial and temporal correspondence. For example, during speech, the area of the mouth opening is highly correlated in space and time to the acoustic envelope of the speech sound (Chandrasekaran, Trubanova, Stillittano, Caplier, and Ghazanfar, 2009). Therefore, it is possible that this observed multisensory-facilitated visual learning might occur for cross-modal stimuli that are not necessarily ecologically associated but share some spatial-temporal correspondence that enhances cross-modal binding.

To this end, we tested the hypothesis that learning to detect the contrast of a Gabor signal undergoing counterphase flicker will benefit from multisensory facilitation by a temporally correlated auditory stimulus. Specifically, we trained two groups of participants: one on a visual only paradigm (VOT) and the other on an audiovisual training paradigm (AVT). We measured the performance of both groups on a visual contrast detection task before and after training. In addition, we investigated how the perceptual system may change during this multisensory-facilitated visual perceptual learning using the perceptual template model (PTM) developed by Dosher and colleagues (Dosher and Lu, 1999; Lu and Dosher, 1999; Lu and Dosher, 2004). Briefly, the PTM is an elaborate version of the simple linear amplifier model with the incorporation of two nonlinear components namely nonlinear transducer and multiplicative noise. Thus, the PTM highlights four possible and distinct mechanisms that can underlie perceptual learning (Dosher and Lu, 2017; Lu and Dosher, 1999). These mechanisms are stimulus enhancement (i.e., reduction in internal noise), external noise exclusion, reduction in multiplicative noise, and changes in nonlinear transducer properties of the system. While stimulus enhancement and external noise exclusion have been previously reported to account for perceptual learning of several visual tasks (Bejjanki et al., 2014; Dosher and Lu, 1998, 1999; Lu, Chu, Dosher, and Lee, 2005; Lu and Dosher, 1999; Lu and Dosher, 2004), to date no studies have reported reduction in multiplicative noise or changes in nonlinear transducer properties as possible mechanisms of visual perceptual learning. In this study, we show that multisensory perceptual learning paradigms are more effective for visual learning of contrast compared to visual only paradigms. Moreover, to our knowledge, we provide the first experimental evidence that improved visual performance after perceptual learning (i.e., multisensory) can occur as a result of changes in nonlinear transducer properties of the system, a finding that reflects the presence of inverse effectiveness arising from multisensory interactions during multisensory perceptual learning. Together, our results imply that multisensory perceptual learning may engage separate and or additional mechanisms compared to visual-only learning and thus, will provide a powerful new set of rehabilitative tools in the quest to improve visual function in patients with low vision.

4.3 Methods

4.3.1 Participants

Twenty-four participants were recruited and randomly assigned to either the visual-only training (VOT) group or the audiovisual training (AVT) group. Each participant had normal or corrected-to-normal vision and normal hearing. Normal or corrected-to-normal vision was confirmed for each participant using the Snellen acuity chart. Before enrolling in the study, participants gave informed consent, and after the study, received monetary compensation. Of the 24 participants initially enrolled, four (2 VOT and 2 AVT) could not complete all sessions due to sickness or failure to comply with the demanding schedule of the training or reasons that could not be determined. The remaining sample comprised 10 VOT participants (6 females, mean (SD) age: 20.3 (1.4) years) and 10 AVT participants (5 females, mean (SD) age: 20.8 (2.4) years). The recruitment and experimental procedures for this work were carried out in accordance with the Declaration of Helsinki and were approved by the Vanderbilt University Institutional Review Board.

4.3.2 Apparatus

The experimental stimuli and all trial-related events were controlled using MATLAB (Math Works Inc., Natick, MA) software in conjunction with the Psychophysics Toolbox Version 3 (Brainard and Vision, 1997; Denis G Pelli, 1997). The visual stimuli were presented on a gamma-corrected NEC CRT monitor (MultiSync FE992 18inch) with ~75cdm^2 background luminance, 85Hz refresh rate and 1024*768 pixel resolution. A video switcher that combines two eight-bit output channels of the video card to produce 14 bits of gray levels was used to enhance the gray level resolution of the CRT monitor (Li and Lu, 2012; Li, Lu, Xu, Jin, and Zhou, 2003). The auditory stimulus was presented binaurally through headphones (Sennheiser HD559). A Minolta Chroma Meter CS-100 and a sound level meter were used to verify the luminance and sound intensity levels, respectively. The durations of all visual and auditory stimuli, as well as their onset times, were confirmed using a Hameg 507 oscilloscope with a photovoltaic cell and microphone.

4.3.3 Stimuli

The visual stimulus was a horizontally oriented Gabor signal undergoing counterphase flickering in the presence of varying levels of external noise. The luminance profile of the Gabor signal can be defined by the equation below (Bejjanki et al., 2014):

$$L(x, y) \ = l_0 \ \times \left(1.0 + csin\left(2\pi f\left(\cos(\theta) + ysin(\theta)\right)\right) \times \exp\left(-\frac{x^2 + \ y^2}{2\sigma^2}\right)\right)$$

where θ, angle of orientation = rad (0); l_0, background luminance = 75cd/m2; f, Gabor center frequency = 1 cycle/degree; σ, the standard deviation of the Gabor's circular Gaussian envelope = 0.5deg; phi, the phase of the Gabor signal alternated between 0 or 180deg at ~21Hz frequency; c, the contrast of the Gabor signal. To generate external noise images, pixel contrasts were drawn independently from identical Gaussian distributions (Lu and Dosher, 2004). Seven external noise contrast levels (0, 0.021, 0.083, 0.165, 0.248, 0.33, and 0.45) were selected for this study based on a brief initial pilot study carried out to ensure that each region of the threshold vs external noise curve (i.e., the region where external noise is smaller than internal noise and vice versa) contain enough external noise contrast levels. Moreover, the external noise images were band-pass filtered from one octave below to one octave above the spatial frequency of the Gabor signal (Lu and Dosher, 2004). The size of both the Gabor patterns and the noise frames were set at 2.0 x 2.0 deg. The auditory stimulus was a white noise signal that was square-wave amplitude modulated with a fundamental frequency of ~21Hz and a modulation depth of 50%. The frequency and duration of the visual and auditory stimuli were intentionally matched to enhance temporal correspondence and cross-modal binding.

4.3.4 Design and Procedure

Each participant in both groups completed 10 sessions consisting of 2 pre-training sessions (sessions 1-2), 6 training sessions (sessions 3-8) and 2 post-training sessions (sessions 9-10). During the pre- and post-training sessions, participants in both groups performed the visual contrast detection task with no feedback on the correctness of their response, and during the training sessions the visual-only training (VOT) group performed the task with feedback while the audiovisual training (AVT) group performed the task with feedback and an auditory white noise. Individual sessions occurred on separates day, with each session lasting approximately an hour. No two consecutive sessions were allowed to occur more than 2 days apart.

All experimental procedures were performed inside a Whisper room (SE 2000 Series) whose only illumination came from the background luminance of the video monitor screen. During each session, participants were seated with their forehead and chin placed comfortably against a head spot (University of Houston Optometry). The visual contrast detection task was a 2AFC task where participants determined which of two locations (left and right of the fixation dot) contained the Gabor signal undergoing counterphase flickering. Each session began with instructions displayed on the screen and started after a key press. Figure 4.1A shows the display sequence of a typical trial. A trial began with a fixation period, which lasted between 500-700ms and was marked by centrally located white dot. After the fixation period, two noise images having the same size as the target stimulus were displayed at 5 degrees on each side of the fixation dot. The RMS contrast level of these noise images was linearly ramped up from zero to the maximum contrast level for a duration of ~94ms, remained at the maximum contrast for ~ 562ms and finally, ramped down to zero for a duration of ~94ms. In total, the noise images were presented for 750ms. The Gabor signal was presented 140ms following the onset of the noise images and in one of the two spatial locations for a duration of 470ms. Similar to the noise images, the contrast of the Gabor signal was ramped up and down linearly over ~94ms. The presentation of the Gabor signal in the middle of the noise period divided the noise period into leading and trailing periods. After the trailing noise period, participants were required to provide their response on which spatial location the Gabor signal appeared using “1” for left or “2” for “right” on the keyboard. Following the response, the next trial was initiated after 300ms. Unlike the pre- and post-training sessions, the training sessions include feedback on accuracy of response. The feedback was either the word “correct” or “wrong” and was displayed for a duration of 200ms following some duration after response. The audiovisual contrast detection task which was performed by the AVT group during the training sessions was similar to the above-described visual contrast detection differing only in the fact that the Gabor signal was always concurrently presented with the auditory white noise for the same duration. In addition, the auditory stimulus was ramped up and down linearly for ~94ms, similar to the visual stimulus. Each session was divided into 6 blocks with a mandatory break of 2 minutes in between them.

Contrast detection thresholds were estimated at two performance criterion levels for each of the seven external noise levels using an interleaved adaptive staircase procedure. The two performance criterion levels were based on two stop rules, a two-down-one-up or 2/1 (i.e., a decrease in signal contrast after two successive correct responses and an increase after every incorrect response) which approximates 70.7% correct performance level (d’ = 1.089) and a three-down-one-up or 3/1 which approximates 79.3% correct performance level (d’ = 1.634). Each staircase had two step sizes dictating how contrast changed after each trial in the staircase. The first step size was 0.5 log units which applied to trials before four reversals were attained and then, 0.1 step size that applied till the end of the trial. While the 3/1 staircases ended after 100 trials, the 2/1 staircase ended after 80 trials. In total, each session consisted of 14 interleaved staircases i.e., 2 threshold levels and 7 external noise levels and 1260 trials. In every new session, the staircase for a particular threshold level and external noise level started with the averaged contrast of the last 40% reversal points of the previous session.

4.3.5 Data Analysis

For each staircase, contrast thresholds were determined by averaging the contrast values for the last 40% of the total number of reversals. To ensure more reliable estimates of contrast thresholds, we averaged the determined contrast thresholds between two successive sessions for each threshold level and external noise level. This strategy meant that the 2 sessions of the pre-training phase were combined, and the 2 sessions of the post-training phase were also combined. From sessions 3 to 8 that comprise the training phase, we combined the data as follows; 3/4, 5/6 and 7/8. All statistical analyses were carried out in JASP software version 0.11 (JASP Team, 2018). To determine the effects of group, threshold level, external noise level and sessions on contrast thresholds, we used multi-factor mixed analysis of variance (ANOVA) with Greenhouse-Geisser correction of sphericity. One factor and multifactor ANOVA as well as pairwise t-tests with Holm-Sidak correction were used for post-hoc analyses. All statistical analyses were two-tailed and significance level, alpha, was set to 0.05.

4.3.6 Modeling

4.3.6.1 Overview of the Perceptual Template Model (PTM)

Across a variety of perceptual tasks, humans are not able to perform beyond a certain limit normally referred to as sensory or perceptual thresholds. This perceptual inefficiency arises from various noise sources such as intrinsic stimulus variability, receptor sampling errors, nonlinear signal transduction, neural response variability and neural transmission related loss of information (Ahumada and Watson, 1985; Barlow, 1956; Burgess, Wagner, Jennings, and Barlow, 1981; Nagaraja, 1964; Denis Guillermo Pelli, 1981). At a system level, the human perceptual system can be expressed as a perfect, noise free system having a separate, equivalent internal noise. In the past, psychologists have quantified the amount of internal noise in perceptual systems using the equivalent input noise method adopted from electrical engineers (Lu and Dosher, 1999). In this method, the dependence of threshold —signal stimulus energy required for an observer to maintain a given performance level— on the amount of external noise is determined by systematically adding increasing amounts of external noise to the signal stimulus (Denis Guillermo Pelli, 1981). The threshold vs external noise contrast (TVC) function has three regions. First, the region where internal noise (Na) is greater than external noise (N_{ext}) and internal noise is the limiting factor to perceptual performance; the region where N_{ext} is equal to Na; and thirdly, the region where N_{ext} is greater than Na, and N_{ext} limits perceptual performance (Lu and Dosher, 1999). It should be noted that this principle assumes that the external noise being applied is akin to the human visual system's internal noise in terms of characteristics (Baldwin, Baker, and Hess, 2016).

The threshold vs external noise contrast (TVC) function has been used widely to identify and quantity the sources of perceptual inefficiencies during a perceptual task in human observer models. The human observer models are models inspired by properties of the visual system such as internal noise, nonlinearity, gain control etc. (Dosher and Lu, 2017; Lu and Dosher, 1999; Lu and Dosher, 2004). They systematize the behavior of perceptual observers in different testing circumstances by characterizing signal and noise processing in perception. In its simplest form, the human observer can be modeled as a noisy linear amplifier or LAM, which is perfect linear amplification plus additive noise. The LAM model has three stages, which are 1) a noise-free linear amplifier, 2) an independent, internal additive noise and 3) a decision stage (Lu and Dosher, 1999; Lu and Dosher, 2004).

Although LAM has been widely used in perceptual studies (D’Zmura and Knoblauch, 1998; Reisbeck and Gegenfurtner, 1997), its ability to explain perceptual performances across different tasks and observers has been questionable. One of the deficits of the LAM model is that it systematically mispredicts the relationship between thresholds at different performance levels because it assumes that the ratio between two thresholds at each external noise contrast is equal to the ratio of the corresponding d′ values, independent of the observer and the particular external noise contrast (Lu and Dosher, 1999; Lu and Dosher, 2004).

To provide to better model, Dosher and colleagues developed the perceptual template model (PTM) by adding two components to the LAM which are nonlinear transducer function, (J. M. Foley and Legge, 1981; Nachmias and Sansbury, 1974) and multiplicative noise (Heeger, 1992; Legge and Foley, 1980; Lu and Sperling, 1996; Ohzawa, Sclar, and Freeman, 1982; Watson and Solomon, 1997). In implementation, the PTM consists of 5 components: 1) a perceptual template (or a processor) that has a contrast gain β to the signal (ii) a nonlinear transducer function that raises both the signal and external noise to the γth power, (iii) a Gaussian-distributed internal multiplicative noise term whose mean is 0 and SD is proportional to (Nm ×) the total energy in the stimulus after the nonlinear transformation, (iv) a Gaussian-distributed additive internal noise term whose mean is 0 and SD (Na) is independent of the stimulus energy, and (v) a decision process that is based on signal to noise ratio. In the PTM, accuracy of perceptual task performance is indexed by d′ (Bejjanki et al., 2014; Lu and Dosher, 1999, 2008):

$$d' = \frac{(\beta c)^{\gamma}}{\sqrt{N_{ext}^{2\gamma} + N_m^2\left[N_{ext}^{2\gamma} + (\beta c)^{2\gamma}\right] + N_a^2}}$$

For a given performance level, d’, the equation above can be rearranged to solve for threshold contrast, c_τ as a function of external noise in the log form:

$$\log(c_\tau) = \frac{1}{2\gamma}\log\left((1 + N_m^2)N_{ext}^{2\gamma} + N_a^2\right)$$

$$= -\frac{1}{2\gamma}\log\left(\frac{1}{d'^2} - N_m^2\right) - \log(\beta)$$

4.3.6.2 Mechanisms of Perceptual Learning under the PTM

Although Lu and Dosher (Lu and Dosher, 1999) highlighted three distinct mechanisms that could underlie perceptual learning within the PTM, a look to the parameters of the PTM indicates that four distinct mechanisms can be described. These mechanisms are stimulus enhancement, external noise exclusion, reduction in multiplicative noise and changes in nonlinear transducer properties of the system (Figure 4.1B).

First, stimulus enhancement improves performance by reducing internal additive noise. Mathematically, this is equivalent to increasing the gain of the output of the signal-relevant perceptual template (i.e., multiplying the contrast of the input stimulus by a factor greater than 1) (Lu and Dosher, 1999). The signature for this mechanism is threshold reduction (performance improvement) in low to zero external noise regions where internal noise is the limiting process. This mechanism does not affect performance or threshold in high external noise region because increasing the gain of the signal equally increases the gain on the external noise. Second, external noise exclusion improves performance by retuning the perceptual template. Because the perceptual template deals with the specific processes involved in the task, retuning the perceptual template focuses perceptual analysis on the appropriate characteristics of the signal stimulus such as the temporal duration, spatial region, or spatial frequency (Lu and Dosher, 1999). The behavioral signature for this mechanism is performance improvements in the region of high external noise. Unlike stimulus enhancement and external noise exclusion mechanisms, the remaining two mechanisms, which are internal multiplicative noise reduction and changes in nonlinear transduction properties, depend on threshold or performance criterion levels. Internal multiplicative noise reduction reduces a source of noise that is proportional to the total energy in the input stimulus. The mechanism produces a behavioral signature of performance improvements in both high and low levels of external noise. However, because higher contrast signals are required to achieve higher level of threshold performance (e.g., d’ of 1.6 instead of 1.0), reduction in multiplicative noise decreases threshold for the high-performance level (or more stringent) criteria compared to low-performance level (or less stringent) criteria. Lastly, changes in nonlinear transducer properties produces performance enhancements that depend on both external noise level and performance level criterion. Typically, the nonlinearity function accelerates at low contrast levels and saturates at high contrast levels (Legge and Foley, 1980). Therefore, within the PTM model, reduction in the strength of the nonlinear transducer decreases threshold contrasts at low compared to high external noise levels and at low compared to high performance levels. Moreover, at high external noise levels and high-performance levels, reduction in transducer strength can result in increase in thresholds or worsening of perceptual performance. It is worth mentioning that while the stimulus enhancement and external noise exclusion mechanisms have been demonstrated in several studies (Dosher and Lu, 2017; Dosher and Lu, 1999; Lu and Dosher, 1999; Lu and Dosher, 2004), the nonlinear mechanisms that are reduction in multiplicative noise and changes in nonlinear transducer properties are yet to be reported.

4.3.6.3 Fitting the PTM

In the PTM, performance improvements across sessions (i.e., pooled across successive session as explained in data analysis section) can be modeled by multiplying Na by a factor Aa(t) < 1.0 for better stimulus enhancement, N_{ext} by a factor Af(t) <1.0 for better external noise exclusion, Nm by factor Am(t) <1.0 for reduced internal multiplicative noise, and γ by factor Ag(t) <1.0 for reduced strength of the nonlinear transducer, where t represents a particular session (Bejjanki et al., 2014; Lu and Dosher, 1999; Lu and Dosher, 2004). When all four mechanisms are combined, the contrast threshold vs. external noise function for a PTM can be expressed as:

$$\log(c_t) = \frac{1}{2\gamma}\left\{log\left[(1+ A_m^2 N_m^2)A_f^{2A_g^2\gamma} N_{ext}^{2A_g^2\gamma} + A_a^2 N_a^2\right] - log\left[\frac{1}{d'^2} - A_m^2 N_m^2\right]\right\}$$
$$- \log(\beta)$$

where c_t denotes the predicted contrast threshold, N_{ext} is the SD of external noises, d′ is the perceptual sensitivity of the observer, corresponding to the thresholds in the two-alternative forced choice task.

To derive TvC curves and parameter values for our data, we fitted 16 models ranging from the null model (which is, no change in any of the four mechanisms) to the full model (which is, changes in all four mechanisms) to 1000 bootstrapped samples (drawn with replacement) determined independently for the VOT group and AVT group. Thus, each bootstrapped sample had 16 best-fitted models, each determined by minimizing the least square difference between the log of the measured threshold contrasts and the log of the model-predicted threshold contrasts. The r^2 statistic was chosen as the measure of goodness of fit. For each bootstrapped sample, we determined the best model (which represents the most plausible underlying mechanism(s)) using the F test. The F test allowed us to select the smallest model in terms of parameter size, which is significantly different from the null model and comparable to the full model. Below is the equation for the F test:

$$F(df_1, df_2) = \frac{\left[(r^2_{full} - r^2_{reduced})/df_1\right]}{\left[(1 - r^2_{full})/df_2\right]},$$

Where $df_1 = k_{full} - k_{reduced}$, and $df_2 = N - k_{full}$. N is the number of predicted data points (i.e., 7 external noise levels x two performance level) and k is the number of parameters in each model.

When two or more models of the same size passed this test, we chose the model with the highest r2 to be the best model. This approach resulted in 1000 best models for each group based on the different combinations of the perceptual learning mechanisms. The underlying mechanism(s) for each group was based on the model with the highest frequency.

