The Nature of Sensory Time Perception – Centralised or Distributed?

By

Aysha Motala

A thesis submitted to Cardiff University

for the degree of

Doctor of Philosophy

School of Optometry and Vision Sciences, Cardiff University

2019

Abstract

Using psychophysical methods and human subjects, this work aims to investigate the role of human sensory systems in the perception and passage of time. Specifically, I question the centralised nature of timing and whether a central clock exists to mediate incoming timing signals across the different sensory modalities. The alternative is that our timing mechanisms are embodied within distributed, modality-specific networks, each operating in a dedicated and independent manner. In my first experiment subjects were exposed to a range of rhythms presented to audio, visual and tactile sensory modalities, and were asked to reproduce a test rhythm via a tapping device. Subjects were able to adapt to a range of rhythms; however, the resulting after-effects were only evidenced when the adapting and test sensory modalities matched. My second experiment questioned how we construct sensory rhythms and, using the same method of rhythm adaptation, I used a single empty interval as a test stimulus. Results show that adapting to a given rhythmic rate strongly influences the temporal perception of a single empty interval. This questions the seemingly unique nature of rhythm, suggesting that adaptive distortions in perceived rate of signals within a sequence are, at least in part, a consequence of distortions in the perception of the inter-stimulus interval between the sequence's component signals. My third experiment focused on more complicated rhythms in the form of anisochrony. I found limited observable after-effects as a result of exposing subjects to patterned rhythms across auditory, visual and tactile sensory modalities. The final experiment demonstrated significant after-effects following exposure to perfectly interleaved auditory and visual rhythms. These results collectively demonstrate mechanisms actively underpinning human perception of time and importantly, present evidence of dynamically distributed mechanisms linked to each sensory modality and processing incoming timing signals in a dedicated manner.

Acknowledgements

This thesis and my doctoral experience would not have been the same without the invaluable contributions of several individuals that I would like to take this opportunity to thank.

Firstly, my sincerest gratitude to my primary supervisor, Professor David Whitaker. It is commonly known advice that you should avoid working with your heroes lest they disappoint you. My experience with David has been the complete opposite, and I am in constant awe of his curiosity, excitement and dedication towards the scientific pursuit. To have worked with him so closely and only have my admiration and respect grow not only professionally but also personally, is a testament to David’s character (and my good fortune to have been given the chance to be supervised by him). It is these qualities that have directed my doctoral experience, and made this a dream come true, and for which I will always be indebted to David for. I would also like to thank my collaborators Dr James Heron and Professor Paul McGraw, for their continued inspiration, support and feedback throughout my PhD. I must also thank Dr Neil Roach for introducing me to the field of sensory time perception during my undergraduate years. A project that I initially undertook as a spontaneous and fun challenge has blossomed into one of my greatest passions, and this is something that was only made possible because of Neil’s guidance and patience. I would additionally like to acknowledge my second supervisor, Dr Tony Redmond for his continued moral support. Tony’s insightful guidance and encouragement to pursue all the opportunities offered to me during my PhD, has made an invaluable difference to my experience and for this I am indelibly grateful.

There are several people within the School of Optometry & Vision Sciences that have contributed to my progress and experience in Cardiff these past three years. I would particularly like to thank Sue Hobbs and the rest of the admin team for their friendship and maternal support that made Cardiff truly feel like home. I would like to thank my friend Nikki Cassells, for her enduring patience and insight and for endless coffee breaks, despite her own workload and personal commitments. Likewise, my gratitude also goes to Julie Albon for her words of wisdom and for being a continual source of encouragement and inspiration. My special thanks also go to the many undergraduates and participants for my experiments (especially those employing lengthy adaptation periods!). The contributions made by this work and the needle moving forward was only made possible due to their commitment and dedication to science. Additionally, to the many friends within the school who have come and gone along the way, and who have brightened each day.