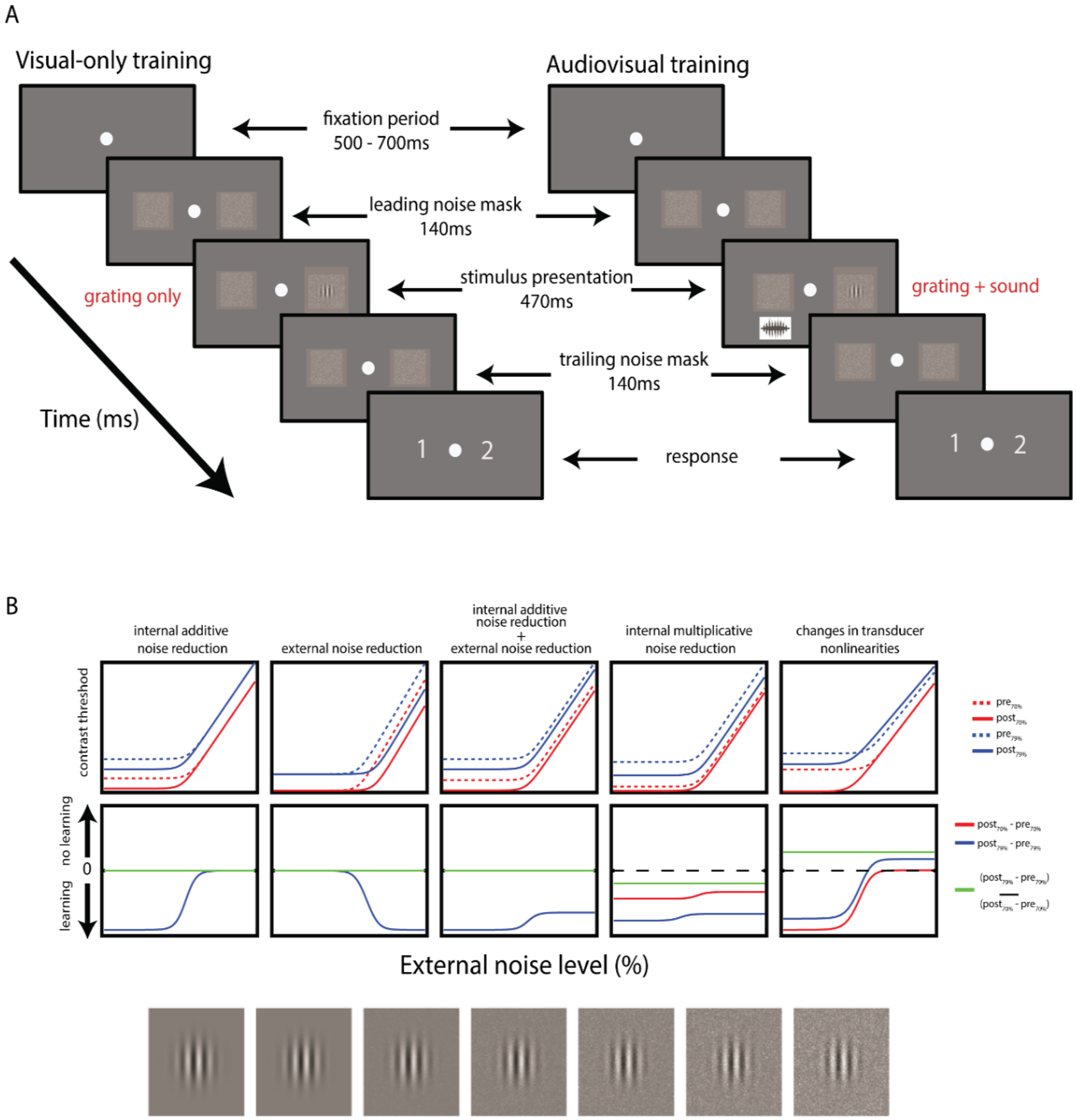

Figure 4. 1 Training paradigms and predictions from PTM.

(A) Participants were randomly assigned to either a visual-only (VOT) or an audiovisual (AVT) training paradigm. Each participant completed 10 sessions which included 2 pre-training, 6 training and 2 post-training sessions. The pre-and post-training sessions resembles with the visual-only training (without feedback), while for the training session, the VOT group performed the visual-only training with feedback while the AVT group performed the audiovisual training with feedback. A typical trial of the visual-only training paradigm started with a fixation period, followed by a leading noise mask, stimulus presentation which was a Gabor patch appearing in one of the two locations, trailing noise mask and response. After that, the feedback was provided on the accuracy of the response and the next trail began about 300ms later. The audiovisual training paradigm was similar to the visual only training except for the inclusion of sound, simultaneously presented with the visual stimulus. (B) Illustrations of typical threshold vs contrast functions and predictions from different mechanisms. The PTM can predict four distinct mechanisms namely internal additive noise reduction, external noise reduction, internal multiplicative noise reduction and changes in transducer nonlinearities.

4.4 Results

We determined the performance of two groups of participants on a 2AFC visual contrast detection task before and after 6 sessions of training. While one group was trained on visual only stimulation (i.e., Gabor signal) with feedback (VOT group), the other group was trained on temporally congruent visual and auditory (i.e., auditory white noise) stimulation with feedback (AVT group). Using an adaptive staircase procedure, contrast thresholds for the detection tasks for each session were determined for two performance criterion levels (70.7% and 79.3%) and seven external noise levels. Pooling over every two sessions, we derived contrast thresholds for the pre-training phase (sessions 1/2), training phase (sessions 3/4, 5/6, 7/8) and post-training phase (sessions 9/10). The mean number of days used to complete the task was comparable between the two groups (VOT: 16.8 ± 1.6 days, AVT: 16.6 ± 2.2 days, t(18) = -0.23, p = 0.821).

4.4 .1 VOT and AVT Participants had Comparable Thresholds at Baseline

Before the training phase, a 2*2*7 ANOVA was performed with group (VOT/AVT) as the between-subject factor; external noise level and performance level as the within-subject factor; and signal contrast thresholds of the pre-training phase as the dependent variable, to determine whether the groups were comparable at baseline. Our results showed no significant main effect of group (F(1, 18) = 0.485, p = 0.495, η^2 = 0.026) indicating that any training or post-training group effects cannot be explained by differences at baseline (Figure 4.2). Unsurprisingly, we found a significant main effect of performance level (F(1, 18) = 140.262, p = 6.235e-10, η^2 = 0.032) indicating lower contrast thresholds for the less stringent performance level (i.e., 70.7%) and higher thresholds for the more stringent performance level (i.e., 79.3%). Moreover, a significant main effect of external noise level

($F(1.7, 29.7) = 99.961$, $p = 4.2e\text{-}13$, $\eta^2 = 0.655$) was found indicating lower contrast thresholds for low external noise levels and higher contrast thresholds for higher external noise levels. The finding of significant main effects of performance level and external noise level are consistent with previous studies (Dosher and Lu, 1998; Lu and Dosher, 1999; Lu and Dosher, 2004) and confirm that our experimental procedures and manipulations were appropriate and plausible.

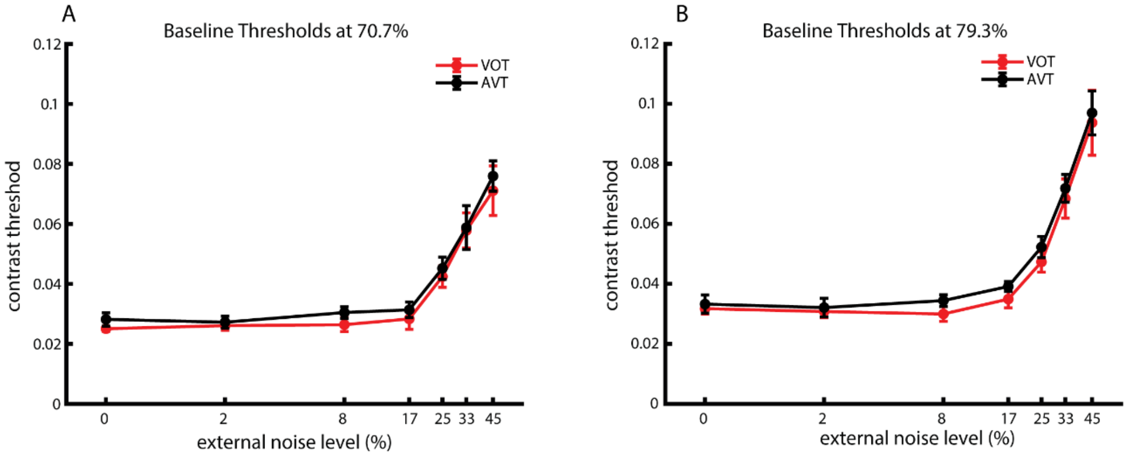

Figure 4. 2 Baseline Contrast Thresholds.

Pre-training contrast thresholds for VOT and AVT groups across all external noise levels for different performance or threshold levels, 70.7% (A) and 79.3% (B). Contrast thresholds are presented in decimal notation while external noise levels are presented in percentages (%).

4.4.2 Adding a Task Irrelevant but Temporally Correlated Sound Improves Contrast Detection Performance during Training

To determine whether audiovisual training (i.e., the presence of auditory white noise) affected participants' performance differently from visual only training, we conducted a 2*4*2*7 ANOVA with group (VOT/AVT) as the between-subject factor; sessions (1/2, 3/4, 5/6, 7/8), performance level (70.7%/79.3%) and external noise level as the within-subject factors; and contrast threshold as the dependent variable (Figure 4.3A). Interestingly, across all participants, we found a significant interaction between group and sessions ($F(2.6, 47) = 4.449$, $p = 0.01$, $\eta^2 = 0.022$), indicating that training-improved performance across the sessions differed between the two groups. Furthermore, we found a significant interaction between session and noise ($F(3.3, 59.8) = 7.798$, $p = 1.071e-4$, $\eta^2 = 0.017$), indicating that across both groups, performance improvements over baseline differed based on the external noise level. However, the 3-way interaction between session, noise and group was not significant ($F(3.3, 59.8) = 0.636$, $p = 0.610$, $\eta^2 = 0.001$), suggesting that the dependence of performance improvements on external noise level was not different between the groups. While not statistically significant, we did find a near significant trend of interaction between session and performance level ($F(2.8, 49.5) = 2.571$, $p = 0.069$, $\eta^2 = 3.05e-4$) indicating a possible dependence of performance improvements on performance level. Again, this near-significant interaction was found not to different between the two groups ($F(2.8, 49.5) = 1.604$, $p = 0.203$, $\eta^2 = 1.902e-4$). Next, we conducted post-hoc analyses to probe further into the significant session vs group interaction. To begin, we investigated the training sessions (i.e., 1/2, 3/4, 5/6, 7/8) where performance improvements differed between the groups. To do this, we conducted 6 separate 2*2 ANOVAs focusing on the interaction between group and session. The six ANOVAs resulted from all pairwise combinations of the three training sessions and the pre-training session (i.e., 1/2, 3/4, 5/6, 7/8). Interestingly, our results showed significant effect for only 1/2-7/8 interaction (after correcting for multiple comparisons with Holm-Sidak correction) indicating that performance during training when sound was included for the AVT group significantly improved for the AVT group compared to the VOT group during the later stages of training (**1/2-7/8**: F(1, 18) = 10.768, p = 0.004, η^2 = 0.033, Holm-Sidak alpha = 0.008).

Furthermore, we conducted analysis separately for each group to understand how performance changed within the group. For each group, we conducted a 2*4*7 ANOVA with sessions (1/2, 3/4, 5/6, 7/8), performance level (70.7%/79.3%) and external noise level as the within subject factors, and contrast threshold as the dependent variable. Findings from these analyses revealed three major differences between training within the two groups. First, our results showed revealed a significant main effect of session for both groups indicating that training indeed improved detection performance for participants in both groups (**VOT**: F(3, 27) = 5.288, p = 0.005, η^2 = 0.009; **AVT**: F(3, 27) = 26.180, p = 3.78e-8, η^2 = 0.053; Figure 4.3A). For the 4 sessions, we conducted 6 pairwise t-tests to determine which sessions showed performance improvement. Interestingly, while training improved performance immediately for the AVT group in the 3/4 session (t(9) = 8.078, p = 0.00002, d = 1.085, Holm-Sidak alpha = 0.0083), for the VOT, significant improvement in performance over baseline was first observed in session 5/6 (t(9) = 3.43, p = 0.0075, d = 1.085, Holm-Sidak alpha = 0.01; Figure 4.3A). Furthermore, once the first significant improvement in performance over baseline was observed, there was no further significant improvement in performance over baseline during subsequent training sessions (Figure

4.3A). However, it is important to note that while performance in the VOT group seemed to plateau getting to the end of the training sessions, there was still a gradual decrease in contrast thresholds (or improvement in performance) for the AVT group towards the end of the training, albeit not significant (Figure 4.3A).

Second, our results revealed a significant interaction between session and external noise level for both groups indicating that improvement in performance over sessions differed between external noise levels (**VOT**: F(18, 162) = 2.329, p = 0.003, $\eta^2 = 0.012$; **AVT**: F(18, 162) = 7.499, p = 6.956e-14, $\eta^2 = 0.027$). To dissect the interaction effects further, we reduced the levels of external noise from 7 to 2 by pooling contrast thresholds from the first 4 external noise levels together into one low external noise level and pooling the remaining 3 external noise levels together into one high external noise level. For each group, we then conducted two tests. First, we investigated the main effect of noise (high vs low) on contrast thresholds and discovered that for both groups, training improved detection performance significantly more for high compared with low external noise levels (**VOT**: F(1, 9) = 89.54, p = 6e-6, $\eta^2 = 0.627$; **AVT**: F(1, 9) = 206.53, p = 1.635e-7, $\eta^2 = 0.634$; Figure 4.3B). Second, we wanted to investigate whether training improved performance in each external noise level for both groups by conducting one-away ANOVA across sessions. Interestingly, while for the AVT group, training improved performance significantly in both low (**AVT**: F(3) = 13.48, p = 1.449e-5) and high external noise levels (**AVT**: F(3) = 24.68, p = 6.76e-8), for the VOT group, training only improved performance in high (**VOT**: F(3) = 5.442, p = 0.005) but not low external noise levels (**VOT**: F(3) = 1.771, p = 0.1763) (Figure 4.3B), indicating differential impact of the training paradigms on improving performance across external noise levels.

Lastly, there was a significant interaction between sessions and performance level (70.7% / 79.3%) for the AVT group (**AVT**: F(3, 27) = 3.393, p = 0.032, η^2 = 8.318e-4) but not the VOT group (**VOT**: F(3, 27) = 0.919, p = 0.445, η^2 = 2.121e-4; Figure 4.3C) indicating that for AVT group but not the VOT group, performance improvement during training depended on the performance criterion level. Because significance was observed only for the AVT group, we conducted further post-hoc analyses to determine which sessions were the difference between low and high-performance level significant. We conducted all six possible two-way interactions between session (1/2, 3/4, 5/6, 7/8) and performance level (70.7%, 79.3%). Interestingly, our results revealed that while the interaction for sessions 1/2-3/4 was not significant (**AVT**: F(1, 9) = 5.432, p = 0.045, η^2 = 0.005, Holm-Sidak alpha = 0.0125) indicating that the initial training improvement was not different between performance levels, the interaction for sessions 3/4-5/6 was significant (**AVT**: F(1, 9) = 25.726, p = 0.0007, η^2 = 0.0055, Holm-Sidak alpha = 0.0083) as well as the interaction for sessions 3/4-7/8 (**AVT**: F(1, 9) = 13.984 , p = 0.0046, η^2 = 0.0056, Holm-Sidak alpha = 0.01). These findings indicate that after the initial improvement in performance observed for both performance levels at session 3/4, training improved performance only for low performance level (i.e., the less stringent criteria) but not high-performance level. Since signal contrast levels for the less stringent criteria are lower and thus, more ambiguous, the continual improvement with sound for this level compared with the more stringent one highlights the presence of an important principle in multisensory integration which is inverse effectiveness, that is, more gains in perceptual performance are likely to be observed when the information in the task relevant modality is weakly effective (Stein and Meredith, 1993).

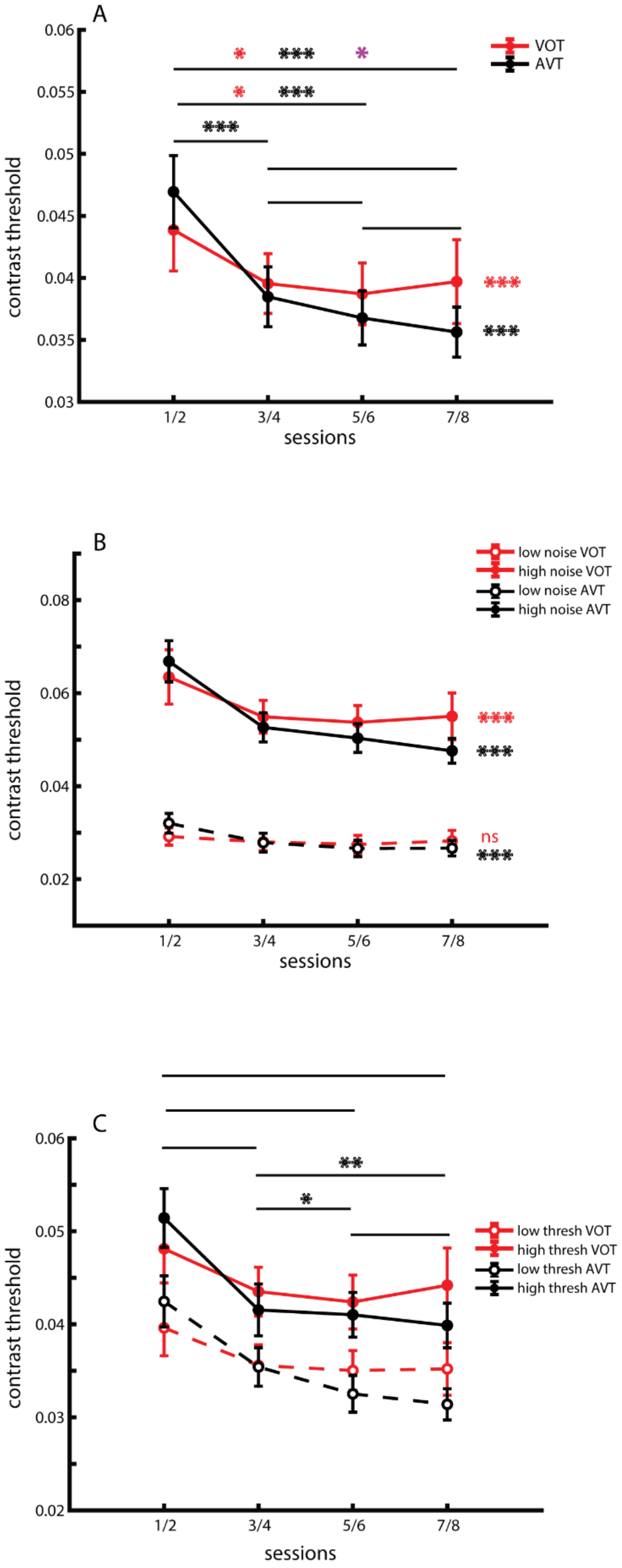
A
0.06
0.055
0.05
0.045
0.04
0.035
0.03
contrast threshold
1/2
3/4
5/6
7/8
sessions
VOT
AVT
* *** *
* ***
***
***
***
B
0.08
0.06
0.04
0.02
contrast threshold
1/2
3/4
5/6
7/8
sessions
low noise VOT
high noise VOT
low noise AVT
high noise AVT
***
***
ns
***
C
0.06
0.05
0.04
0.03
0.02
contrast threshold
1/2
3/4
5/6
7/8
sessions
**
*
low thresh VOT
high thresh VOT
low thresh AVT
high thresh AVT

Figure 4. 3 Training results.

(A) Mean contrast thresholds pooled across both external noise levels and performance levels for the groups, VOT (in red) and AVT (in black) and sessions (1/2, 3/4, 5/6, 7/8). (B) Mean contrast thresholds pooled across performance levels for the groups, VOT (in red) and AVT (in black), external noise level (low: [first 4 out 7 contrast levels, white-filled circles, dashed lines], high: [remaining 3 contrast levels, filled circles, solid lines]) and sessions (1/2, 3/4, 5/6, 7/8). (C) Mean contrast thresholds pooled across external noise levels for the groups, VOT (in red) and AVT (in black); performance level (low: [white-filled circles, dashed lines], high: [filled circles, solid lines]) and sessions (1/2, 3/4, 5/6, 7/8). Asterisks on the right side of the line plots represent statistical results of post-hoc one-way ANOVA for that particular factor. Asterisks on top of the plots represent statistical results of post-hoc pairwise t-tests between sessions for each group (VOT in red, AVT in black) and the interaction effects for post-hoc 2*2 ANOVA with group and session as factors (purple). * $p<0.5$, **$p<0.01$, ***$p<0.001$

4.4.3 Audiovisual Training Reduces Contrast Threshold More for the AVT Group after Training and in the Absence of Sound

Because the AVT paradigm involved the use of both auditory white noise and Gabor signal while the VOT paradigm involved only the Gabor signal, we investigated training-induced improvement in performance bereft of the accompanying sound for the AVT group. To do this, we conducted a 2*2*2*7 ANOVA with group (VOT/AVT) as the between-subject factor; sessions (1/2, 9/10), performance level (70.7%/79.3%) and external noise level as the within-subject factors; and contrast threshold as the dependent variable. Our analysis revealed two interesting results. First, there was a significant interaction between session and group (F(1, 18) = 5.693 , p = 0.028, $\eta^2 = 0.005$; Figure 4.4A) indicating that participants in the AVT group had significantly more improved detection performance than those in the VOT group even in the absence of sound after the training. Second, we found a significant interaction between session and external noise level (F (3.1, 55.1) = 23.578, p = 4.802e-10, $\eta^2 = 0.027$; Figure 4.4A) indicating post-training improved performance differed across different external noise levels.