Finally, I would like to express my deepest gratitude to my family. They have been a limitless source of encouragement, support and dependable good humour. I would like to thank my brother Mohammed, for his always-appropriate words of wisdom and for being able to lighten any situation with his good humour and positive energy. I would especially like to thank my sister Zainab, for the countless hours spent proofreading my abstracts, chapters, and manuscripts and for helping me practice many conference presentations, always without hesitation and with the utmost care and attention to detail, despite her own extensive workload and commitments.

Importantly, I would like to thank my best friend and fiancé, Klodi, for his everlasting patience and understanding. For always knowing how to brighten the darkest day and believing in me, even when I could not do so myself. Lastly, I would like to thank my mother, Gori, whose strength and selflessness has been a continual source of inspiration, and whose determination and resilience has been my greatest example. Ultimately, to whom I owe each and every achievement.

To my Mother,

Publications and Conference Presentations

Selected work in this thesis has been presented in written form in peer-reviewed journals, and also through poster and oral presentations at various national and international meetings. Details on these are provided below:

Journal Articles

Motala, A., Heron, J., McGraw, P. V., Roach, N. W., & Whitaker, D. Rate after-effects fail to transfer cross-modally: Evidence for distributed sensory timing mechanisms. Scientific Reports, 8(1), 924. (2018).

Motala, A. & Caceres, L. Disentangling neural synchronization and sustained neural activity in the processing of auditory temporal patterns. Frontiers in Human Neuroscience, 12, 497. (2018).

Motala, A., Heron, J., McGraw, P. V., Roach, N. W., & Whitaker, D. Sensory Rate Perception – Simply the sum of its parts? (In preparation).

Invited Talks

Motala, A. Sensory Time Perception - Using auditory, tactile and visual rhythms to explore theories of time and temporal processing in humans. University of Giessen, Germany, 2018

Motala, A. Perceiving time across the senses - Using auditory, tactile and visual rhythms to explore theories of time and temporal processing in humans. City University, London, 2018

Motala, A. Starting Out and Enjoying Your PhD. Doctoral Academy, Cardiff University, Wales, 2018.

Conference Presentations

Motala, A. & Whitaker, D. Rate aftereffects fail to transfer cross-modally: Evidence towards distributed timing mechanisms. British Congress of Optometry & Vision Sciences, Ulster University, N. Ireland, 2016

Motala, A. & Whitaker, D. Rate aftereffects fail to transfer cross-modally: Evidence towards distributed timing mechanisms. Applied Vision Association Christmas Meeting, Queen Mary, University of London, London, 2016

Motala, A. & Whitaker, D. Rate aftereffects fail to transfer cross-modally: Evidence towards distributed timing mechanisms. Experimental Psychology Society Timing Workshop, John Moores University, Liverpool, 2017

Motala, A. & Whitaker, D. Sensory Clocks - One or Many? Evidence towards distributed timing mechanisms. Speaking of Science Interdisciplinary Conference, Cardiff University, Wales, 2017

Motala, A. & Whitaker, D. Modality Specific Rate Aftereffects: Evidence Towards Distributed Timing Mechanisms, Bristol Young Vision Researcher's Colloquium, Bath University, Bath, 2017

Motala, A. & Whitaker, D. The rhythm aftereffect – Adapting to sensory rate? British Congress of Optometry &Vision Sciences, Plymouth University, Plymouth, 2017

Motala, A. & Whitaker, D. Modality Specific Rate Aftereffects: Evidence towards Distributed Timing Mechanisms, International Conference of the Timing Research Forum, Strasbourg, France 2017

Motala, A. & Whitaker, D. Visual Rate Perception – More than the sum of its parts? Applied Vision Association Christmas Meeting, Queen Mary, University of London, London, 2017

Motala, A. & Whitaker, D. Sensory Rate Perception – Simply the sum of its parts? International Multisensory Research Forum, University of Toronto, Canada, 2018

Motala, A. & Whitaker, D. Visual Rhythm Perception – Simply the sum of its parts? Bristol Young Vision Researcher's Colloquium, Bristol University, Bristol, 2018