To probe these analyses further, we conducted separate 2*4*7 ANOVAs for each group with sessions (1/2, 9/10), performance level (70.7%/79.3%) and external noise level as the within subject factors, and contrast threshold as the dependent variable. For each group, there was a significant main effect of session indicating improved performance after training (**VOT**: F(1, 9) = 5.667, p = 0.041, $\eta^2 = 0.01$; **AVT**: F(1, 9) = 38.401, p = 1.55e-4, $\eta^2 = 0.066$). Similar to the findings from the training data, we discovered significant interactions between session and external noise level for both group (**VOT**: F(2.7, 24.6) = 8.085, p = 8.483e-4, $\eta^2 = 0.018$; **AVT**: F(2.7, 24.1) = 17.633, p = 4.832e-6, $\eta^2 = 0.04$). To understand this further, we reduced the number of external noise levels from 7 to 2 as described earlier and conducted two analyses. First, a 2*2 ANOVA with session and external noise level (low vs high) as factors showed a significant main effect of external noise for both groups indicating that after training, performance improved better in high compared to low external noise level (**VOT**: F(1, 9) = 68.914, p = 1.6e-5, η^2 = 0.55; **AVT**: F(1, 9) = 143.444, p = 7.828e-7, η^2 = 0.615; Figure 4.4C). However, when post-hoc simple main effects of session (i.e., pre and post) were conducted each external noise level for each group, we discovered that while the AVT group, after training performance improved significantly in both low (**AVT**: F(1) = 15.363, p = 0.0035) and high external noise levels (**AVT**: F(1) = 46.367, p = 7.822e-5), for the VOT group, only in high (**VOT**: F(1) = 11, p = 0.009) but not low external noise level (**VOT**: F(1) = 4.3e-5, p = 0.995) was improvement in performance observed. Unlike the training results, there was no significant interaction between sessions and performance levels for both groups (**VOT**: F(1, 9) = 0.232, p = 0.641, η^2 = 7.6e-5; **AVT**: F(1, 9) = 0.376, p = 0.555, η^2 = 8.136e-5; Figure 4.4D). Because the AVT group was trained with sound for all six training sessions, it is possible that removing the sound in the post-training could have impacted detection performance. However, a 2*2*2*7 ANOVA with group (VOT/AVT) as the between-subject factor; sessions (7/8, 9/10), performance level (70.7%/79.3%) and external noise level as the within-subject factors; and contrast threshold as the dependent variable revealed no significant interactions including the interaction between session and group (F(1, 18) = 0.33, p = 0.573, η^2 = 1.919e-4; Figure 4.4B), indicating that removal of sound did not affect performance of the AVT group. This further implies that mechanisms that may mediate performance changes during the training session may not differ from those during the post-training session.

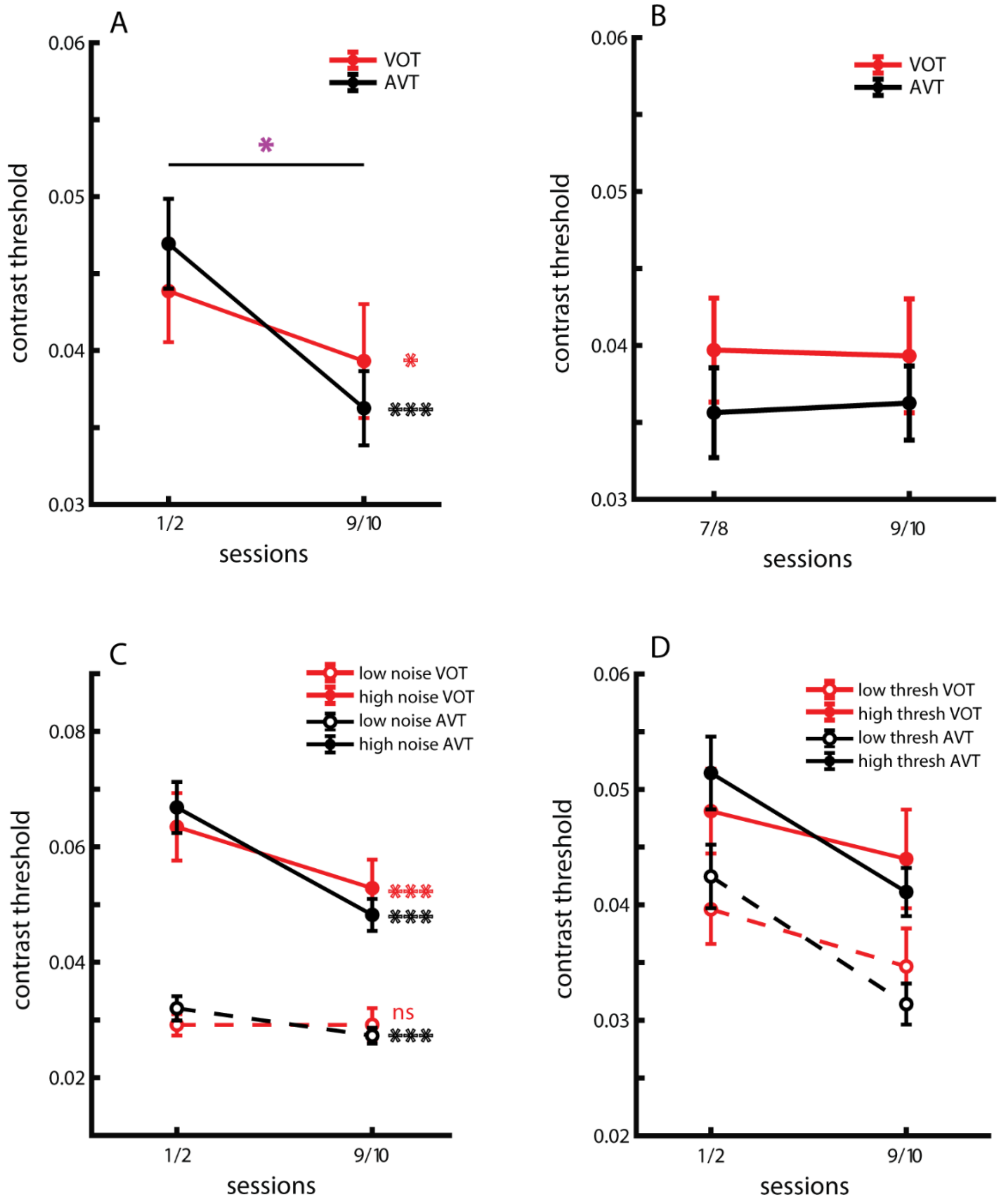

Figure 4. 4 Post training results.

(A) Mean contrast thresholds pooled across both external noise levels and performance levels for the groups, VOT (in red) and AVT (in black) and sessions (1/2, 9/10). (B) Mean contrast thresholds pooled across both external noise levels and performance levels for the groups, VOT (in red) and AVT (in black) and sessions (7/8, 9/10). (C) Mean contrast thresholds pooled across performance levels for the groups, VOT (in red) and AVT (in black), external noise level (low: [first 4 out 7 contrast levels, white-filled circles, dashed lines], high: [remaining 3 contrast levels, filled circles, solid lines]) and sessions (1/2, 9/10). (D) Mean contrast thresholds pooled across external noise levels for the groups, VOT (in red) and AVT (in black); performance level (low: [white-filled circles, dashed lines], high: [filled circles, solid lines]) and sessions (1/2, 9/10). Asterisks on the right side of the line plots represent statistical results of post-hoc t-tests for that particular factor. Asterisks on top of the plots represent statistical results of interaction effects for post-hoc 2*2 ANOVA with group and session as factors (purple). * $p<0.5$, **$p<0.01$, ***$p<0.001$

4.4.4 Modeling Results

We employed the PTM to infer mechanisms underlying the observed superiority of AVT compared to VOT training by fitting threshold vs external noise functions to our data separately for each group. Because training and post-training contrast detection thresholds did not differ within and between the groups (see Figure 4.4B), we fitted the PTM to all 5 sessions (i.e., 1/2, 3/4, 5/6, 7/8, 9/10) together to identify perceptual learning mechanisms across the entire pre-, during and post-training phases. The PTM framework allows us to distinguish between four mechanisms, which are stimulus enhancement (i.e., reduction in internal noise), external noise exclusion, reduction in multiplicative noise, and changes in nonlinear transducer properties of the system. Based on these four mechanisms, we fitted 16 models ranging from no change in all four mechanisms to a change in all mechanisms to 1000 bootstrapped samples (i.e., sampling with replacement) determined independently for each group. For each bootstrapped sample, we determined the best fitting model of the 16 possible models using the F test and eventually, determined the best overall model across the 1000 bootstrapped samples as the model with the highest frequency of selection. Our analysis revealed that while improved external noise exclusion was the best-fit model for the VOT group (Figure 4.5A), for the AVT, the best-fit model was a combination of improved external noise exclusion and reduced strength of the nonlinear transduction properties of the perceptual system (Figure 4.5B). Interestingly, for the external noise reduction mechanisms that was common between the groups, the amount of noise reduction for the AVT group was consistently higher than that of the VOT group, confirming the superiority of AVT paradigm observed in the behavioral analysis (Figure 4.5C, D).

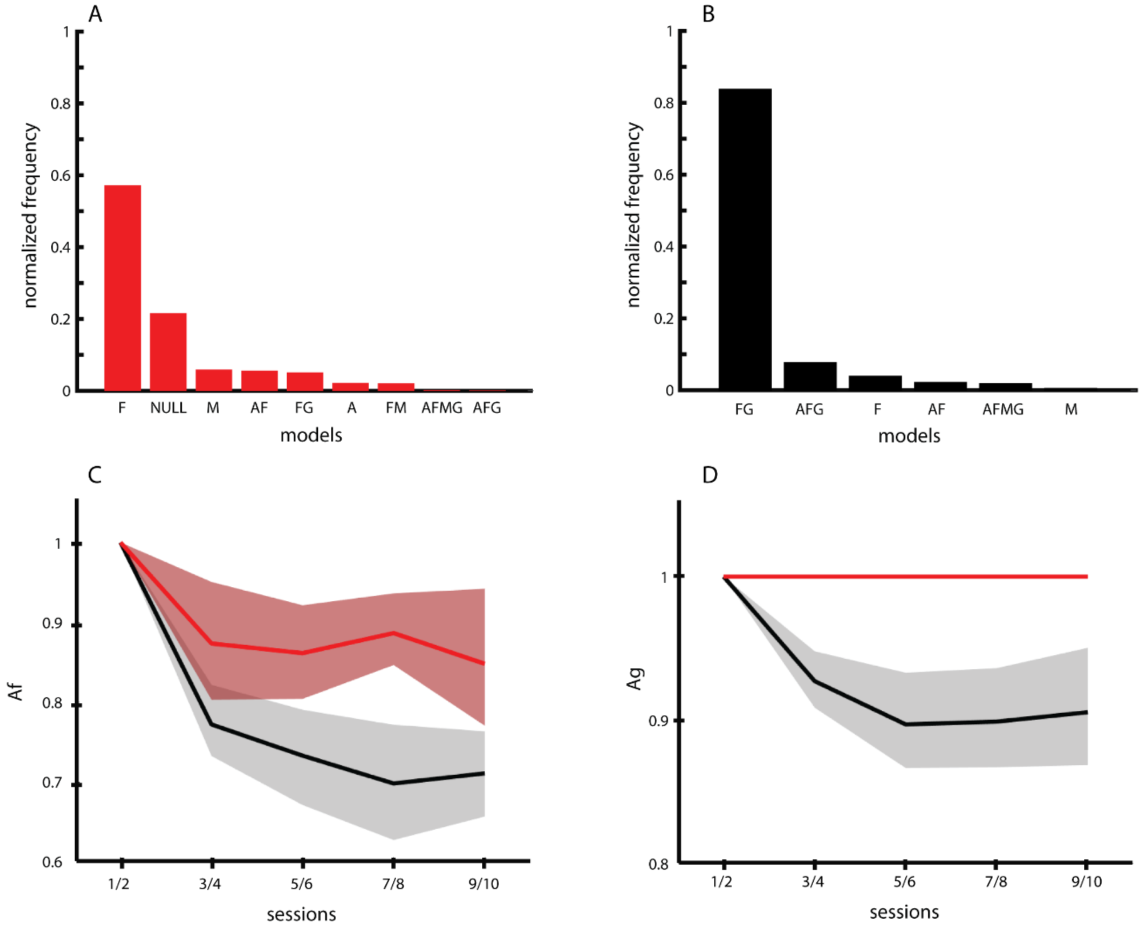

Figure 4. 5 PTM modelling results

(A-B) Normalized frequencies for PTM models with above zero frequency for VOT (A) and AVT (B) groups. F, external noise reduction; M, internal multiplicative noise reduction; A, internal additive noise reduction; G, change in transducer nonlinearities. Mean Af (C) and Ag(C) scores with 95% confidence interval (shaded area) across sessions (1/2, 3/4, 5/6, 7/8, 9/10) for VOT (red) and AVT (black) groups.

4.5 Discussion

Training on a visual perceptual task with additional information from another sensory modality (e.g., audition) produces more effective and efficient learning outcomes than using only the visual information (Kim, Seitz, and Shams, 2008; Seitz et al., 2006; Shams and Seitz, 2008; Zilber et al., 2014). However, it is still unclear whether this benefit can apply to visual features that may not be ecologically associated such as motion and speech stimuli. In this study, we asked whether learning to detect a visual Gabor signal undergoing counterphase flicker would benefit from hearing a task irrelevant but temporally correlated sound. We determined pre and post training visual contrast detection performance of two groups of participants trained either on a visual-only or on an audiovisual paradigm. To provide a mechanistic understanding of observed training differences, we employed the PTM, an observer model that can distinguish four mechanisms of perceptual learning using signal to noise detection theory (Dosher and Lu, 2017; Lu and Dosher, 1999; Lu and Dosher, 2004; Lu and Dosher, 2009). Our findings show that incorporating task irrelevant sound in a visual contrast detection training improved visual detection performance not only in the training phase when sound was present, but also after training when the sound was absent. In addition, audiovisual training showed significant improvement in detection performance in conditions where the visual stimulus was weakly effective highlighting an important principle guiding multisensory interactions and integration known as inverse effectiveness. According to our PTM analysis, audiovisual training reduced perceptual inefficiencies associated with external noise more than visual only training. Further highlighting the superiority of AVT compared to VOT training, we show for the first time that a type of perceptual learning (i.e., audiovisual) can produce changes in nonlinear transduction properties of the perceptual system. Taken together, our results imply that multisensory perceptual learning may engage separate and or additional mechanisms compared to visual-only learning and thus, will provide a powerful new set of rehabilitative tools in the quest to improve visual function in patients with low vision.

Indeed, our findings are consistent with previous studies that demonstrated that a task irrelevant but simultaneously presented sound could enhance the salience or detection of visual targets. For example, a significant enhancement of perceived intensity of LED light stimulus was observed when paired with concurrent broadband auditory stimulus (Stein, London, Wilkinson, and Price, 1996). Moreover, Lippert et al. (2007) demonstrated that an informative sound significantly enhanced detection rates in a contrast detection task. In our study, the presence of the sound provided no task relevant information about the crucial dependent variable on the task, i.e., the location of the Gabor patch. Previous studies have indicated that the sound-induced perceptual enhancements on visual detection may reflect changes in low level multisensory interactions (indexed as changes in sensitivity, d') and or high level cognitive or decisional effects (indexed as response bias). For instance, while Lippert et al. (2007) observed changes in both sensitivity and response bias using different experiment approaches, Noesselt et al. (2010) reported only changes in sensitivity but not response bias. In this study, it is more likely that the increase in detection performance observed during the presentation of sound for the AVT group reflected changes in sensitivity due to the 2AFC nature of the paradigm which permits a bias-free examination of subject's detection performance (Arieh and Marks, 2008; Odgaard, Arieh, and Marks, 2003).

Compared to previous literature, the lack of improvement in low external noise conditions for the VOT group is not surprising and does not suggest a poorly implemented training paradigm. Lu and Dosher (2004) observed performance improvements only in high external noise conditions in a foveal orientation identification task. However, other similar experiments that tested in the peripheral visual field found improvements in both low and high noise (Dosher & Lu, 1998, 1999). Comparing the findings in central and peripheral vision training indicate that perceptual improvement in low external noise is difficult to achieve when the information reliability is high because of the visual pathway involved in performing the task. In general, information reliability and thus, performance in a detection task can be enhanced by increasing the duration of the visual targets (J. Foley and Tyler, 1976; Graham and Kemp, 1938; Kahneman, 1966) as explained by Bloch's Law (Bloch, 1885; Gorea and Tyler, 1986). Therefore, in our study, it is possible that the lack of performance enhancement at low external noise conditions for the VOT group is a result of performance ceiling effects due to the highly effective visual stimulus characterized by its long duration. It is therefore interesting that even under these conditions audiovisual training was able to improve performance both during and after training.

The principle of inverse effectiveness was introduced by Stein and colleagues (Stein & Meredith, 1993) using cell recordings in animals. Their findings suggested that the greatest neural and perceptual gains were observed when the stimuli to be combined were individually weakly effective, that is, they produced weak responses in isolation. Under signal to noise detection paradigms, a weakly effective stimulus can mean either lower signal energy under constant external noise or a higher noise energy relative to constant signal strength. First, our study found that audiovisual training improved detection performance in the higher external noise conditions compared to the visual only training. This finding agrees with studies that examined the effect of applying different levels of background acoustic noise on the magnitude of visually facilitated speech comprehension and intelligibility and reported a monotonic relationship where the greatest multisensory gains were achieved under very high background noise conditions (Grant and Seitz, 2000). However, contrary to these findings, (Ross, Saint-Amour, Leavitt, Javitt, and Foxe, 2007) suggested the maximal multisensory gains might occur within a range of intermediate signal-to-noise ratios. Additionally, Chen et al. (2011) examined the impact of simultaneously presented sound on visual contrast detection and discrimination tasks under different levels of external noise and discovered perceptual enhancements only in intermediate noise conditions. The lack of agreement between Chen et al. (2011) finding and that of this study may be explained by the fact that our study involved first, an audiovisual stimulus pair that shared a more complex temporal correspondence instead of mere simultaneity and second, continued presentation of the audiovisual stimuli across several sessions, allowing for enhanced temporal binding and associative learning. Regardless of external noise, our study showed that audiovisual but not visual only training improved detection sensitivity for low performance level (70.7%) stimuli where the signal contrasts are low compared to high performance level stimuli. In agreement with this finding, Noesselt et al. (2010) demonstrated that co-occurring sound enhanced detection sensitivity of a low but not high intensity Gabor patch. Taken together, these findings demonstrating the principle of inverse effectiveness during the multisensory training indicate the presence of multisensory interactions mediating the observed superiority of audiovisual training.

In the past, multisensory-facilitated unisensory perceptual learning had been demonstrated by several studies using mostly visual motion (Kim et al., 2008; Seitz et al., 2006; Shams and Seitz, 2008; Zilber et al., 2014) and speech stimuli (Lidestam, Moradi, Pettersson, and Ricklefs, 2014; Sheffert and Olson, 2004; Zäske, Mühl, and Schweinberger, 2015) which may be described as "ecologically valid" (Shams et al., 2011). However, when arbitrary stimuli were paired, training did not result in improved detection performance in the absence of the auxiliary modality (Shams et al., 2011; Wozny et al., 2008). This finding raised an important question about the nature of cross-modal features and their associations that can benefit from multisensory facilitated unisensory perceptual learning. Our findings indicate that criteria for achieving this kind of learning may reside in the presence of some spatial-temporal correspondence that can enhance cross-modal binding, which indeed can be observed in ecologically valid features as well. However, whether ecologically valid features may yield better learning outcomes in terms of more gains and stability compared with features with artificial associations require future investigation.

Evidence from prior studies suggests that multisensory training may facilitate unisensory learning through the enhancement of connectivity between unisensory and multisensory brain areas leading to an activation of a wider network during subsequent presentation of unisensory stimuli. For example, a short period of audiovisual training on motion discrimination recruited a wider network of brain areas (including pSTS, mSTS, and AC), which was activated above baseline during the post-training phase when the auditory stimulus was absent (Zilber et al., 2014). In context of target detection, presenting simultaneous visual and auditory targets strengthened the interregional coupling between auditory and visual thalamic structures with their respective sensory specific cortices and more importantly with STS, a known multisensory hub in the brain involved in multisensory temporal processing (Noesselt et al., 2007; Noesselt et al., 2010). These findings corroborate the hypothesis that repeated exposure to audiovisual stimulation during training enhances connectivity or association between the features or brain representations, leading to enhanced processing of the features in subsequent stimulation (Shams et al., 2011).

According to our PTM analysis, audiovisual training reduced perceptual inefficiencies associated with external noise more than visual only training. Indeed, the external noise exclusion mechanism has been reported by several studies using PTM (Dosher and Lu, 1998, 1999; Lu and Dosher, 1999; Lu and Dosher, 2004), indicating the prominent role it plays in how the brain improves perceptual performance during training. Under the PTM, external noise exclusion improves performance by retuning the perceptual template. Given that the perceptual template deals with the specific processes involved in the task, this finding suggests that AVT improved learning outcomes by enhancing the perceptual analysis on the task and stimulus relevant characteristics. Importantly, we show for the first time that a type of perceptual learning (i.e., audiovisual) can produce changes in nonlinear transduction properties of the perceptual system. Under the PTM model, a behavioral signature of a dependence of sensitivity/discriminability on performance level reflects two types of mechanisms namely multiplicative noise reduction and changes in transducer nonlinearities. While multiplicative noise reduction is indexed by more improvement in high vs low performance level, changes in nonlinear transducer properties (as implemented in our model) is indexed by more improvement in low vs high performance levels. In addition, changes in transducer nonlinearities manifests as improvement in low external noise levels and sometimes no improvement or even worsening of performance at high external noise, mimicking internal noise reduction mechanisms (Baldwin et al., 2016). Based on our behavioral and modelling analyses, changes in transducer nonlinearities accounts for our findings compared with multiplicative noise reduction.

Depending on different models of perception, changes in transducer nonlinearities may reflect two mechanisms, which are changes in nonlinear properties of the neurons or changes in stimulus uncertainty. While the PTM model used in this study does not parameterize stimulus uncertainty, Lu and Dosher (1999) explained that the nonlinear transducing properties implemented in the PTM are mathematically comparable to the uncertainty principle implemented in a similar model developed by Ahumada and Watson (1985), especially when the d' lies between 0.75 and 2.50 (the range within which threshold was determined in this study). Interestingly, these two mechanisms underlying changes in nonlinear transducer properties parallel the role of temporally correlated sound in perceptual enhancements on a contrast detection task. Given that any contributions from response criterion shifts have been ruled out due to the 2AFC nature of the task, observed cross-modal facilitation in a contrast detection task can be attributed to either the transient boosting hypothesis (Andersen and Mamassian, 2008) or the uncertainty reduction hypothesis (Lippert et al., 2007). The transient boosting hypothesis suggests that audiovisual integration may occur when auditory and visual intensity transients are perceived and is thought to underlie mechanisms such as the influence of auditory flutter rate on perceived visual flicker rate. Brain areas such as the superior colliculus involved in the processing of signal transients and the magnocellular pathway involved in the processing of low spatial, high temporal frequency stimuli have been thought to mediate the transient boosting of sensory energies and thus, parallels the Lu and Dosher (1999) explanation. On the other hand, the uncertainty reduction hypothesis, which parallels the Ahumada and Watson (1985) explanation, reflects a decision strategy rather than a genuine perceptual enhancement where the presentation of the transient sound signifies the onset of the visual stimulus, thereby reducing the ambiguity about when the visual information is presented (Chen et al., 2011). While both mechanisms could separately or in tandem account for the role of sound in our AVT paradigm, we consider the uncertainty explanation less likely for two reasons. First, if sound produced this alerting effect in our results, then the difference between the AVT and VOT groups should have been observed in the first sessions of training, however, both performance in both groups were comparable in the first session and only differed later in the training sessions (i.e., 7/8), indicating that the benefits of sound stemmed from repeated exposure and perhaps, enhanced associative learning. Second, if sound only alerted the temporal window of the visual stimulus, the detection performance should be better in the last training session (i.e., 7/8) compared with the post training session (i.e., 9/10) where sound was absent. However, in our study, we did not find any significant effect of group when an ANOVA was conducted with group and sessions (7/8 and 9/10) as the factors. Together, these findings imply that audiovisual training is superior to visual only training and may engage separate and or additional mechanisms to improve perceptual performance during and after training.

Finally, our findings have important clinical implications for using perceptual learning paradigms in low vision rehabilitation. In low vision rehabilitation, the ultimate goal is to improve patient’s visual performance to a level that translates into useful improvements in the performance of daily tasks, increasing their independence and enhancing their quality of life. However, current approaches have been mostly visual in nature and often limited gains and long training times are reported. Apart from the established effectiveness and efficiency of multisensory perceptual training paradigms, the finding of inverse effectiveness suggests that AVT training may improve perceptual outcomes better than VOT training in low vision patients where information reliability is reduced due to impaired structural and functional processing. Moreover, the fact that we can use simple stimuli (such as Gabor patches and sounds) and still obtain greater improvements open more avenue for the strategic design and implementation of AVT paradigms using artificially induced cross-modal associations. Despite the promising nature of multisensory perceptual learning paradigms, in order to facilitate their acceptance into clinical management, further research is needed to determine the degree of transferability of learning outcomes from trained to untrained tasks or locations as well as the stability of the learned effects.

4.6 References

Ahumada, A., and Watson, A. (1985). Equivalent-noise model for contrast detection and discrimination. *JOSA A, 2*(7), 1133-1139.

Andersen, T. S., and Mamassian, P. (2008). Audiovisual integration of stimulus transients. *Vision research, 48*(25), 2537-2544.

Arieh, Y., and Marks, L. E. (2008). Cross-modal interaction between vision and hearing: A speed—accuracy analysis. *Perception and psychophysics, 70*(3), 412-421.

Baldwin, A. S., Baker, D. H., and Hess, R. F. (2016). What do contrast threshold equivalent noise studies actually measure? Noise vs. nonlinearity in different masking paradigms. *PLoS One, 11*(3), e0150942.

Ball, K., and Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science, 218*(4573), 697-698.

Barlow, H. B. (1956). Retinal noise and absolute threshold. *JOSA, 46*(8), 634-639.

Bejjanki, V. R., Zhang, R., Li, R., Pouget, A., Green, C. S., Lu, Z.-L., and Bavelier, D. (2014). Action video game play facilitates the development of better perceptual templates. *Proceedings of the National Academy of Sciences, 111*(47), 16961-16966.

Bloch, A.-M. (1885). Experiences sur la vision. *Comptes Rendus de la Societé de Biologie, 37*, 493.

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

Burgess, A., Wagner, R., Jennings, R., and Barlow, H. B. (1981). Efficiency of human visual signal discrimination. *Science, 214*(4516), 93-94.

Camilleri, R., Pavan, A., Ghin, F., Battaglini, L., and Campana, G. (2014). Improvement of uncorrected visual acuity and contrast sensitivity with perceptual learning and transcranial random noise stimulation in individuals with mild myopia.

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., and Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Comput Biol, 5*(7), e1000436.

Chen, Y.-C., Huang, P.-C., Yeh, S.-L., and Spence, C. (2011). Synchronous sounds enhance visual sensitivity without reducing target uncertainty. *Seeing and perceiving, 24*(6), 623.

D'Zmura, M., and Knoblauch, K. (1998). Spectral bandwidths for the detection of color. *Vision research, 38*(20), 3117-3128.

Dosher, B., and Lu, Z.-L. (2017). Visual perceptual learning and models. *Annual Review of Vision Science, 3*, 343-363.

Dosher, B. A., and Lu, Z.-L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences, 95*(23), 13988-13993.

Dosher, B. A., and Lu, Z.-L. (1999). Mechanisms of perceptual learning. *Vision research, 39*(19), 3197-3221.

Durrie, D., and McMinn, P. S. (2007). Computer-based primary visual cortex training for treatment of low myopia and early presbyopia. *Trans Am Ophthalmol Soc, 105*, 132-138.

Fiorentini, A., and Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency.

Foley, J., and Tyler, C. W. (1976). Effect of stimulus duration on stereo and vernier displacement thresholds. *Perception and psychophysics, 20*(2), 125-128.

Foley, J. M., and Legge, G. E. (1981). Contrast detection and near-threshold discrimination in human vision. *Vision research, 21*(7), 1041-1053.

Gorea, A., and Tyler, C. W. (1986). New look at Bloch's law for contrast. *JOSA A, 3*(1), 52-61.

Graham, C. H., and Kemp, E. H. (1938). Brightness discrimination as a function of the duration of the increment in intensity. *The Journal of general physiology, 21*(5), 635.

Grant, K. W., and Seitz, P.-F. (2000). The use of visible speech cues for improving auditory detection of spoken sentences. *The Journal of the Acoustical Society of America, 108*(3), 1197-1208.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual neuroscience, 9*(2), 181-197.

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

JASP Team. (2018). JASP (Version 0.11)[Computer software].

Kahneman, D. (1966). Time-intensity reciprocity in acuity as a function of luminance and figure-ground contrast. *Vision research, 6*(3-4), 207-215.

Karni, A., and Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences, 88*(11), 4966-4970.

Kasten, E., Poggel, D. A., and Sabel, B. A. (2000). Computer-based training of stimulus detection improves color and simple pattern recognition in the defective field of hemianopic subjects. *Journal of cognitive neuroscience, 12*(6), 1001-1012.

Kim, R. S., Seitz, A. R., and Shams, L. (2008). Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS One, 3*(1), e1532.

Legge, G. E., and Foley, J. M. (1980). Contrast masking in human vision. *JOSA, 70*(12), 1458-1471.

Levi, D. M., and Li, R. W. (2009). Perceptual learning as a potential treatment for amblyopia: a mini-review. *Vision research, 49*(21), 2535-2549.

Li, X., and Lu, Z.-L. (2012). Enabling high grayscale resolution displays and accurate response time measurements on conventional computers. *Journal of visualized experiments: JoVE*(60).

Li, X., Lu, Z.-L., Xu, P., Jin, J., and Zhou, Y. (2003). Generating high gray-level resolution monochrome displays with conventional computer graphics cards and color monitors. *Journal of neuroscience methods, 130*(1), 9-18.

Lidestam, B., Moradi, S., Pettersson, R., and Ricklefs, T. (2014). Audiovisual training is better than auditory-only training for auditory-only speech-in-noise identification. *The Journal of the Acoustical Society of America, 136*(2), EL142-EL147.

Lippert, M., Logothetis, N. K., and Kayser, C. (2007). Improvement of visual contrast detection by a simultaneous sound. *Brain research, 1173*, 102-109.

Lu, Z.-L., Chu, W., Dosher, B. A., and Lee, S. (2005). Independent perceptual learning in monocular and binocular motion systems. *Proceedings of the National Academy of Sciences, 102*(15), 5624-5629.

Lu, Z.-L., and Dosher, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *JOSA A, 16*(3), 764-778.

Lu, Z.-L., and Dosher, B. A. (2004). Perceptual learning retunes the perceptual template in foveal orientation identification. *Journal of Vision, 4*(1), 5-5.

Lu, Z.-L., and Dosher, B. A. (2008). Characterizing observers using external noise and observer models: assessing internal representations with external noise. *Psychological review, 115*(1), 44.

Lu, Z.-L., and Dosher, B. A. (2009). Mechanisms of perceptual learning. *Learning and perception, 1*(1), 19-36.

Lu, Z.-L., and Sperling, G. (1996). Contrast gain control in first-and second-order motion perception. *JOSA A, 13*(12), 2305-2318.

Mayer, M. J. (1983). Practice improves adults' sensitivity to diagonals. *Vision research, 23*(5), 547-550.

Nachmias, J., and Sansbury, R. V. (1974). Grating contrast: discrimination may be better than detection. *Vision research*.

Nagaraja, N. (1964). Effect of luminance noise on contrast thresholds. *JOSA, 54*(7), 950-955.

Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H.-J., and Driver, J. (2007). Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *Journal of Neuroscience, 27*(42), 11431-11441.

Noesselt, T., Tyll, S., Boehler, C. N., Budinger, E., Heinze, H.-J., and Driver, J. (2010). Sound-induced enhancement of low-intensity vision: multisensory influences on human sensory-specific cortices and thalamic bodies relate to perceptual enhancement of visual detection sensitivity. *Journal of Neuroscience, 30*(41), 13609-13623.

Odgaard, E. C., Arieh, Y., and Marks, L. E. (2003). Cross-modal enhancement of perceived brightness: sensory interaction versus response bias. *Perception and psychophysics, 65*(1), 123-132.

Ohzawa, I., Sclar, G., and Freeman, R. (1982). Contrast gain control in the cat visual cortex. *Nature, 298*(5871), 266-268.

Pelli, D. G. (1981). *Effects of visual noise.* Citeseer.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision, 10*(4), 437-442.

Polat, U., Ma-Naim, T., Belkin, M., and Sagi, D. (2004). Improving vision in adult amblyopia by perceptual learning. *Proceedings of the National academy of Sciences of the United States of America, 101*(17), 6692-6697.

Polat, U., Ma-Naim, T., and Spierer, A. (2009). Treatment of children with amblyopia by perceptual learning. *Vision research, 49*(21), 2599-2603. doi: 10.1016/j.visres.2009.07.008

Polat, U., Schor, C., Tong, J.-L., Zomet, A., Lev, M., Yehezkel, O., . . . Levi, D. M. (2012). Training the brain to overcome the effect of aging on the human eye. *Scientific reports, 2*.

Powers, A. R., Hevey, M. A., and Wallace, M. T. (2012). Neural correlates of multisensory perceptual learning. *Journal of Neuroscience, 32*(18), 6263-6274.

Reisbeck, T., and Gegenfurtner, K. (1997). *Velocity tuned mechanisms in human motion perception.* Paper presented at the Annual Meeting of the Association for Research in Vision and Ophthalmology (ARVO 1997).

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

Sahraie, A., MacLeod, M.-J., Trevethan, C. T., Robson, S. E., Olson, J. A., Callaghan, P., and Yip, B. (2010). Improved detection following Neuro-Eye Therapy in patients with post-geniculate brain damage. *Experimental brain research, 206*(1), 25-34.

Seitz, A. R., Kim, R., and Shams, L. (2006). Sound facilitates visual learning. *Current Biology, 16*(14), 1422-1427.

Shams, L., and Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in cognitive sciences*, *12*(11), 411-417.

Shams, L., Wozny, D. R., Kim, R., and Seitz, A. (2011). Influences of multisensory experience on subsequent unisensory processing. *Frontiers in psychology, 2*.

Sheffert, S. M., and Olson, E. (2004). Audiovisual speech facilitates voice learning. *Perception and psychophysics, 66*(2), 352-362.

Stein, B. E., London, N., Wilkinson, L. K., and Price, D. D. (1996). Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. *Journal of cognitive neuroscience, 8*(6), 497-506.

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

Watson, A. B., and Solomon, J. A. (1997). Model of visual contrast gain control and pattern masking. *JOSA A, 14*(9), 2379-2391.

Wozny, D., Seitz, A., and Shams, L. (2008). Learning associations between simple visual and auditory features. *Journal of Vision, 8*(6), 171-171.

Zäske, R., Mühl, C., and Schweinberger, S. R. (2015). Benefits for voice learning caused by concurrent faces develop over time. *PLoS One*, *10*(11), e0143151.

Zilber, N., Ciuciu, P., Gramfort, A., Azizi, L., and Van Wassenhove, V. (2014). Supramodal processing optimizes visual perceptual learning and plasticity. *Neuroimage, 93*, 32-46.

Chapter 5

General Discussion

5.1 Low Vision: Clinical Model of Adult-onset Visual Deprivation

The sense of vision is vitally important in the vast majority of mammalian species, including Homo sapiens, and represents a critical means for perceiving the environment and guiding actions and behavior. As such, structural and/or functional damage to the visual system has devastating consequences on perception and behavior. Injury at various levels of the visual hierarchy results in visual impairments that differ in their nature, severity, and overall functional consequence (Huxlin, 2008; Sabel, Henrich-Noack, Fedorov, & Gall, 2011). For instance, while an insult to structures anywhere along the processing hierarchy from the retina to primary visual cortex (V1) may deprive higher areas of basic sensory input and lead to either focal or global loss of vision, injury to higher-level extrastriate areas can result in more selective deficits in visual perception such as weaknesses in perceiving faces (Huxlin, 2008).

Damage to visual structures including the retina and visual cortex can result in chronic loss of vision that cannot be remedied by refractive procedures or medication and that leads to profound difficulties in performance of visually guided activities (Legge & Chung, 2016). This condition, termed low vision, is clinically assessed using measures of visual acuity, contrast sensitivity and visual field tests. Quantitatively, a visual acuity measure worse than 20/60 (i.e., inability to discriminate between two black bars separated by 3 minutes of an arc or more) or a visual field extent of less than 20° in the better eye represent the criteria for low vision (Legge & Chung, 2016). The large majority of disorders resulting in low vision are a result of damage to the retina. Age-related macular degeneration (AMD), glaucoma and diabetic retinopathy accounted for almost a half of all blindness and low vision cases in United States according to Congdon et al (Congdon et al., 2004). Depending upon the location of the retinal damage, central or peripheral vision can be selectively compromised in these conditions. Post-retinal injuries normally result in a condition termed hemianopia, that is loss of vision in the visual field location represented by the damaged region; given that, one striking characteristic of many visual brain regions is a retinotopic organization. Regardless of the nature of vision loss, the devastating consequences of low vision including reduced quality of life (Scott, Smiddy, Schiffman, Feuer, & Pappas, 1999) and high global economic burden (Frick, Gower, Kempen, & Wolff, 2007) call for the development of improved rehabilitative approaches that can restore some degree of visual function in patients with low vision.

5.2 Clinical Implications for Injury in Low Vision

5.2.1 Multisensory temporal perception is inherently plastic

The findings from this dissertation support previous findings in the literature demonstrating that the multisensory temporal system is inherently plastic and can rapidly adapt to changes in the environment. For instance, human participants can shift their point of subjective simultaneity after being exposed to several minutes of an asynchronous audiovisual stimulus of fixed stimulus onset asynchrony (Fujisaki, Shimojo, Kashino, & Nishida, 2004; Van der Burg, Alais, & Cass, 2013; Vroomen, Keetels, De Gelder, & Bertelson, 2004).

This phenomenon is generally referred to as temporal recalibration. Interestingly, this short-term plasticity is independent of the nature of the judgment required, i.e., whether participants were asked to report synchronous/asynchronous as in SJ task or to judge the order of the stimulus presentations as in a TOJ task (Vroomen et al., 2004). Similar to the impact of short-term monocular deprivation on multisensory temporal perception, temporal recalibration declines in magnitude even if the individual was re-exposed to the adaptation stimulus (Machulla, Di Luca, Froehlich, & Ernst, 2012). However, the aftereffect of exposure to asynchrony dissipates when stimulus conditions involving a different asynchrony than the one used during exposure is introduced as counterevidence (Machulla et al., 2012). Recently, there is evidence showing that temporal recalibration can occur at very short and rapid timescales including on a trial-to-trial basis (Van der Burg et al., 2013). Rapid recalibration is thought to reflect mechanisms involved in overcoming inevitable variations in audiovisual timing and realigning audiovisual signals at onset to maximize the perceptual benefits of audiovisual integration (Van der Burg et al., 2013). Furthermore, the plasticity of multisensory temporal processing has been evinced through perceptual learning (Powers, Hevey, & Wallace, 2012; Powers, Hillock, & Wallace, 2009). For instance, training participants on audiovisual temporal simultaneity task with feedback narrowed their temporal binding window (Powers et al., 2009). Although perceptual learning occurs on a longer timescale compared to the effects of monocular deprivation and exposure to asynchronous stimuli, in Powers et al. (2009) study, the only significant enhancement occurred during the first hour of training indicating that the temporal binding window can rapidly change following perceptual learning. Taken together, these findings together with the effects of short-term monocular deprivation observed in Chapters 2 and

3 of this dissertation emphasize the idea that the multisensory temporal processing and perception is malleable and can rapidly adapt to changes in the environment.

5.2.2 Where in the brain do deprivation-induced changes in multisensory temporal processing occur?

Multisensory temporal processing involves a network of brain regions including sensory specific regions such as the primary auditory and visual cortices, and higher association regions such as superior temporal sulcus, intraparietal sulcus, insula, and several foci in the frontal lobe, including within the superior and ventromedial frontal gyri (Calvert et al., 1999; Noesselt et al., 2007; Noesselt et al., 2010). According to Bayesian causal inference models of multisensory perception, changes in the multisensory temporal perception could occur in early sensory areas reflecting changes in unisensory encoding and/or in high order association areas reflecting changes in the tendency to bind the audiovisual information (Körding et al., 2007; Magnotti, Ma, & Beauchamp, 2013). Although, our findings do not directly implicate a subset of those regions, we can make inferences based on our behavioral and neuroimaging findings. First, from our behavioral results in chapter 2 and chapter 3, we observe that short-term monocular deprivation caused a reduction in the maximum probability of simultaneity perception and a reduction in the temporal binding window. While changes in the temporal binding alone may support both changes in early sensory and high order association areas (see discussion in chapter 2), changes in the maximum probability of simultaneity perception imply changes in participant’s tendency to bind audiovisual stimuli, which reflects changes in higher order decisional areas (Körding et al., 2007; Magnotti et al., 2013). Second, power analysis of post-stimulus neural oscillatory activity in alpha band showed significant clusters in both occipital-parietal and frontal areas during visual-only stimulation after short-term monocular deprivation. While occipital-parietal alpha is associated with encoding of sensory information, frontal alpha power reflects mechanisms influencing perceptual decisional processes (Grabot, Kösem, Azizi, & Van Wassenhove, 2017). Considering the behavioral and EEG results together, we can hypothesize that short-term monocular deprivation effects on multisensory temporal perception may reflect changes in both early sensory areas and late-stage decisional areas in the brain.

In the future, Bayesian casual inference model can be employed to provide a mechanistic understanding of whether these changes in multisensory temporal perception after monocular deprivation can be expressed in parameters that reflect sensory encoding processes and or those that reflect changes in decisional priors. (See appendix for an implementation of Bayesian causal inference model in understanding binocular summation in multisensory temporal perception). In addition, from a neuroimaging standpoint, source localization analysis of EEG signals can be used to probe which areas in the brain are involved in the changes observed after monocular deprivation. Furthermore, it will be interesting to know whether there is strengthening vs weakening of functional relationships between various areas such early sensory and high association areas following short-term monocular deprivation. Lastly, it will be informative to determine whether correlations exist between parameters of Bayesian causal inference models and source localization measures obtained from EEG source localization analysis.

5.2.3 Changes in multisensory temporal processing may occur in low vision

Based on chapter 2 and 3, we found that short-term monocular deprivation enhanced multisensory temporal acuity (i.e., the ability to discriminate between the timing of auditory and visual events) indexed as reduced temporal binding window for the deprived eye while worsening acuity for the non-deprived eye (note this finding was only significant in chapter 2 but not in chapter 3). These findings were consistent with studies using pure visual tasks such as binocular rivalry and contrast detection tasks, showing that short-term monocular deprivation boosted the perceptual abilities of the deprived eye and vice versa for the non-deprived (Lunghi, Burr, & Morrone, 2011; Lunghi, Burr, & Morrone, 2013; Zhou, Clavagnier, & Hess, 2013). These changes were thought to reflect the upregulation of contrast-gain control mechanisms geared towards boosting the visual signal in the deprived eye following the absence of visual information during the deprivation period (Lunghi et al., 2011; Lunghi, Emir, Morrone, & Bridge, 2015).