Motala, A. & Whitaker, D. Adapting to crossmodal rhythms in pursuit of a central timing mechanism. British Congress of Optometry & Vision Sciences, Anglia Ruskin University, Cambridge, 2018

Table of Contents

Appendices

List of Figures

1. Introduction to Sensory Systems and Perception

2. Introduction to Time Perception and Timing Models

3. Methodology

4. Investigating Sensory Timing using Psychophysics

5. Assessing the modality-specificity of the rhythm after-effect

6. Sensory rate perception – Simply the sum of its parts?

7. Dissociating rhythm and interval discrimination through unimodal temporal pattern adaptation

8. Adapting to cross-modal rhythms in search of a centralised timer

List of Tables

5. Assessing the modality-specificity of the rhythm after-effect

6. Sensory rate perception – Simply the sum of its parts?

7. Dissociating rhythm and interval discrimination through unimodal temporal pattern adaptation

8. Adapting to cross-modal rhythms in search of a centralised timer

List of Abbreviations

LGN	Lateral Geniculate Nucleus
V1	Visual Cortex
MLE	Maximum Likelihood Estimation
SBF	Striatal Beat-Frequency
tRNS	transcranial Random Noise Stimulation
IPL	Inferior Parietal Lobule
SMG	Supra Marginal Gyrus
SMA	Supplementary Motor Area
MEG	Magnetoencephalography
SDT	Signal Detection Theory
2AFC	Two-Alternative Forced Choice
SOA	Stimulus Onset Asynchrony
PSE	Point of Subjective Equality
TOJ	Temporal Order Judgement
JND	Just Noticeable Difference
ITD	Inter-aural Time Difference
ERP	Event-Related Potential
CRF	Clinical Research Facility
A	Auditory
T	Tactile
V	Visual
AV	Audio-visual

1. Introduction to Sensory Systems and Perception

1.1 Visual Perception

Visual perception is perhaps one of the most important functions humans have evolved through time, and in an increasingly visual society, this system remains fundamental to not only our survival but also our quality of life. The visual system detects and interprets light signals to build a perceptual representation of the physical world. Anatomically, it is mediated through a system consisting of retinal photoreceptor cells, the optic nerve and optic chiasm, lateral geniculate nucleus (LGN), optic radiations, and V1 (also known as the primary visual cortex/striate cortex). Higher levels of the visual system include areas V2, V3, V4 and V5/MT in mammals. The following chapter will elaborate on these structures in more detail.

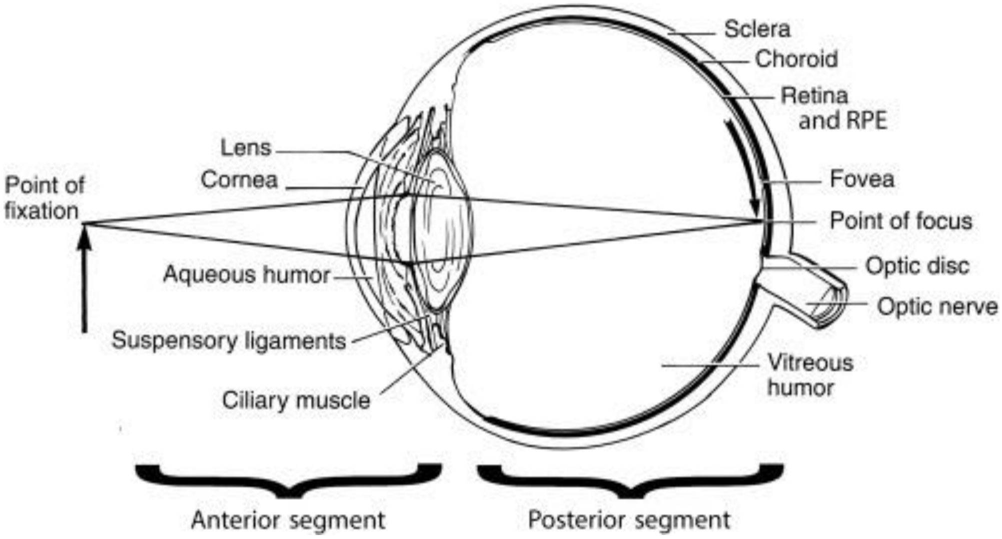


Figure 1.1. Anatomy of the eye, (Hejtmancik & Nickerson, 2015).