Considering that low vision represents a state of partial visual deprivation, how do we relate this to changes in sensory and perceptual abilities that may occur in low vision patients? While in general, our findings suggest that low vision patients may experience not only changes in visual perception but also changes in multisensory perception, it is not obvious what the exact nature of these changes might be. Still, we believe the early perceptual changes that may occur in low vision patients as a result of spontaneous plasticity may resemble findings from our studies together with those in pure visual studies. These early perceptual changes in low vision patients may present as periods of both unisensory and multisensory perceptual boosting which are a result of compensatory mechanisms to the early periods of visual deprivation. On the other hand, it is possible that as low vision progresses, the ensuing visual deprivation may be detrimental to visual processing and thus, multisensory processing as well. For example, while about 2 hours of monocular deprivation resulted in enhanced cortical excitability (Lunghi, Berchicci, Morrone, & Di Russo, 2015; Lunghi et al., 2011), in a study where transcranial magnetic stimulation was used to induce phosphine perception, 48 hours of monocular deprivation resulted in reduced cortical excitability (Lou et al., 2011). Given the evidence that patients who suffered early visual deprivation had enlarged temporal binding windows in adult life (Chen, Lewis, Shore, & Maurer, 2017; Richards, Goltz, & Wong, 2017), it is plausible to hypothesize that following a long period of visual deprivation occurring when low vision progresses, multisensory temporal perceptual abilities would deteriorate in low vision. This would be consistent with the finding that blurring a visual image increased the temporal binding window as a result of increased sensory noise affecting the measures of physical asynchrony (Magnotti et al., 2013). Nevertheless, as low vision progresses, it is possible that changes in multisensory temporal perception will also reflect changes in tendency to bind audiovisual stimuli reflecting high order decisional changes as shown by cross-modal recalibration findings in blind cohorts (C Büchel, 1998; Christian Büchel, Price, Frackowiak, & Friston, 1998; Burton et al., 2002; Held, Freedman, & Harris, 1996).

Surprisingly, the literature on multisensory processing in low-vision conditions - especially those due to retinal degeneration - remains sparse and hence, provides an opportune avenue for more research. In the near future, it would be interesting to investigate the impact of short-term and long-term visual impairment on multisensory processing, and the mechanisms (reduced sensory encoding or improved audiovisual binding tendency or both) underlying changes in the multisensory interaction in both instances. A further key study on this topic will be to investigate the neural correlates of these enhancements observed under visually impaired conditions. Undoubtedly, the findings from these future studies will provide a better and more thorough picture of spontaneous plasticity mechanisms in visual and multisensory cortical processing during visual impairment and the impact on perception and behavior. Lastly, these would also inform the creation of more efficient, effective, and comprehensive therapies to promote improved visual function in these conditions.

5.3 Clinical Implications for Recovery and Rehabilitation in Low Vision

The ability of the brain to reorganize its structure and function does not only imply changes following an insult but also imply hope for possible recovery and rehabilitation. In low vision patients, the goal of rehabilitation is to significantly improve patient's visual performance to a level that translates into useful behavioral improvements in the performance of daily tasks (Legge and Chung, 2016). Ultimately, this is expected to increase patients' independence and ameliorate their quality of life. However, training has to be more effective and efficient in order not to lose patients due to lack of commitment and compliance or possibly to doubt about utility of the paradigm. Some of these issues have been at the forefront in terms of honing the efficacy of rehabilitative approaches in low vision research.

5.3.1 Perceptual Learning, a means for rehabilitation in Low vision

Perceptual learning approaches used in vision rehabilitation are founded in the concept of plasticity (W. Li, 2016). Perceptual learning has been used as a tool for improving visual function in conditions including but not limited to amblyopia (Levi and Li, 2009; Polat, Ma-Naim, Belkin, and Sagi, 2004; Polat, Ma-Naim, and Spierer, 2009), hemianopia (Kasten, Poggel, and Sabel, 2000; Sahraie et al., 2010), glaucoma, AMD, myopia and presbyopia (Camilleri, Pavan, Ghin, Battaglini, and Campana, 2014; Durrie and McMinn, 2007; Polat et al., 2012). Despite the unique etiologies of these different conditions, studies have managed to employ successfully perceptual learning paradigms that are tailored to the rehabilitative needs of patients with these conditions. In the following, we briefly review paradigms that have been used in patients suffering from central and peripheral vision loss.

5.3.1.1 Perceptual Learning for Central Vision Loss

Perceptual learning protocols used in the rehabilitation of central vision loss are designed to train intact retinal regions referred to as preferred retinal loci (PRL) (Schuchard, 1994). These regions are peripheral locations in the visual field that, following training, can be used by patients for the performance of everyday perceptual activities such as reading and walking that are typically mediated by central vision (Cheung and Legge, 2005; Crossland, Culham, Kabanarou, and Rubin, 2005; Cummings, Whittaker, Watson, and Budd, 1985; Schuchard, 1994). In effect, PRL functions as the new 'fovea'. Perceptual learning paradigms in this domain have utilized several techniques such as rapid serial presentation of text or words and contrast detection training to improve vision and reading abilities at the PRL. For example, (Tarita-Nistor, Brent, Steinbach, Markowitz, and González, 2014) trained patients with binocular central vision loss at their PRL on continually presented text adjusted to each patient's reading acuity threshold. After four consecutive training sessions, the patients showed significantly reduced average required reading time, enhanced reading acuity, faster reading speed, and improved stability of fixation. More recently, (Maniglia et al., 2016) trained central vision loss patients on a contrast detection task using two different procedures and recorded significant improvements on both procedures after training. These improvements were found to be stable for 4 and 6 months after training (Maniglia et al., 2016).

5.3.1.2 Perceptual Learning for Peripheral Vision Loss

In peripheral vision loss, perceptual learning approaches typically involve the stimulation of areas of residual vision surrounding the focal retinal lesions or damage. This procedure has been termed Visual Restoration Therapy (VRT). According to the residual vision activation theory put forth by (Sabel et al., 2011), areas of total blindness in visual field defects do not sharply transition into areas of normal vision. Instead, there exists an intervening region of partial damage with surviving neurons. The survival of a certain minimum number of neurons can serve as a foundation for neural reorganization through synaptic plasticity (Sabel et al., 2011). Indeed, evidence from animal model studies has shown that as few as 10-20% of normal retinal ganglion cell numbers following optic nerve damage is sufficient to support some degree of recovery of visually guided behavior (Sabel, 1997; Sautter and Sabel, 1993). VRT has been successfully employed in conditions such as hemianopia, optic neuritis, and glaucoma. For example, Sabel and Gudlin (2014) used a perceptual learning paradigm in 30 glaucoma patients (half in the experimental group, half in the control group) where areas of residual vision were stimulated for one-hour every day for 3 months. Significant enhancements in detection accuracy and response times in high-resolution perimetry were observed following training (Sabel and Gudlin, 2014). Another study employed VRT for 6 months in 302 patients with visual field defects due to conditions such as ischemia, hemorrhage, head trauma, tumor removal or anterior ischemic optic neuropathy (Mueller, Mast, and Sabel, 2007). After therapy, significant improvements were seen in the detection of supra-threshold stimuli, response times were speeded and there was an enlargement of visual fields (Mueller et al., 2007). Despite the seeming success of VRT, these approaches came under intense criticism by some who proposed that improvements observed after training were attributable to excessive saccadic eye movements (Dundon, Bertini, Làdavas, Sabel, and Gall, 2015). Nonetheless, other studies have confirmed that patients who undergo VRT seem to have good fixation abilities, making this less of a concern (Dundon et al., 2015).

5.3.2 Multisensory perspective on perceptual learning in low vision rehabilitation

Previous evidence and those produced by chapter 4 of this dissertation support the notion that multisensory perceptual learning facilitates visual learning more than visual-only learning paradigms do (Kim, Seitz, and Shams, 2008; Seitz, Kim, and Shams, 2006; Shams and Seitz, 2008; Shams, Wozny, Kim, and Seitz, 2011). In low vision conditions, visual sensory processing is affected in ways that lead to reduced gray matter volume (C. Li et al., 2012) and activations (Duncan, Sample, Weinreb, Bowd, and Zangwill, 2007) in visual cortical areas. In addition, outflow of information from visual cortical areas to other sensory systems and upstream target areas such as attentional networks may be impaired. Whereas visual learning might deal with enhancing activation and representations in visual processing areas and strengthening connectivity with upstream targets, through evidence provided, multisensory learning might improve upon the outcomes from visual learning while improving connectivity with other sensory systems (Figure 5.1).

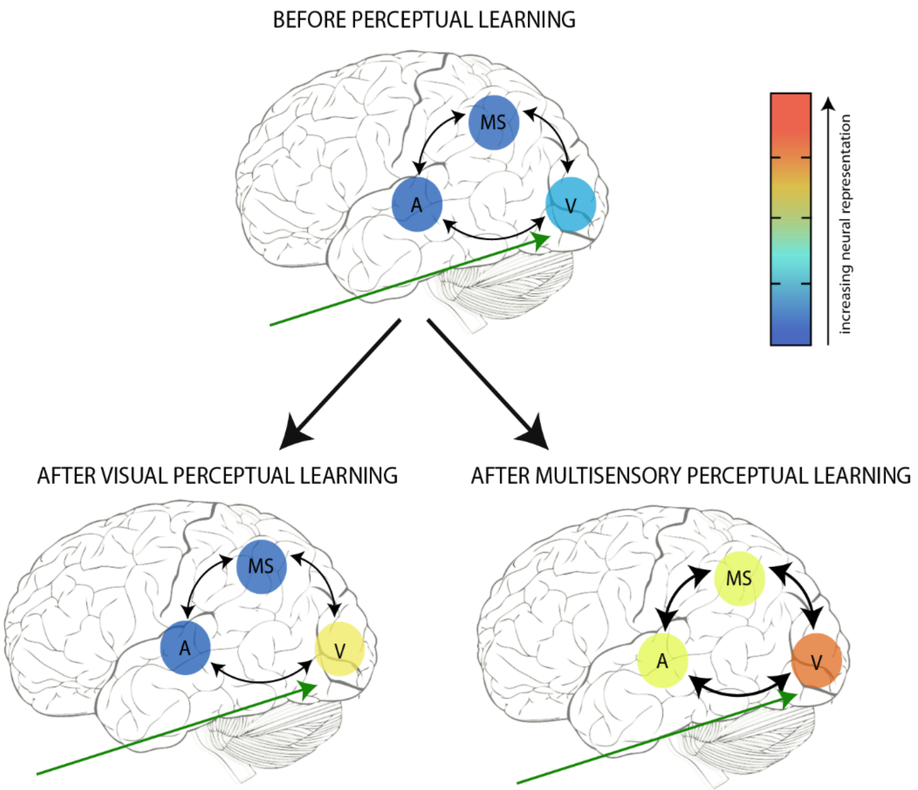

Figure 5.1 Proposed Neural Mechanisms Underlying Visual and Multisensory Perceptual Learning (PL) in Low Vision.

Top: Before PL, incoming visual information (green arrow) will weakly activate visual areas (V) due to impaired visual processing. Auditory (A) and multisensory areas (MS) may not be activated over baseline. Bottom left: Visual PL only engages structures in the visual system leading to enhanced neural representations in the visual system. Bottom right: On the other hand, multisensory PL activates all three sensory areas leading to the formation or modification of neural representations in all three systems and strengthening of connectivity (thick arrows) among sensory areas. In addition, multisensory PL may lead to increased neural activation in visual areas than visual PL.

Low vision rehabilitation may benefit from some attributes of multisensory perceptual learning paradigms. First, the increased robustness of multisensory paradigms compared to unisensory paradigms makes them more suitable for low vision rehabilitation, as the more improvement in visual performance, the more likely a patient may gain independence in carrying out daily visual tasks and experience an increase in quality of life. Secondly, not only are multisensory approaches more effective, but they also require a relatively small number of sessions for performance to reach its asymptotic level. This attribute is very crucial because the commitment in time and effort required by these patients sometimes serve as a discouragement to enrolling for visual training. The realization of this attribute in low vision rehabilitation approaches will remarkably augment patient compliance and commitment levels, which are key to successful therapy and rehabilitation. Another attribute that makes the application of multisensory perceptual learning approaches plausible in low vision rehabilitation is its comprehensibility. As discussed earlier, operations of different sensory systems are very intertwined even at the level of primary cortical areas, implying that damage to one modality may affect how it interacts with other modalities, leading to degradation of connectivity patterns among these brain regions. This demands that rehabilitative approaches consider the affected sensory modality as well as other modalities whose interaction with the former may be impaired.

Considering the fact that low vision rehabilitation programs utilize a broad range of stimuli and tasks including simple light and contrast detection tasks for their training paradigms, it was crucial to establish that multisensory-based perceptual learning could produce similar effects in improvement across a wide variety of visual stimuli (e.g., flashes, Gabor patch) and tasks. Our investigation is relevant as previous studies have mostly employed cross-modal stimuli (visual and auditory motion stimuli) that shared some complex congruent characteristics such as direction of motion and have demonstrated the importance of stimulus congruency to visual learning under multisensory facilitation training. Moreover, Shams et al. (2011) proposed that mechanisms leading to unisensory benefits after multisensory learning might require sensory features that are ecologically associated, such as auditory and visual motion, or lip movements and voice, etc. The findings presented in chapter 4 imply that cross-modal stimuli with artificial associations that can enhance binding can produce learning outcomes that are better than vision-only paradigms. This opens up new avenues for the strategic design and implementation of perceptual learning paradigms for low vision rehabilitation. Nevertheless, we do not assert that artificially induced cross-modal associations may produce more effective and stable learning outcomes compared with ecological valid cross-modal features or stimuli. Further studies will be needed to test that possibility.

Despite the promising nature of the multisensory approach for low vision rehabilitation, further research is necessary for its acceptance into the clinical domain. First, it is important to investigate the generalizability of these multisensory training protocols to performance of daily life activities. According to the ‘reverse hierarchy theory’ of visual perceptual learning, generalization may occur when training drives more of the processing and attentional resources to higher level cortical association areas (Ahissar & Hochstein, 2004). Hence, this theory implies that multisensory perceptual learning paradigms may enhance generalization of learning outcomes due to their ability to recruit and engaged higher order multisensory brain areas (Proulx, Brown, Pasqualotto, & Meijer, 2014; Zilber, Ciuciu, Gramfort, Azizi, & Van Wassenhove, 2014). Understanding this aspect of multisensory training is crucial for its acceptance as a form of therapy in low vision patients as one of the prominent aims of perceptual therapy in patients should be to rehabilitate; that is the translation of perceptual benefits into measurable improvements in the daily life of the patient.

Another crucial study will be to determine the duration of effectiveness of these enhancements because, after all, training benefits lasting for only a few hours or days would not justify the time and expense patients would have to invest in the treatment. On the same argument of matching patients' effort to benefits after training, we also need to learn what are the structural and functional predictors of possible improvement in these patients after multisensory training. This knowledge could facilitate development of sensitive and specific inclusion and exclusion criteria that will enable proper screening prior to training. Lastly, the relevance of findings that will evolve from the investigation of the neural correlates of the enhancements seen in these patients cannot be overemphasized.

5.4 References

Ahissar, M., and Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in cognitive sciences, 8*(10), 457-464.

Büchel, C. (1998). Functional neuroimaging studies of Braille reading: cross-modal reorganization and its implications. *Brain: a journal of neurology, 121*(7), 1193-1194.

Büchel, C., Price, C., Frackowiak, R., and Friston, K. (1998). Different activation patterns in the visual cortex of late and congenitally blind subjects. *Brain: a journal of neurology, 121*(3), 409-419.

Burton, H., Snyder, A. Z., Conturo, T. E., Akbudak, E., Ollinger, J. M., and Raichle, M. E. (2002). Adaptive changes in early and late blind: a fMRI study of Braille reading. *Journal of neurophysiology, 87*(1), 589-607.

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

Camilleri, R., Pavan, A., Ghin, F., Battaglini, L., and Campana, G. (2014). Improvement of uncorrected visual acuity and contrast sensitivity with perceptual learning and transcranial random noise stimulation in individuals with mild myopia.

Chen, Y.-C., Lewis, T. L., Shore, D. I., and Maurer, D. (2017). Early binocular input is critical for development of audiovisual but not visuotactile simultaneity perception. *Current Biology, 27*(4), 583-589.

Cheung, S. H., and Legge, G. E. (2005). Functional and cortical adaptations to central vision loss. *Vis Neurosci, 22*(2), 187-201. doi: 10.1017/s0952523805222071

Congdon, N., O'Colmain, B., Klaver, C., Klein, R., Muñoz, B., Friedman, D. S., . . . Mitchell, P. (2004). Causes and prevalence of visual impairment among adults in the United States. *Archives of Ophthalmology (Chicago, Ill.: 1960), 122*(4), 477-485.

Crossland, M. D., Culham, L. E., Kabanarou, S. A., and Rubin, G. S. (2005). Preferred retinal locus development in patients with macular disease. *Ophthalmology, 112*(9), 1579-1585.

Cummings, R. W., Whittaker, S. G., Watson, G. R., and Budd, J. M. (1985). Scanning characters and reading with a central scotoma. *American journal of optometry and physiological optics, 62*(12), 833-843.

Duncan, R. O., Sample, P. A., Weinreb, R. N., Bowd, C., and Zangwill, L. M. (2007). Retinotopic organization of primary visual cortex in glaucoma: Comparing fMRI measurements of cortical function with visual field loss. *Progress in retinal and eye research, 26*(1), 38-56.

Dundon, N. M., Bertini, C., Làdavas, E., Sabel, B. A., and Gall, C. (2015). Visual rehabilitation: visual scanning, multisensory stimulation and vision restoration trainings. *Frontiers in behavioral neuroscience, 9*.

Durrie, D., and McMinn, P. S. (2007). Computer-based primary visual cortex training for treatment of low myopia and early presbyopia. *Trans Am Ophthalmol Soc, 105*, 132-138.

Frick, K. D., Gower, E. W., Kempen, J. H., and Wolff, J. L. (2007). Economic impact of visual impairment and blindness in the United States. *Archives of Ophthalmology, 125*(4), 544-550.

Fujisaki, W., Shimojo, S., Kashino, M., and Nishida, S. y. (2004). Recalibration of audiovisual simultaneity. *Nature neuroscience, 7*(7), 773-778.

Grabot, L., Kösem, A., Azizi, L., and Van Wassenhove, V. (2017). Prestimulus alpha oscillations and the temporal sequencing of audiovisual events. *Journal of cognitive neuroscience, 29*(9), 1566-1582.

Held, R., Freedman, S., and Harris, C. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature, 380*, 11.

Huxlin, K. R. (2008). Perceptual plasticity in damaged adult visual systems. *Vision Research, 48*(20), 2154-2166.

Kasten, E., Poggel, D. A., and Sabel, B. A. (2000). Computer-based training of stimulus detection improves color and simple pattern recognition in the defective field of hemianopic subjects. *Journal of cognitive neuroscience, 12*(6), 1001-1012.

Kim, R. S., Seitz, A. R., and Shams, L. (2008). Benefits of stimulus congruency for multisensory facilitation of visual learning. *PloS one, 3*(1), e1532.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., and Shams, L. (2007). Causal inference in multisensory perception. *PloS one, 2*(9), e943.

Legge, G. E., and Chung, S. T. (2016). Low vision and plasticity: Implications for rehabilitation. *Annual Review of Vision Science, 2*, 321-343.

Levi, D. M., and Li, R. W. (2009). Perceptual learning as a potential treatment for amblyopia: a mini-review. *Vision Research, 49*(21), 2535-2549.

Li, C., Cai, P., Shi, L., Lin, Y., Zhang, J., Liu, S., . . . Li, S. (2012). Voxel-based morphometry of the visual-related cortex in primary open angle glaucoma. *Current eye research, 37*(9), 794-802.

Li, W. (2016). Perceptual Learning: Use-Dependent Cortical Plasticity. *Annual Review of Vision Science, 2*, 109-130.

Lou, A. R., Madsen, K. H., Paulson, O. B., Julian, H. O., Prause, J. U., Siebner, H. R., and Kjaer, T. W. (2011). Monocular visual deprivation suppresses excitability in adult human visual cortex. *Cerebral Cortex, 21*(12), 2876-2882.

Lunghi, C., Berchicci, M., Morrone, M. C., and Di Russo, F. (2015). Short-term monocular deprivation alters early components of visual evoked potentials. *The Journal of physiology, 593*(19), 4361-4372.

Lunghi, C., Burr, D. C., and Morrone, C. (2011). Brief periods of monocular deprivation disrupt ocular balance in human adult visual cortex. *Current Biology, 21*(14), R538-R539.

Lunghi, C., Burr, D. C., and Morrone, M. C. (2013). Long-term effects of monocular deprivation revealed with binocular rivalry gratings modulated in luminance and in color. *Journal of Vision, 13*(6), 1-1.

Lunghi, C., Emir, U. E., Morrone, M. C., and Bridge, H. (2015). Short-term monocular deprivation alters GABA in the adult human visual cortex. *Current Biology, 25*(11), 1496-1501.

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

Magnotti, J. F., Ma, W. J., and Beauchamp, M. S. (2013). Causal inference of asynchronous audiovisual speech. *Frontiers in psychology, 4*, 798.

Maniglia, M., Pavan, A., Sato, G., Contemori, G., Montemurro, S., Battaglini, L., and Casco, C. (2016). Perceptual learning leads to long lasting visual improvement in patients with central vision loss. *Restorative neurology and neuroscience, 34*(5), 697-720.

Mueller, I., Mast, H., and Sabel, B. A. (2007). Recovery of visual field defects: a large clinical observational study using vision restoration therapy. *Restorative neurology and neuroscience, 25*(5, 6), 563-572.

Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H.-J., and Driver, J. (2007). Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *Journal of Neuroscience, 27*(42), 11431-11441.

Noesselt, T., Tyll, S., Boehler, C. N., Budinger, E., Heinze, H.-J., and Driver, J. (2010). Sound-induced enhancement of low-intensity vision: multisensory influences on human sensory-specific cortices and thalamic bodies relate to perceptual enhancement of visual detection sensitivity. *Journal of Neuroscience, 30*(41), 13609-13623.

Polat, U., Ma-Naim, T., Belkin, M., and Sagi, D. (2004). Improving vision in adult amblyopia by perceptual learning. *Proceedings of the National academy of Sciences of the United States of America, 101*(17), 6692-6697.

Polat, U., Ma-Naim, T., and Spierer, A. (2009). Treatment of children with amblyopia by perceptual learning. *Vision research, 49*(21), 2599-2603. doi: 10.1016/j.visres.2009.07.008

Polat, U., Schor, C., Tong, J.-L., Zomet, A., Lev, M., Yehezkel, O., . . . Levi, D. M. (2012). Training the brain to overcome the effect of aging on the human eye. *Scientific reports, 2*.

Powers, A. R., Hevey, M. A., and Wallace, M. T. (2012). Neural correlates of multisensory perceptual learning. *Journal of Neuroscience, 32*(18), 6263-6274.

Powers, A. R., Hillock, A. R., and Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *The Journal of neuroscience, 29*(39), 12265-12274.

Proulx, M. J., Brown, D. J., Pasqualotto, A., and Meijer, P. (2014). Multisensory perceptual learning and sensory substitution. *Neuroscience and Biobehavioral Reviews, 41*, 16-25.

Richards, M. D., Goltz, H. C., and Wong, A. M. (2017). Alterations in audiovisual simultaneity perception in amblyopia. *PloS one*, *12*(6), e0179516.

Sabel, B. A. (1997). Unrecognized Potential of Surviving Neurons:: Within-systems Plasticity, Recovery of Function, and the Hypothesis of Minimal Residual Structure. *The Neuroscientist*, *3*(6), 366-370.

Sabel, B. A., and Gudlin, J. (2014). Vision restoration training for glaucoma: a randomized clinical trial. *JAMA ophthalmology, 132*(4), 381-389.

Sabel, B. A., Henrich-Noack, P., Fedorov, A., and Gall, C. (2011). Vision restoration after brain and retina damage: the "residual vision activation theory". *Prog Brain Res, 192*(1), 199-262.

Sahraie, A., MacLeod, M.-J., Trevethan, C. T., Robson, S. E., Olson, J. A., Callaghan, P., and Yip, B. (2010). Improved detection following Neuro-Eye Therapy in patients with post-geniculate brain damage. *Experimental brain research, 206*(1), 25-34.

Sautter, J., and Sabel, B. A. (1993). Recovery of brightness discrimination in adult rats despite progressive loss of retrogradely labelled retinal ganglion cells after controlled optic nerve crush. *European Journal of Neuroscience, 5*(6), 680-690.

Schuchard, R. (1994). Preferred retinal locus: a review with application in low vision rehabilitation. *Low Vision and Vision Rehabil, 7,* 243-256.

Scott, I. U., Smiddy, W. E., Schiffman, J., Feuer, W. J., and Pappas, C. J. (1999). Quality of life of low-vision patients and the impact of low-vision services. *American Journal of Ophthalmology, 128*(1), 54-62.

Seitz, A. R., Kim, R., and Shams, L. (2006). Sound facilitates visual learning. *Current Biology, 16*(14), 1422-1427.

Shams, L., and Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in cognitive sciences, 12*(11), 411-417.

Shams, L., Wozny, D. R., Kim, R., and Seitz, A. (2011). Influences of multisensory experience on subsequent unisensory processing. *Frontiers in psychology, 2.*

Tarita-Nistor, L., Brent, M. H., Steinbach, M. J., Markowitz, S. N., and González, E. G. (2014). Reading training with threshold stimuli in people with central vision loss: a feasibility study. *Optometry and Vision Science, 91*(1), 86-96.

Van der Burg, E., Alais, D., and Cass, J. (2013). Rapid recalibration to audiovisual asynchrony. *Journal of Neuroscience, 33*(37), 14633-14637.

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

Zhou, J., Clavagnier, S., and Hess, R. F. (2013). Short-term monocular deprivation strengthens the patched eye's contribution to binocular combination. *Journal of Vision, 13*(5), 12-12.

Zilber, N., Ciuciu, P., Gramfort, A., Azizi, L., and Van Wassenhove, V. (2014). Supramodal processing optimizes visual perceptual learning and plasticity. *Neuroimage, 93*, 32-46.

Appendix A

Appendix Ch A: Binocular enhancement of multisensory temporal perception

The contents of this chapter are adapted from

Opoku-Baah, C. and Wallace, M.T., 2021. Binocular Enhancement of Multisensory Temporal Perception. Investigative ophthalmology and visual science, 62(3), pp.7-7.

A.1 Abstract

The goal of this study was to examine the behavioral effects and possible underlying mechanism of binocularity on audiovisual temporal perception in normally sighted individuals. Participants performed two audiovisual simultaneity judgment tasks, one using simple flashes and beeps and the other using audiovisual speech stimuli. Each participant performed the task with the left eye, right eye and both eyes in separate, randomized blocks. Two measures, which were the point of subjective simultaneity (PSS) and the temporal binding window (TBW), an index for audiovisual temporal acuity, were derived for each viewing condition, stimulus type and participant. The data were then modeled using causal inference, allowing estimates to be made of the level of sensory noise affecting audiovisual simultaneity perception for each viewing condition, stimulus type and participant. While for the PSS, a two-way repeated measures ANOVA with Greenhouse-Geisser correction revealed no significant effect of viewing condition (p=0.695), for the TBW, a significant interaction between stimulus type and viewing condition (p=0.04) was found. Post-hoc

simple effects analysis with Bonferroni correction revealed a significant difference between binocular and monocular viewing (average of left and right eyes) for the flash-beep stimuli (p=0.001) but no difference for the speech stimuli (p=0.698). Modeling results showed a significant reduction in sensory noise during binocular performance on flash-beep trials (p=0.025). Binocular viewing was found to enhance audiovisual temporal acuity as indexed by the TBW for simple low-level audiovisual stimuli, and modeling results suggest that this effect may stem from a reduction in sensory noise.

A.2 Introduction

A fundamental component of human vision is the combination of the signals received separately from the two eyes into a single image (Blake and Fox, 1973; Blake, Sloane, and Fox, 1981). Besides stereopsis and a widened field of view, using two eyes compared to one often yields improved performance on a number of measures, a phenomenon termed binocular summation; see detailed reviews by Blake and Fox (1973); (Blake et al., 1981). These summation effects are seen on tasks using both threshold (i.e., contrast detection) (Baker, Lygo, Meese, and Georgeson, 2018; Legge, 1984a, 1984b; Meese, Georgeson, and Baker, 2006) and suprathreshold stimuli (i.e., contrast discrimination), (Georgeson, Meese, and Baker, 2007; Legge, 1984a) Vernier acuity, (Banton and Levi, 1991) visual acuity, (Cagenello, Arditi, and Halpern, 1993; Home, 1978) reaction times (Blake, Martens, and Di Gianfilippo, 1980; Westendorf and Blake, 1988; Yehezkel, Sterkin, Sagi, and Polat, 2015) etc.). Collectively, these psychophysical studies have revealed that using two eyes compared one eye can result in performance improvements ranging from 30 to 70%. In addition, evidence from electrophysiological studies in humans has shown that binocular viewing elicits evoked potentials of approximately 25% greater amplitude when compared with monocular viewing (Harter, Seiple, and Salmon, 1973; Pardhan, Gilchrist, Douthwaite, and Yap, 1990).

While most work is consistent with the general finding of binocular summation, the magnitude of summation differs across studies, and can even include instances where binocular viewing results in poorer performance or lower evoked potential amplitudes compared to that of one eye (Curtis and Rule, 1978; Levelt, 1965; Pardhan et al., 1990). Factors such as task and stimulus characteristics, individual differences as well as differences in monocular performance can influence the magnitude of binocular summation (Baker et al., 2018; Frisén and Lindblom, 1988; Pardhan, 1996; Pardhan and Gilchristt, 1990). For example, Frisén and Lindblom (1988) discovered that binocular summation was relatively high (resulting in performance gains of about 40%) for tasks with low stimulus complexity (i.e., differential light sensitivity of target luminance) and non-existent for tasks with high stimulus complexity (i.e., pattern recognition of digits against a random checkboard background). Among clinical populations, such as patients with amblyopia, a neurodevelopmental disorder of the visual system associated with disrupted binocular vision (Birch, 2013; Hamm, Black, Dai, and Thompson, 2014; D. M. Levi, Knill, and Bavelier, 2015), studies have reported reduced magnitude of binocular summation compared to age-matched controls (Dorr et al., 2019; Pardhan and Gilchrist, 1992; Thompson et al., 2011).

While binocular summation has been well studied for a variety of visual tasks, the study of the effects of binocularity on tasks that involve the interaction of visual and non-visual stimuli (i.e., multisensory tasks) has received much less attention. Although humans are highly visual, a large number of real-world events are multisensory, giving rise to information that concurrently stimulates multiple senses. In fact, there is mounting evidence that supports the view that multisensory processing (i.e., the interaction and or integration of information from multiple senses) may be a ubiquitous operation in the brain occurring at various levels of sensory processing hierarchies, including areas once considered classical unisensory processing hubs (Driver and Noesselt, 2008; Ghazanfar and Schroeder, 2006). The integration of multisensory information has both neural and perceptual consequences (Stein and Meredith, 1993; Stein and Stanford, 2008). At the neural level, studies have reported increased spiking activity of neurons in response to stimulus combinations (with responses that can exceed the simple summation of unisensory spiking responses), while at the perceptual level,(Stein and Meredith, 1993; Stein and Stanford, 2008) multisensory integration has been shown to increase performance in detection, discrimination, localization and reaction time tasks (Diederich and Colonius, 2004; Ernst and Banks, 2002; Frassinetti, Bolognini, and Làdavas, 2002; Lovelace, Stein, and Wallace, 2003; Zou, Müller, and Shi, 2012).

One of the key facets of this multisensory integration is the determination as to which signals arose from the same source. Important information about which stimuli should be integrated or bound is found in some of the low-level features of the multisensory pairing, such as their spatial and temporal coincidence (Murray and Wallace, 2011; Stein and Meredith, 1993). For example, in the temporal realm, sensory signals generated by the same event are likely to arrive at the sensory organs in close temporal proximity and hence, and this proximity represents a powerful statistical cue as to the likelihood that the signals originated from the same event.

Psychophysically, a number of studies have focused on understanding how the brain deals with multisensory temporal factors using simultaneity judgment (SJ) tasks (Zampini, Guest, Shore, and Spence, 2005). In a typical SJ task, participants are presented with paired multisensory stimuli (such as a visual flash and an auditory beep) with varying stimulus onset asynchronies (SOAs) and are asked to determine whether the stimulus pair was "synchronous" or "asynchronous". In other multisensory temporal tasks, subjects are asked to make temporal order judgements (TOJ) as to which stimulus of the multisensory pairing appeared first (Zampini, Shore, and Spence, 2003). Participant's reports of synchrony across the various SOAs can be used to create response distributions and allow the derivation of two important measures of multisensory temporal function - namely the point of subjective simultaneity (PSS) and the temporal binding window (TBW). The PSS is defined as the SOA at which perceived simultaneity is maximal. Interestingly, the PSS is not always at objective simultaneity (i.e., zero) but is usually found on the visual-leading side of the response distributions; see more discussion, Murray and Wallace (2011). In addition, as opposed to being a fixed construct, the PSS tends to vary dependent upon a variety of factors. These factors can be stimulus related (such as stimulus duration and intensity) (Boenke, Deliano, and Ohl, 2009; Jaśkowski, 1999; Sanford, 1971), task related (such as judging the onset vs the offset in an SJ task) (Wen, Opoku-Baah, Park, and Blake, 2020), or attention related (such as being asked to attend to one modality) (Schneider and Bavelier, 2003; Stelmach and Herdman, 1991; Zampini, Shore, and Spence, 2005). On the other hand, the TBW is the range of stimulus onset asynchronies within which two stimuli are likely to be perceptually bound or integrated (Wallace and Stevenson, 2014), thus serving as a proxy measure for multisensory temporal acuity. Like the PSS, the TBW is modulated by stimulus-related factors such as effectiveness or reliability (Fister, Stevenson, Nidiffer, Barnett, and Wallace, 2016) and stimulus complexity (e.g., flash-beep versus speech) (Stevenson and Wallace, 2013).

Although the PSS and the TBW have served as key constructs for understanding audiovisual temporal perception, the fact that they are descriptive measures derived by fitting Gaussian models limits the ability to make direct connections to neural mechanisms underlying audiovisual temporal perception (Magnotti, Ma, and Beauchamp, 2013). Consequently, Magnotti et al. (2013) developed a variant of the causal inference model (see Körding et al. (2007)) in an effort to provide greater mechanistic insights into how an observer makes synchrony judgments using the temporal relationship between the multisensory cues. This model breaks the processes involved in audiovisual simultaneity perception into low-level unisensory processes involving the encoding and processing of the individual cues and higher-level multisensory processes involving the binding or integration of these multiple sensory stimuli (Magnotti et al., 2013). In the implementation of the model, the reliability of unisensory encoding is indexed by σ, which represents the level of sensory noise in the measurement of the physical asynchrony while the tendency to bind or integrate the multisensory cues is indexed by $p(c=1)$, the observer's prior probability of inferring a common cause. Thus, as σ or $p(c=1)$ increases, there is a decrease in the precision of measuring physical asynchrony or an increase in the tendency to bind the audiovisual signals, both of which result in resulting in a widening of the TBW. For example, Magnotti et al. (2013) demonstrated that decreasing the reliability of the visual information in a speech SJ task via blurring increased the level of sensory noise.

Clinically, patients with conditions such as autism, schizophrenia and amblyopia exhibit widened TBW compared to age-matched controls, suggesting that impaired multisensory temporal function may have cascading effects into domains of clinical interest (Noel, Stevenson, and Wallace, 2018; Richards, Goltz, and Wong, 2017; Wallace and Stevenson, 2014). Although these patients show a similar phenotype (i.e., widened TBW), using the causal inference model, Noel et al. (2018) demonstrated that the widened TBW in patients with autism may result from atypical priors (i.e., increased p(c=1)), while that of patients with schizophrenia may stem from a combination of atypical priors and weakened sensory representations (i.e., increased σ). In the case of amblyopia, there still remains questions about whether the widened TBW is due to impaired binocular vision (i.e., deficits in formation of sensory representations) or impaired multisensory integration (i.e., deficits in priors) which could occur due to abnormal visual experience during development (Carriere et al., 2007; Wallace, Perrault, Hairston, and Stein, 2004).

Hence, our goal in this study was to understand the effect of binocularity on audiovisual temporal perception in normally sighted individuals. Specifically, our objective was to determine whether binocular viewing could affect audiovisual temporal perception as indexed via the PSS and TBW. Moreover, we were interested in determining whether differences in monocular vs. binocular viewing were dependent on the nature of the stimuli used in the task, and thus employed both simple low-level stimuli (i.e., flashes and beeps) and complex higher-level stimuli (i.e., speech). Lastly, we employed the causal inference model to determine whether binocular viewing influenced the level of sensory noise affecting the measurement of physical asynchrony in SJ task. Based on evidence from prior studies, we established several hypotheses. First, we hypothesized that binocular viewing would shift the PSS towards the auditory leading side (signifying more visual-biased responses) and reduce the size of the TBW (signifying improved audiovisual temporal acuity). This hypothesis was based on the well-established fact that binocular viewing enhances perceived stimulus intensity (Georgeson et al., 2007; Legge and Rubin, 1981; Levelt, 1965) and the fact that increasing intensity of the visual stimulus in an SJ task shifts the PSS toward the visual leading side and reduces the TBW. Second, we hypothesized that the effects of binocular viewing on these measures would be greater for the simple flash-beep stimuli when compared with the speech stimuli based on prior evidence that binocular summation tend to decrease with increasing stimulus complexity. Lastly, given the fact that binocular viewing enhances stimulus reliability, we hypothesized that binocular viewing will reduce sensory noise estimated using the causal inference model. Importantly, the findings of this study would contribute to the understanding of the effects of binocular vision and to some degree, visual processes on multisensory perception.

A.3 Methods

A.3.1 Participants

Nineteen participants (Male 5, age (mean ± SD): 19.8 ± 1.7 years) performed audiovisual SJ tasks with the flash-beep stimuli and with the speech stimuli and were compensated with either gift cards or course credits. All participants presented normal or corrected-to-normal vision, normal binocular vision and normal hearing. Normal vision was defined as both eyes having a visual acuity better than 20/30 while binocular vision was defined as stereo acuity better than 60 arc-seconds. Visual acuity and stereoacuity measurements were made using a Snellen chart at 6m and a Randot stereo chart respectively. Each participant gave an informed consent before being allowed to participate. All recruitment and experimental procedures were approved by the Vanderbilt University Institutional Review Board (IRB) and were carried out in accordance with the Declaration of Helsinki. Four participants were excluded from further analysis due to high proportions of synchrony reports for high SOA values in one or more experiments.

A.3.2 Stimulus and Apparatus

All experimental procedures for both the flash-beep and speech SJ tasks took place inside a dimly lit WhisperRoom^{TM} (SE 2000 Series). The visual stimuli for both stimulus types were displayed on a gamma-corrected monitor (21-inch Asus LCD) with 120-Hz refresh rate while the auditory stimuli were presented binaurally through headphones (Sennheiser HD559). For the flash-beep task, the visual stimulus was a white annular ring with an outer and inner diameter of 6 and 3 degrees respectively. The ring was displayed at the center of fixation and at 50cd/m2 luminance on a screen with luminance of 10cd/m^2. The auditory stimulus was an 1800Hz brief tone presented at ~70dB. While the visual stimulus was presented for 17ms, the auditory stimulus was presented for 10ms and was linearly ramped up and down each for 2ms. Both the visual and auditory stimuli for the flash-beep task were generated and presented using MATLAB (Math Works Inc., Natick, MA) software with the Psychophysics Toolbox Version (Brainard, 1997; Pelli, 1997). On the other hand, the stimuli for the audiovisual speech task consisted of a video of a female talker uttering the phoneme /ba/, including all prearticulatory movements, with a pixel resolution of 1920*1080 and a duration of ~2300ms. The auditory component of the video was presented at ~ 70dB. All speech stimuli for SJ tasks were presented using E-Prime version 2.0.8. A Minolta Chroma Meter CS-100 and a sound level meter were used to verify the luminance and sound intensity levels respectively. The durations of all visual and auditory stimuli, as well as the SOAs, were confirmed using a Hameg 507 oscilloscope with a photovoltaic cell and microphone.

A.3.3 Procedure

Each participant completed two sessions of the flash-beep (FB) task and two sessions of the speech (SP) task, arranged in an FB-SP-FB-SP or SP-FB-SP-FB order. This order was randomized and counterbalanced across participants. In each sub-session, participants performed the task with either the left eye, the right eye or both eyes in separate, randomized blocks. During monocular viewing, the untested eye was covered with an opaque patch and after each monocular viewing block, participants took a 5-minute break in order to reduce the effects of deprivation on subsequent sessions. For both tasks, participants judged whether the visual stimulus (which was flash ring for the FB task and lip movements for the speech task) and the auditory stimulus (which was brief tone for the FB task and “/ba/" sound for the speech task) occurred at the same time or at different times. From trial to trial, the onsets of the visual and auditory stimuli were separated by a set of pre-defined SOAs (FB task: ±400, ±300, ±200, ±150, ±100, ±50 and 0; SP task: ±500, ±400, ±300, ±250, ±200, ±150, ±100 and 0) where negative and positive SOA values corresponded to auditory-preceding-vision and vision-preceding-auditory SOAs respectively. For each block, each SOA was presented 10 times in randomized fashion totaling 260 trials for each viewing condition for the FB task and 300 trials for each viewing condition for the SP task. Each trial began with a brief fixation period which lasted between 700 and 1000ms (Figure A.1). During this period, participants viewed a centrally displayed plus sign on the screen. After the fixation period, the audiovisual stimulus was presented and participants were then asked to provide their responses by pressing ‘1’ on the keyboard if the pair of audiovisual stimuli was synchronous or by pressing ‘2’, if the pair was asynchronous. Before participants began the main experiment, each was given brief initial practice sessions using the highest SOAs for each task in order to ensure task familiarization and comprehension. Participants were not provided with feedback on the correctness of their responses during the main experiment.

A.3.4 Derivation of Behavioral Measures

For each participant, we pooled responses from blocks for each viewing condition and stimulus type and then computed proportions of synchrony reports as a function of SOA using the pooled data. To determine the PSS and TBW values for each viewing condition and stimulus type, we fitted a single-term Gaussian distribution model with the amplitude, mean and standard deviation as free parameters. While the mean and standard deviation parameters ranged from negative infinity to positive infinity, the range of possible values for the amplitude parameter was bound between 0 and 1. The averaged r2 values for flash-beep task (0.92 ± 0.05) and the speech task (0.91 ± 0.06) showed reasonable fits to the data. We derived the PSS and the TBW as the mean and standard deviation of the best fitting Gaussian model respectively.