Light enters the eye and is refracted via the cornea (Figure 1.1). On passing through the pupil it is then further refracted by the lens and an inverted image is then projected to the retina. The retina contains a large number of receptor cells called rods and cones (and collectively referred to as photoreceptor cells). Photoreceptors remain crucial to our visual abilities as photoreceptor proteins absorb photons, activating a change in the cell's membrane potential and stimulating biological processes. This action represents the process of transduction (Goldstein, 2007). Strictly, the retina contains three different types of photoreceptor cells – rods, cones and photosensitive retinal ganglion cells. Rods and cones are understood to directly contribute information to build a representation of the world whereas photosensitive retinal ganglion cells (discovered much more recently) are thought to not directly contribute to vision but support pupillary reflexes and circadian rhythms (Berson, Dunn, & Takao, 2002). Rods are extremely sensitive to light and are therefore the driving photoreceptor in environments with low light levels (as colour vision and the contribution from cones becomes less essential). Cones, on the other hand, require a larger number of photons and therefore significantly brighter light to produce a signal. Their primary role includes responsibility for daytime vision and visual acuity (sharpness of vision as they provide us information on fine detail of our environment). Humans possess three different types of cone cells, each responsible for a different wavelength of light (of short, medium and long wavelengths) and the ability to perceive colour is deduced by evaluating these signals. It is understood that, on average, the human retina possesses 120 million rods and 6 million cones (Osterberg, 1935).

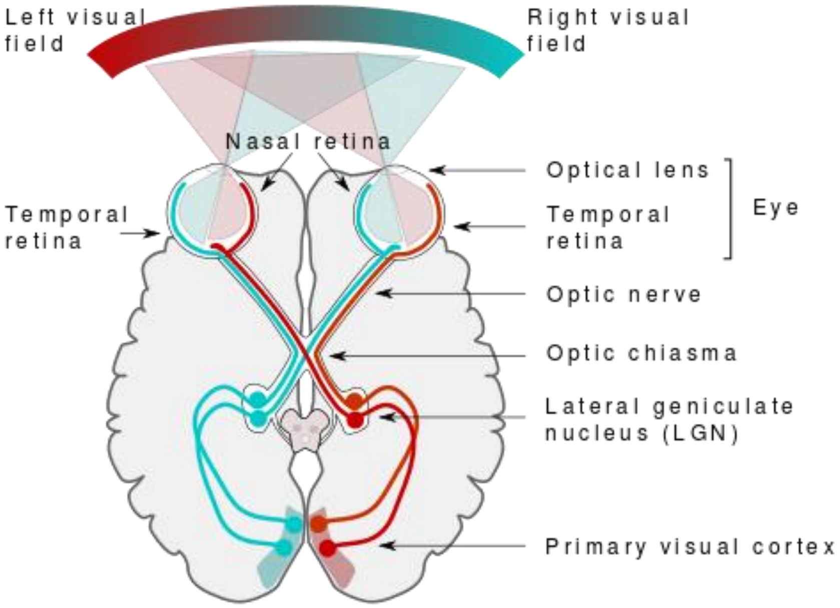


Figure 1.2. Schematic demonstration of the human visual pathway (Nieto, 2015).