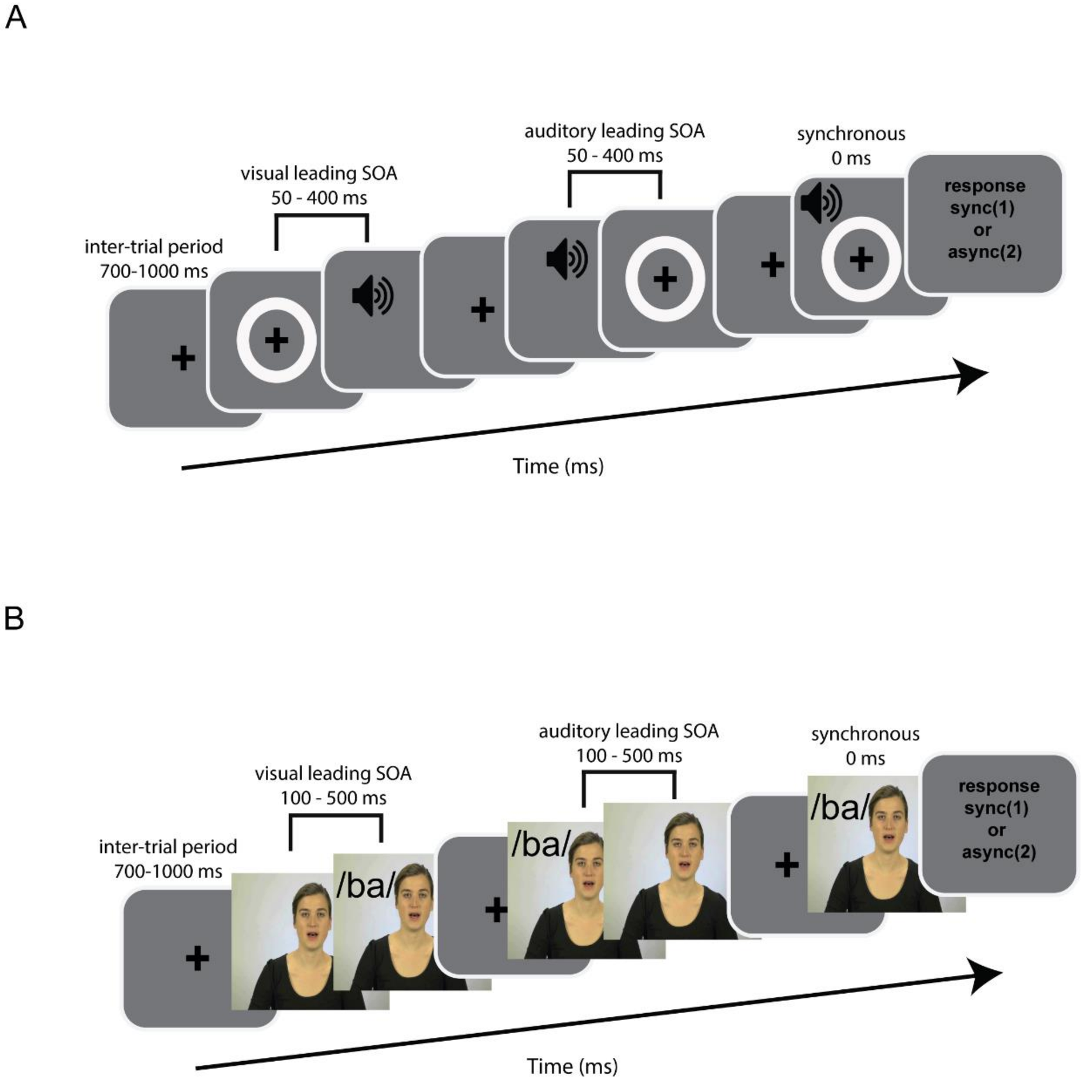

Figure A. 1 Schematic of the procedure for the (A) flash-beep SJ task and (B) speech SJ task.

Participants judged the simultaneity of a visual stimulus (flash of light (A) and lip movements (B)) and an auditory stimulus (auditory beep (A) and phoneme /ba/ (B)) presented with varying stimulus onset asynchronies. On each trial, there was a brief fixation period (700-1000ms), followed by the stimulus presentation. Participants were then asked to respond by pressing the keyboard after which the next trial began automatically.

A.3.5. Fitting the Causal Inference Model

The causal inference model provides a mechanistic understanding of how an observer makes synchrony judgments between two stimuli from different sensory modalities during the performance of an SJ task (Magnotti et al., 2013). We will point to Magnotti et al. (2013) for a more detailed derivation of this model. Moreover, while the model was originally derived using speech stimuli, in principle the model should work for other stimuli such as flash-beep employed in SJ tasks.

According to the causal inference model, the brain first infers the underlying causal structure of cues from multiple sensory modalities before combining them (Figure A.2). This underlying causal structure can be one of two possibilities which are 1) the events having a common cause (C=1) or 2) the two events having different causes (C=2). Naturally, events emanating from a common source such as auditory and visual speech results in a narrow distribution of physical asynchronies with a mean that is characteristic of the relationship between the two cues. For instance, the asynchrony distribution of audiovisual speech has a positive mean owing to the small delay between the visual and the auditory onsets. This delay stems from the fact that pre-articulatory facial movements occur before the engagement of the vocal cords during speech. In the case of non-speech stimuli, the auditory and visual stimuli most likely have similar onsets and thus, may result in an asynchrony distribution with a mean of zero. When the two events have different causes, the distribution of physical asynchronies is broad and has a mean of zero due to the lack of relationship between the cues. Furthermore, the model posits that the observer's measured asynchrony is subject to sensory noise and hence, follows a broader distribution than physical asynchrony. When these component distributions are overlaid, a window of measured asynchronies for which the probability of inferring a common cause outweighs the probability of inferring different causes emerges. This window termed the Bayes-optimal synchrony window is independent of the physical asynchrony between the cues observed and hence, represents a decisional structure used by the observer in making synchrony judgments. In its implementation, the causal inference model uses six parameters, which can be grouped into two subject parameters and four stimulus parameters. The first subject parameter is σ, which represents sensory noise that corrupts the measurement of the physical asynchrony and thus, as σ increases, there is a decrease in the precision of measuring physical asynchrony. The second subject parameter is $p_{C=1}$, which represents the prior probability of a common cause. When $p_{C=1}$ is high, there is an increased tendency to report synchrony. The stimulus parameters include the mean and standard deviation of the C=1 ($\mu_{C=1}$, $\sigma_{C=1}$) and C=2 ($\mu_{C=2}$, $\sigma_{C=2}$) distributions.

To fit the model to our data, we employed routines from source codes available freely on this website: http://openwetware.org/wiki/Beauchamp:CIMS. Following procedures in Magnotti et al. (2013), we fitted a single model for each viewing condition and for each subject for both the flash-beep stimulus condition and the speech condition. Each model had five free parameters which were the σ, $p_{C=1}$ and the three stimulus-based parameters ($\sigma_{C=1}$, $\mu_{C=2}$, $\sigma_{C=2}$); $\mu C_{=1}$ is set to zero. The parameter range for the parameters were set as follows: $p_{C=1}$ [0.01, 0.99], $\sigma_{C=1}$ [0, 400], $\mu_{C=2}$ [-200, 200] and $\sigma_{C=2}$ [150, 800]. For each subject, we determined the parameters for the best fitting model using 128 initial positions of the starting parameter values and by maximizing the binomial log-likelihood function on the observed data. The above procedure was repeated 10 times and the final set of parameters that went into further analysis was the mean of the parameter values from the 10 best fitting models.

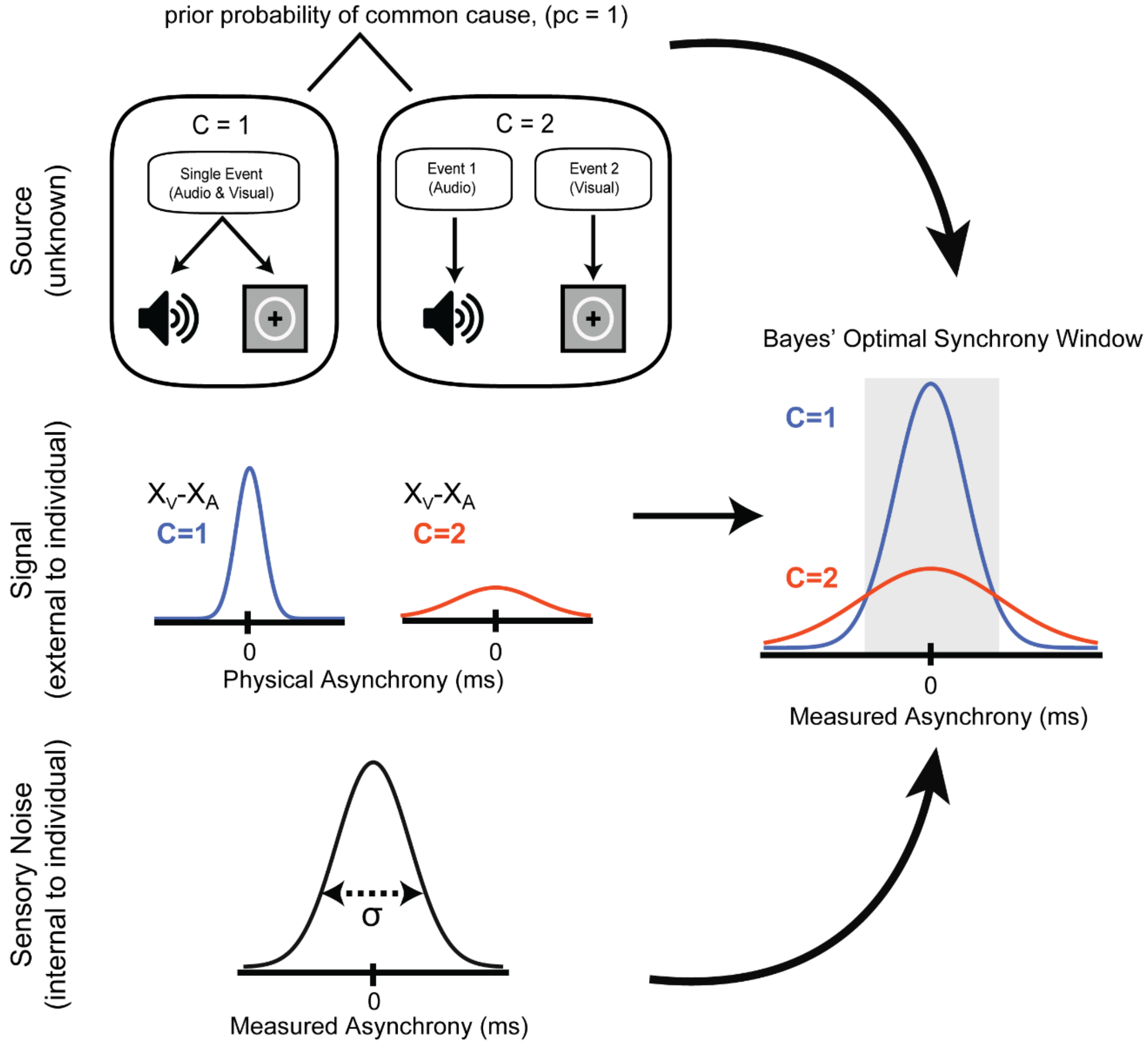

Figure A. 2 Causal inference model for audiovisual SJ tasks.

Before multiple cues are combined, the brain determines whether they originate from a common source (C=1) or different sources (C=2). Auditory and visual stimuli that share a common source have a narrow distribution of physical asynchronies (middle, blue) and a mean that suggest a relationship between the cues (e.g., positive mean for speech or zero mean for flash-beep). When the paired stimuli have different sources, the distribution is broad, and the mean is zero due (middle, red). According to the model, each participant possesses a prior tendency to bind multiple sensory information across time (pc=1, top) and samples information from the sensory world with a certain level of noisiness (sensory noise, bottom). Combining these components creates of window of measured asynchronies where the probability of inferring a common cause is more likely than that of separate causes (middle right). This window termed the Bayes' optimal window is asynchrony serves a decision structure for judging the simultaneity of these events. Figure modified from (Noel et al, 2016).

A.4 Results

We recorded synchrony judgments on two SJ tasks, one using a flash-beep stimuli and the other using speech stimuli, from 19 subjects, out of which, 4 were excluded from further analysis (see methods sections). Figure A.3 shows the mean proportions of synchrony reports plotted as a function of SOA for the binocular condition (blue) and the averaged monocular conditions (orange) for (A) the flash-beep stimulus and (B) the speech stimulus. Audiovisual temporal perception was indexed for the two stimulus types via two perceptual measures - the PSS and the TBW.

In order to determine the effect of binocularity on audiovisual temporal perception, we conducted 2 x 2 repeated measures ANOVA with Greenhouse-Geisser correction on each of the performance measures (i.e. PSS and TBW) with viewing condition (i.e., binocular vs monocular) and stimulus type (i.e., flash-beep and speech) as the within-subject factors using the JASP software version 0.11.1 (JASP Team, 2019) Here, monocular performance was defined as the averaged performances of the left and right eye conditions. We were able to pool the results for the right and left eyes since there was no statistically significant difference between them for both the PSS and TBW for both stimulus types (i.e., flash beep and speech); all $p>0.3$. All statistical analyses were two-tailed with an alpha of 0.05.

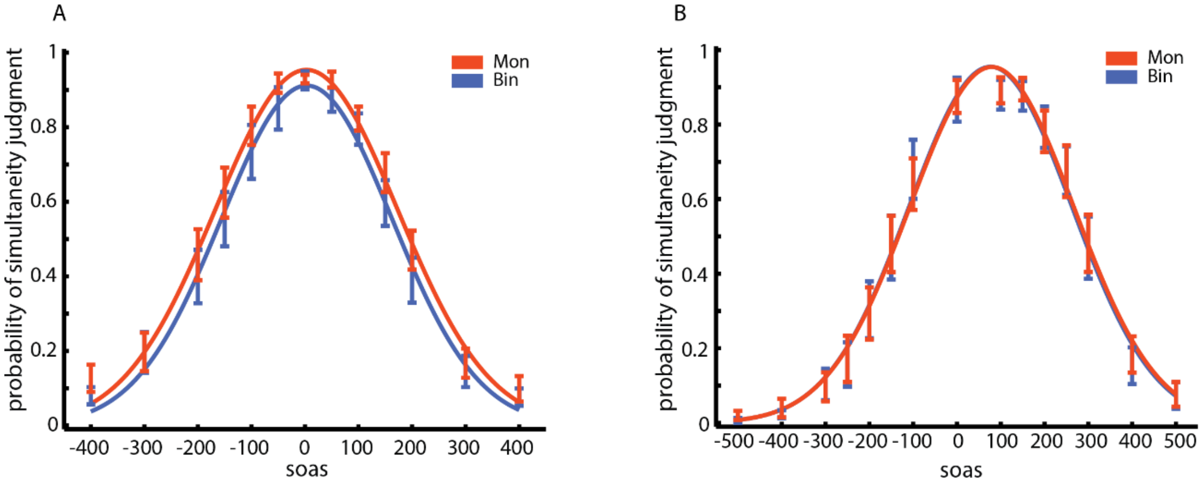

Figure A. 3 Mean proportions of synchrony reports.

Proportion of synchrony reports averaged across participants for the binocular condition (blue) and mean monocular condition (orange) is plotted as a function of SOA for (A) the flash-beep SJ task and (B) the speech SJ task.

A.4.1 Binocular viewing has no effect on PSS for SJ tasks using either flash-beep or speech stimuli

For the PSS, a two-way repeated measures ANOVA revealed a significant main effect of stimulus type ($F(1,14) = 58.74$, $p<0.001$, $\eta^2_p = 0.81$; Figure A.4A). The PSS averaged across all viewing conditions was significantly shifted towards more positive values for the speech stimulus (85.87 ms) compared to the flash-beep stimulus (-0.51 ms). Surprisingly, our analysis showed no significant effect of viewing condition ($F(1,14) = 0.16$, $p=0.695$, $\eta^2_p = 0.011$) and no significant interaction between viewing condition and stimulus type ($F(1,14) = 0.113$, $p= 0.742$, $\eta^2_p = 0.008$), indicating no effect of binocular viewing on the PSS for either stimulus type.

A.4.2 Binocular viewing enhances audiovisual temporal acuity for SJ tasks using flash-beep stimuli but not speech stimuli.

Consistent with our hypothesis, a two-way repeated measures ANOVA conducted on the TBW revealed a significant effect of stimulus type ($F(1,14) = 6.08$, $p=0.027$, $\eta^2_p = 0.303$), a significant effect of viewing condition ($F(1,14) = 8.13$, $p=0.013$, $\eta^2_p = 0.367$) and a significant interaction between stimulus type and viewing condition ($F(1,14) = 5.13$, $p=0.04$, $\eta^2_p = 0.268$; Figure A.4B). To further investigate the dependence of this TBW difference on stimulus type, we conducted a post-hoc simple effects analysis with Bonferroni correction on the ANOVA results. Our analysis revealed that for the flash-beep stimulus, the TBW for binocular viewing was significantly narrower than that for monocular viewing ($t(14) = -4.14$, $p = 0.001$, $d = -1.069$, adjusted alpha = 0.025). In contrast, for the speech stimulus there was no significant difference between the TBW for binocular viewing and that for monocular viewing (t(14) = -0.4, p = 0.698, d = -0.102, adjusted alpha = 0.025).

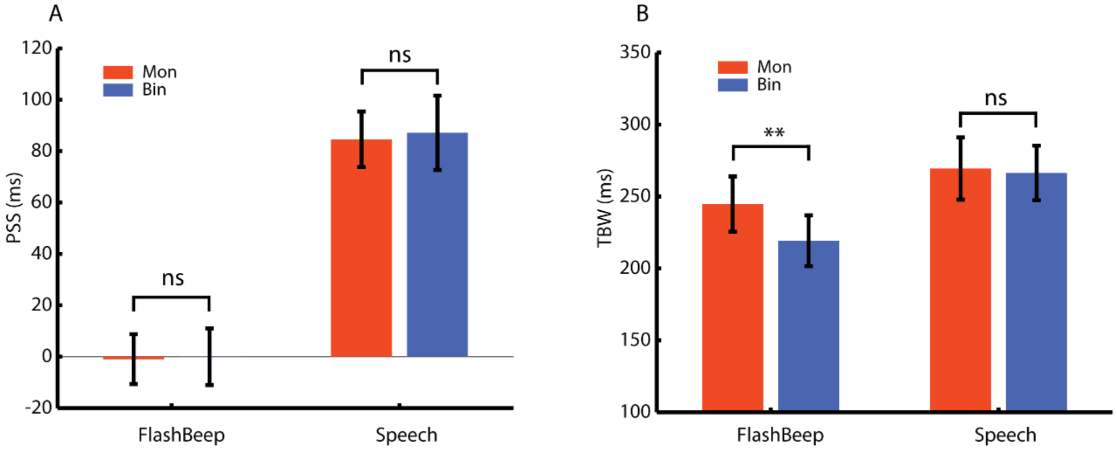

Figure A. 4 Effects of viewing condition and stimulus type on PSS and TBW.

Mean PSS (A) and TBW (B) results plotted for task and viewing conditions (binocular (blue) and mean monocular (orange). The error bars represent ±SEM.

A.4.3 Binocular enhancement in audiovisual temporal acuity could be explained by a reduction in sensory noise

In order to provide more mechanistic insights into the binocular enhancement observed for the flash-beep stimuli, we employed the causal inference model developed by Magnotti et al. (2013). As described earlier, this model provides a first-principles analysis of how the temporal relationship between cues can be leveraged to determine whether these cues originate from a common source (C=1) or different sources (C=2). The model accomplishes this using six parameters including a sensory noise parameter which is a proxy for reliability of unisensory encoding or the level of noisiness in the formation of sensory representations.

We took this approach based on the following rationale. First, the stimuli that were presented across the viewing conditions shared the same characteristics and therefore, we expected no difference in the stimulus-based parameters across the conditions. Second, because binocular integration is predominantly a low-level visual phenomenon, we assumed that binocular viewing would most likely affect the sensory noise parameter more than the participant's prior probability of inferring a common cause.

Because our hypothesis was that binocular viewing might affect sensory reliability, we fitted a single model for all three viewing conditions. This model comprised a separate σ for each condition with all remaining parameters constrained across the viewing conditions. We then determined values of the sensory noise parameter from best fitting models on the data for all viewing conditions. A paired t-test was conducted between these values for the binocular and monocular conditions where the monocular condition was defined as the mean of the left and right eye values. From our findings, there was a statistically significant reduction in sensory noise during binocular viewing (t(14) = -2.51, p = 0.025, d = -0.648; Figure A.5A). Moreover, our results showed that this difference in sensory noise was strongly and positively correlated with the difference observed in the TBW (Pearson's r = 0.78, p < 0.001; Figure A.5B). These findings indicate that the effect of binocular viewing on audiovisual temporal perception observed for the flash-beep stimuli may stem from a reduction in sensory noise.

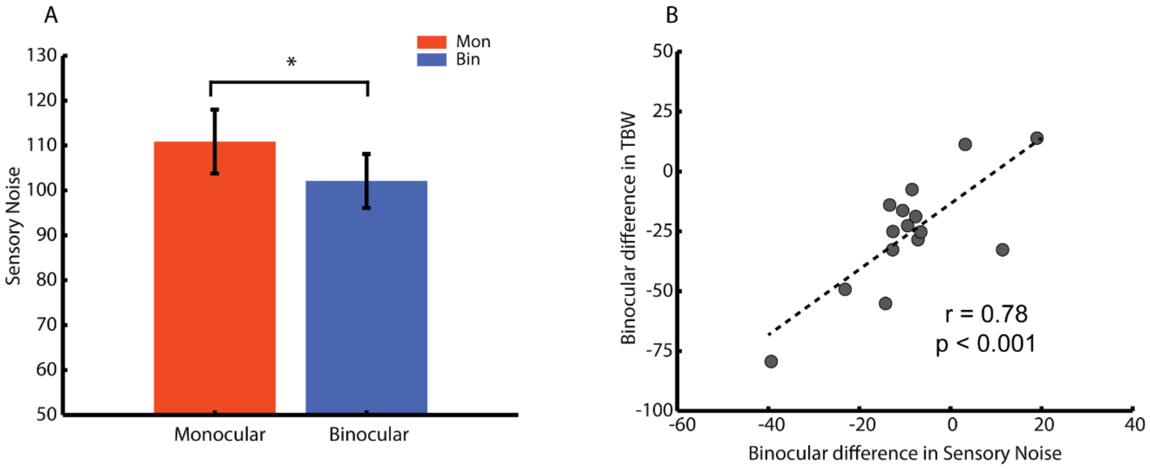

Figure A. 5 The effects of viewing condition of sensory noise for the flash-beep SJ task.

(A) Mean values of the sensory noise parameter plotted for binocular (blue) and averaged monocular (orange) conditions for the flash-beep SJ task. The error bars represent ±SEM. (B) Pearson's correlation between the binocular-monocular difference in TBW and binocular-monocular difference in sensory noise parameter.

A.5 Discussion

This study provides the first clear evidence of binocular summation in audiovisual temporal perception in normally sighted individuals. The key finding was that audiovisual temporal acuity, as indexed by the TBW, was improved under binocular viewing conditions. Consistent with prior studies, this benefit was only seen when low-level audiovisual stimuli were used and was absent with the use of audiovisual speech stimuli. Causal inference modeling suggests that the binocular benefit was a result of a reduction in sensory noise affecting the measurement of physical asynchrony during audiovisual temporal perception

Although our study investigated binocular summation using a multisensory (i.e., audiovisual) task, our finding that binocular viewing enhances audiovisual temporal acuity is in line with studies that have reported binocular summation in several suprathreshold visual tasks such as contrast and orientation discrimination tasks, visual and Vernier acuity tasks and reaction times tasks (Baker et al., 2018; Banton and Levi, 1991; Blake et al., 1980; Cagenello et al., 1993; Georgeson et al., 2007; Home, 1978; Legge, 1984a, 1984b; Meese et al., 2006; Westendorf and Blake, 1988; Yehezkel et al., 2015). Previous physiologically plausible models explaining these findings of binocular summation in visual tasks (especially using contrast and luminance detection and discrimination tasks) posited that the inputs from the corresponding retinal points in the two eyes are linearly transduced before they undergo binocular summation and finally, suppressive ocular interactions, mostly in the primary visual cortex (Legge, 1984b). However, recent work challenges this framework and demonstrates that models that include suppressive ocular interactions before summation may provide better fits and explanation to these findings of binocular summation (Georgeson et al., 2007; Meese et al., 2006).

In the case of audiovisual temporal perception, studies have benefitted from Bayesian modelling approaches including the causal inference model applied in this study (Ernst and Banks, 2002; Körding et al., 2007; Magnotti et al., 2013). Generally, these models comprise parameters that index processes occurring at the unisensory level and those that involve the binding and or the integration of multisensory cues. Considering the fact that binocular integration is a low-level visual phenomenon occurring predominantly in the primary visual cortex (Blake and Fox, 1973), we believe that the role binocular integration plays in audiovisual simultaneity perception can be explained by considering the summation of the luminance energies of the suprathreshold visual stimuli received from the two eyes prior to multisensory integration. Following the evidence that binocular viewing enhances perceived stimulus intensity, our finding of binocular summation of audiovisual temporal acuity (i.e., reduction in the TBW) for the flash-beep task fits studies that have demonstrated that increasing the effectiveness of the stimuli in an SJ task improves audiovisual temporal acuity (Fister et al., 2016; Magnotti et al., 2013). For instance, Fister et al. (2016) investigated the effect of increasing stimulus intensity on the probability of making synchrony judgments for visual-leading SOAs in an SJ task. They discovered that as SOA increased, the probability of making synchrony judgments fell more rapidly for the highly effective stimuli than for the lowly effective stimuli. This finding implied that increasing the effectiveness of the stimuli decreased the tolerance for audiovisual asynchrony, which manifests as a narrowing of the TBW.

Using the causal inference model, our study showed that the binocular enhancement in the TBW observed for the flash-beep could be explained by a reduction in the level of sensory noise affecting the observer’s judgment of asynchrony. Indeed, this finding agrees with

Magnotti et al. (2013) study that demonstrated that manipulating stimulus reliability affects the noisiness in the formation of sensory representations, parameterized in the causal inference model as sensory noise, σ. Specifically, Magnotti et al. (2013) showed that when the reliability of the visual stimulus during the performance of an SJ task was decreased through blurring, there was an increase in the level of sensory noise (σ) affecting the judgment. Fitting Gaussian models to the data showed that the non-blurry stimulus condition (i.e., more reliable) had a narrower TBW, in concordance with our results where the binocular viewing condition decreased the TBW for the flash-beep stimuli. Besides SJ tasks, Beierholm, Quartz, and Shams (2009) applied the causal inference model to an audiovisual spatial localization task and showed that high contrast stimuli decreased the standard deviation of visual likelihood parameter signifying decreased noisiness in visual sensory representations. Although the two models (i.e., Magnotti et al. (2013) and Beierholm et al. (2009)) were developed for different problems (i.e., audiovisual speech perception and audiovisual spatial localization respectively), Magnotti et al. (2013) highlighted that both problems are mathematically similar and that, the models share the same theoretical framework. Hence, it is plausible to conclude that the sensory noise parameter in Magnotti et al. (2013) model and the standard deviation of visual likelihood in Beierholm et al. (2009) model serve a similar function, as both relate to the noisiness in sensory representations.

Although the causal inference model is able to differentiate between the contributions of low-level unisensory mechanisms (i.e., level of sensory noise) and high-level multisensory mechanisms (i.e., prior probability of inferring common cause) to changes in audiovisual temporal perception, when it comes to the unisensory mechanisms, it does not provide any insight into the type of sensory noise (i.e., whether internal or external) driving these changes. Moreover, the causal inference model does not make explicit assumptions about the sources of the sensory noise. However, considering the nature of the audiovisual temporal paradigm, it is plausible to hypothesize that the estimation of the physical synchrony using the visual and auditory cues may be based on the reliability of the binocular and the binaural outputs. Consequently, this may suggest that the source of the sensory noise in the model is found after binocular and binaural integration. Blake and colleagues (Blake and Fox, 1973; Blake et al., 1981) discussed the plausibility of a model with late-stage noise. Nevertheless, we believe that these details about the types and sources of noise should be incorporated into future developments of this model to facilitate the understanding of the different sensory noise mechanisms affecting audiovisual temporal perception

While binocular viewing reduced the TBW for the flash-beep stimuli, it did not affect the TBW for the speech task. Based on prior studies, there are several possible explanations. First, prior work has shown that binocular summation is more likely to occur for tasks or stimuli with low complexity (i.e., differential light sensitivity of target luminance) as opposed to those with high complexity (i.e., pattern recognition of digits against a random checkboard background). Indeed, Frisén and Lindblom (1988) posited that the more complex the stimuli, the higher the level of cortical processing required and the smaller the magnitude of binocular summation. Second, the lack of binocular summation for the speech stimuli could be explained by studies that have shown that stimuli with higher energy (i.e., luminance or contrast) yield less binocular summation. For example, Home (1978) showed that for a pattern recognition task, binocular summation was high for low target contrasts and absent at higher contrasts. Additionally, the dependence of binocular summation on stimulus contrast has been demonstrated for discrimination tasks of contrast (Georgeson et al., 2007; Legge, 1984a), orientation (Bearse Jr and Freeman, 1994) and Vernier acuity (Banton and Levi, 1991). Thus, the lack of binocular summation for the speech stimulus may stem from high stimulus complexity and/or high stimulus reliability. If the latter is true, then it can be hypothesized that the TBW of the speech stimulus may benefit from binocular enhancement if the reliability of the stimulus is reduced through blurring or addition of noise.

Considering the evidence that binocular viewing can increase the neural response and perceived intensity of viewed visual targets, our finding of no effect of viewing condition on the PSS contradicts studies that have shown that increasing stimulus effectiveness affects the PSS (Boenke et al., 2009; Smith, 1933). For example, Boenke et al. (2009) revealed that increasing the intensity of the visual stimulus in a TOJ task (a variant of SJ task) significantly shifted the PSS towards the auditory leading side, in other words, maximum perceived simultaneity was achieved with a stimulus pair of larger auditory-lead under increased visual intensity. However, in order to explain the seeming discrepancy here, it is essential that we consider how amenable the PSS is to changes in stimulus intensity assuming all other factors remain constant. For instance, in the Boenke et al. (2009) study, increasing the intensity of the visual stimulus by approximately five-fold shifted the PSS by 27ms to the left (i.e., towards the auditory leading side). Based on this analysis, one would expect that for binocular viewing, which enhances perceived brightness by approximately 40%, there would be a shift in the PSS of only 2ms assuming a linear relationship between PSS and stimulus intensity. In fact, the lack of PSS shift under binocular viewing is consistent with studies that have assessed the impact on audiovisual temporal perception by visual phenomena that modulate perceived stimulus effectiveness. For example, Opoku-Baah and Wallace (2020) showed that a brief period of monocular deprivation, a phenomenon known to boost perceived contrast in the deprived eye, did not significantly affect the PSS, although changes in the TBW were observed.

Importantly, we believe our findings have clinical implications for understanding the underlying mechanisms of the multisensory perceptual deficits observed in patients with impaired binocular vision such as in amblyopia. Several studies have shown that patients with amblyopia suffer several visual deficits including reduced visual acuity, reduced stereopsis (D. Levi and Harwerth, 1977; D. M. Levi et al., 2015; D. M. Levi, Waugh, and Beard, 1994) and even deficits in higher-level perceptual functions such as global shape detection (Hess, Wang, Demanins, Wilkinson, and Wilson, 1999), motion processing (Simmers, Ledgeway, Hess, and McGraw, 2003) and real-world scene perception (Mirabella, Hay, and Wong, 2011). Recently, amblyopia has been linked with deficits in audiovisual integration (Chen, Lewis, Shore, and Maurer, 2017; Narinesingh, Goltz, and Wong, 2017; Richards et al., 2017). For instance, Narinesingh et al. (2017) showed that adult patients with amblyopia exhibited reduced susceptibility to the McGurk effect compared to age-matched controls. In regard to audiovisual temporal perception, Richards et al. (2017) demonstrated that amblyopes compared to age-matched controls exhibited significantly widened TBW but no difference in the PSS when tested on an SJ task with the flash-beep stimuli. Using a subset of six amblyopes, they also showed that the size of the TBW was not different across viewing conditions, which were binocular, better eye and amblyopic eye conditions (Richards et al., 2017). Interestingly, while the widened TBW observed in amblyopes indicates impaired multisensory temporal integration, the absence of an effect of viewing condition on the TBW measured in amblyopes and the finding of binocular enhancement in audiovisual simultaneity perception in normally sighted individuals provided by this study suggests a possible role of impaired binocular vision in the observed multisensory deficits in amblyopia. These suggestions warrant further studies geared at understanding the relative contributions of impaired binocular vision and impaired multisensory integration to the observed deficits in multisensory temporal function. It will also be interesting to investigate how the relative contributions of these mechanisms differ based on factors such as amblyopia severity and etiology. Furthermore, we believe that the causal inference model as applied in our study will be a useful tool in providing an interesting picture of whether the deficits in audiovisual temporal perception observed in amblyopia stem from impaired binocular vision (formation of sensory representations) and or impaired multisensory processing (prior probability of inferring a common cause, also known as the binding tendency). Such a finding will help inform whether multisensory perceptual training paradigms should be developed to target these mechanisms separately in the management of amblyopia. Together, these studies will enrich our understanding of the overall sensory and perceptual deficits in amblyopia and their underlying mechanisms and enable the development of behavioral therapies that address these mechanisms.

A.6 References

Baker, D. H., Lygo, F. A., Meese, T. S., and Georgeson, M. A. (2018). Binocular summation revisited: Beyond√ 2. *Psychological bulletin, 144*(11), 1186.

Banton, T., and Levi, D. M. (1991). Binocular summation in vernier acuity. *JOSA A, 8*(4), 673-680.

Bearse Jr, M. A., and Freeman, R. D. (1994). Binocular summation in orientation discrimination depends on stimulus contrast and duration. *Vision research, 34*(1), 19-29.

Beierholm, U. R., Quartz, S. R., and Shams, L. (2009). Bayesian priors are encoded independently from likelihoods in human multisensory perception. *Journal of Vision, 9*(5), 23-23.

Birch, E. E. (2013). Amblyopia and binocular vision. *Progress in retinal and eye research, 33*, 67-84.

Blake, R., and Fox, R. (1973). The psychophysical inquiry into binocular summation. *Perception and Psychophysics, 14*(1), 161-185.

Blake, R., Martens, W., and Di Gianfilippo, A. (1980). Reaction time as a measure of binocular interaction in human vision. *Investigative ophthalmology and visual science, 19*(8), 930-941.

Blake, R., Sloane, M., and Fox, R. (1981). Further developments in binocular summation. *Perception and Psychophysics, 30*(3), 266-276.

Boenke, L. T., Deliano, M., and Ohl, F. W. (2009). Stimulus duration influences perceived simultaneity in audiovisual temporal-order judgment. *Experimental brain research, 198*(2-3), 233-244.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision, 10*(4), 433-436.

Cagenello, R., Arditi, A., and Halpern, D. L. (1993). Binocular enhancement of visual acuity. *JOSA A, 10*(8), 1841-1848.

Carriere, B. N., Royal, D. W., Perrault, T. J., Morrison, S. P., Vaughan, J. W., Stein, B. E., and Wallace, M. T. (2007). Visual deprivation alters the development of cortical multisensory integration. *Journal of neurophysiology, 98*(5), 2858-2867.

Chen, Y.-C., Lewis, T. L., Shore, D. I., and Maurer, D. (2017). Early binocular input is critical for development of audiovisual but not visuotactile simultaneity perception. *Current Biology, 27*(4), 583-589.

Curtis, D. W., and Rule, S. J. (1978). Binocular processing of brightness information: A vector-sum model. *Journal of Experimental Psychology: Human Perception and Performance, 4*(1), 132.

Diederich, A., and Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Perception and Psychophysics, 66*(8), 1388-1404.

Dorr, M., Kwon, M., Lesmes, L. A., Miller, A., Kazlas, M., Chan, K., . . . Bex, P. J. (2019). Binocular summation and suppression of contrast sensitivity in strabismus, fusion and amblyopia. *Frontiers in human neuroscience, 13*, 234.

Driver, J., and Noesselt, T. (2008). Multisensory interplay reveals cross-modal influences on ‘sensory-specific’brain regions, neural responses, and judgments. *Neuron, 57*(1), 11-23.

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

Fister, J. K., Stevenson, R. A., Nidiffer, A. R., Barnett, Z. P., and Wallace, M. T. (2016). Stimulus intensity modulates multisensory temporal processing. *Neuropsychologia, 88*, 92-100.

Frassinetti, F., Bolognini, N., and Làdavas, E. (2002). Enhancement of visual perception by cross-modal visuo-auditory interaction. *Experimental brain research, 147*(3), 332-343.

Frisén, L., and Lindblom, B. (1988). Binocular summation in humans: evidence for a hierarchic model. *The Journal of physiology, 402*(1), 773-782.

Georgeson, M., Meese, T., and Baker, D. (2007). Binocular interaction: Contrast matching and contrast discrimination are predicted by the same model. *Spatial vision, 20*(5), 397-413.

Ghazanfar, A. A., and Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in cognitive sciences, 10*(6), 278-285.

Hamm, L. M., Black, J., Dai, S., and Thompson, B. (2014). Global processing in amblyopia: a review. *Frontiers in psychology, 5*, 583.

Harter, M. R., Seiple, W. H., and Salmon, L. (1973). Binocular summation of visually evoked responses to pattern stimuli in humans. *Vision research, 13*(8), 1433-1446.

Hess, R. F., Wang, Y.-Z., Demanins, R., Wilkinson, F., and Wilson, H. R. (1999). A deficit in strabismic amblyopia for global shape detection. *Vision research, 39*(5), 901-914.

Home, R. (1978). Binocular summation: a study of contrast sensitivity, visual acuity and recognition. *Vision research, 18*(5), 579-585.

Jaśkowski, P. (1999). Reaction time and temporal-order judgment as measures of perceptual latency: The problem of dissociations *Advances in psychology* (Vol. 129, pp. 265-282): Elsevier.

JASP Team. (2019). JASP (Version 0.11. 1)[Computer software]. *JASP Team: Amsterdam, Netherlands*.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., and Shams, L. (2007). Causal inference in multisensory perception. *PLoS One, 2*(9), e943.

Legge, G. E. (1984a). Binocular contrast summation—I. Detection and discrimination. *Vision research, 24*(4), 373-383.

Legge, G. E. (1984b). Binocular contrast summation—II. Quadratic summation. *Vision research, 24*(4), 385-394.

Legge, G. E., and Rubin, G. S. (1981). Binocular interactions in suprathreshold contrast perception. *Perception and Psychophysics, 30*(1), 49-61.

Levelt, W. J. (1965). Binocular brightness averaging and contour information. *British journal of psychology, 56*(1), 1-13.

Levi, D., and Harwerth, R. S. (1977). Spatio-temporal interactions in anisometropic and strabismic amblyopia. *Investigative ophthalmology and visual science, 16*(1), 90-95.

Levi, D. M., Knill, D. C., and Bavelier, D. (2015). Stereopsis and amblyopia: A mini-review. *Vision research, 114*, 17-30.

Levi, D. M., Waugh, S. J., and Beard, B. L. (1994). Spatial scale shifts in amblyopia. *Vision research, 34*(24), 3315-3333.

Lovelace, C. T., Stein, B. E., and Wallace, M. T. (2003). An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Cognitive brain research, 17*(2), 447-453.

Magnotti, J. F., Ma, W. J., and Beauchamp, M. S. (2013). Causal inference of asynchronous audiovisual speech. *Frontiers in psychology, 4*, 798.

Meese, T. S., Georgeson, M. A., and Baker, D. H. (2006). Binocular contrast vision at and above threshold. *Journal of Vision, 6*(11), 7-7.

Mirabella, G., Hay, S., and Wong, A. M. (2011). Deficits in perception of images of real-world scenes in patients with a history of amblyopia. *Archives of Ophthalmology, 129*(2), 176-183.

Murray, M. M., and Wallace, M. T. (2011). *The neural bases of multisensory processes*: CRC Press.

Narinesingh, C., Goltz, H. C., and Wong, A. M. (2017). Temporal binding window of the sound-induced flash illusion in amblyopia. *Investigative ophthalmology and visual science, 58*(3), 1442-1448.

Noel, J. P., Stevenson, R. A., and Wallace, M. T. (2018). Atypical audiovisual temporal function in autism and schizophrenia: similar phenotype, different cause. *European Journal of Neuroscience, 47*(10), 1230-1241.

Opoku-Baah, C., and Wallace, M. T. (2020). Brief period of monocular deprivation drives changes in audiovisual temporal perception. *Journal of Vision, 20*(8), 8-8.

Pardhan, S. (1996). A comparison of binocular summation in young and older patients. *Current Eye Research*, *15*(3), 315-319.

Pardhan, S., and Gilchrist, J. (1992). Binocular contrast summation and inhibition in amblyopia. *Documenta ophthalmologica, 82*(3), 239-248.

Pardhan, S., Gilchrist, J., Douthwaite, W., and Yap, M. (1990). Binocular inhibition: psychophysical and electrophysiological evidence. *Optometry and vision science, 67*(9), 688-691.

Pardhan, S., and Gilchristt, J. (1990). The effect of monocular defocus on binocular contrast sensitivity. *Ophthalmic and Physiological Optics, 10*(1), 33-36.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision, 10*(4), 437-442.

Richards, M. D., Goltz, H. C., and Wong, A. M. (2017). Alterations in audiovisual simultaneity perception in amblyopia. *PloS one*, *12*(6), e0179516.

Sanford, A. (1971). Effects of changes in the intensity of white noise on simultaneity judgements and simple reaction time. *The Quarterly Journal of Experimental Psychology, 23*(3), 296-303.

Schneider, K. A., and Bavelier, D. (2003). Components of visual prior entry. *Cognitive psychology, 47*(4), 333-366.

Simmers, A. J., Ledgeway, T., Hess, R. F., and McGraw, P. V. (2003). Deficits to global motion processing in human amblyopia. *Vision Research, 43*(6), 729-738.

Smith, W. F. (1933). The relative quickness of visual and auditory perception. *Journal of Experimental Psychology, 16*(2), 239.

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

Stein, B. E., and Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience, 9*(4), 255-266.

Stelmach, L. B., and Herdman, C. M. (1991). Directed attention and perception of temporal order. *Journal of Experimental Psychology: Human Perception and Performance, 17*(2), 539.

Stevenson, R. A., and Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Experimental brain research*, *227*(2), 249-261.

Thompson, B., Richard, A., Churan, J., Hess, R. F., Aaen-Stockdale, C., and Pack, C. C. (2011). Impaired spatial and binocular summation for motion direction discrimination in strabismic amblyopia. *Vision research*, *51*(6), 577-584.

Wallace, M. T., Perrault, T. J., Hairston, W. D., and Stein, B. E. (2004). Visual experience is necessary for the development of multisensory integration. *Journal of Neuroscience, 24*(43), 9580-9584.

Wallace, M. T., and Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia, 64*, 105-123.

Wen, P., Opoku-Baah, C., Park, M., and Blake, R. (2020). Judging Relative Onsets and Offsets of Audiovisual Events. *Vision, 4*(1), 17.

Westendorf, D., and Blake, R. (1988). Binocular reaction times to contrast increments. *Vision research*, *28*(2), 355-359.

Yehezkel, O., Sterkin, A., Sagi, D., and Polat, U. (2015). Binocular summation of chance decisions. *Scientific reports, 5*, 16799.

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

Zampini, M., Shore, D. I., and Spence, C. (2003). Audiovisual temporal order judgments. *Experimental brain research, 152*(2), 198-210.

Zampini, M., Shore, D. I., and Spence, C. (2005). Audiovisual prior entry. *Neuroscience letters*, *381*(3), 217-222.

Zou, H., Müller, H. J., and Shi, Z. (2012). Non-spatial sounds regulate eye movements and enhance visual search. *Journal of Vision, 12*(5), 2-2.