The optic nerve then transmits information about the retinal image to the optic chiasm, a cross-shaped structure (Figure 1.2). Here, information from both eyes is amalgamated and split according to visual field. The corresponding half fields of views are contralaterally sent to those halves of the brain to be processed (so information from the right field of view *of both eyes* is sent to the left half of the brain and information from the left field of view *of both eyes* is sent to the right half of the brain to be processed). Both the left and right optic tract (now carrying visual information from contralateral visual fields) continue to the lateral geniculate nucleus (LGN). Neurons in the LGN then transfer the visual image to the occipital lobe and visual cortex where the image is further processed. The primary visual cortex (V1) receives information directly from the LGN and visual information then travels via a cortical hierarchy to higher-processing areas in the cortex.

Areas V1 and V2 are involved in processing basic visual features, as neurons in these regions respond selectively to specific orientation and position, and are believed to process basic information about size and space. Area V3 is involved in shape perception whereas V4 is involved in colour vision. Areas V3 and MT/V5 are involved in motion detection, spatial localisation and hand and eye movements. The complexity of neural responses increases as information passes through the visual hierarchy. For example, where a V1 neuron may selectively respond to a particular orientation, neurons in the visual association cortex may respond selectively to faces. Further specialisation occurs when visual information is split into dorsal and ventral streams (Mishkin & Ungerleider, 1982).The dorsal stream, often referred to as the ‘where’ stream deals with spatial attention. However, this particular area has also been referred to as the ‘how’ stream to demonstrate its influential role in guiding movements towards spatial locations (Goodale & Milner, 1992). Conversely, the ventral stream is known more commonly as the ‘what’ stream as it is involved in acknowledging and categorizing visual input. Whilst substantial documentation exists supporting these two visual streams, there still exists some debate as to how independent they are as there exists a substantial cross-over between the two streams.

Central to our visual abilities are the contrast sensitivity and visual acuity functions. Contrast is specifically the change in luminance over the overall luminance level (ΔL/L). Contrast sensitivity is the log of the aforementioned function, and is understood as the detection of minimal luminance levels of an object of visual focus (compared to its respective background) (Figure 1.4) (Amesbury & Schallhorn, 2003).

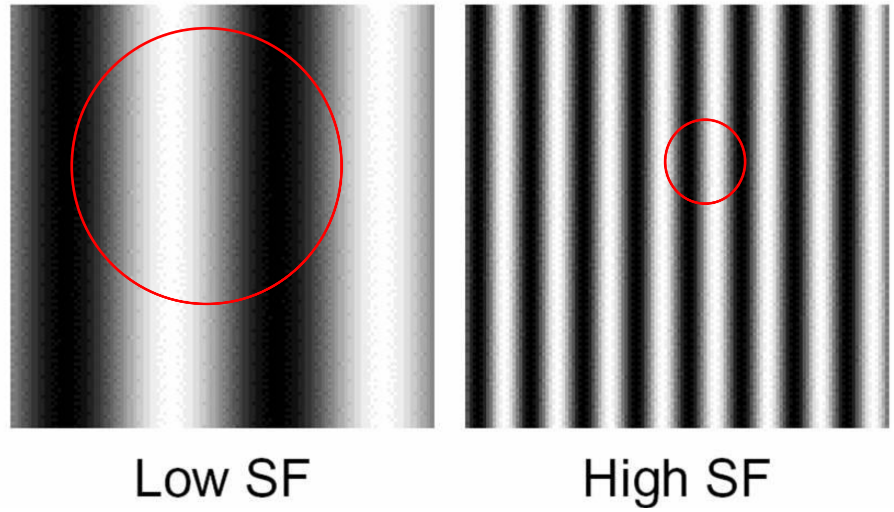


Figure 1.3. Low and high spatial frequency sine wave gratings. (*New York University Website,* retrieved 20 September, 2016, from http://www.cns.nyu.edu/~david/courses/perception/lecturenotes/channels/channels.html, (Landy, 2006)). A sinusoidal grating consists of light and dark bars, the intensity of which is determined by the sine function in trigonometry. The red circles indicate the centre and surrounding concentric areas (see text for further description).

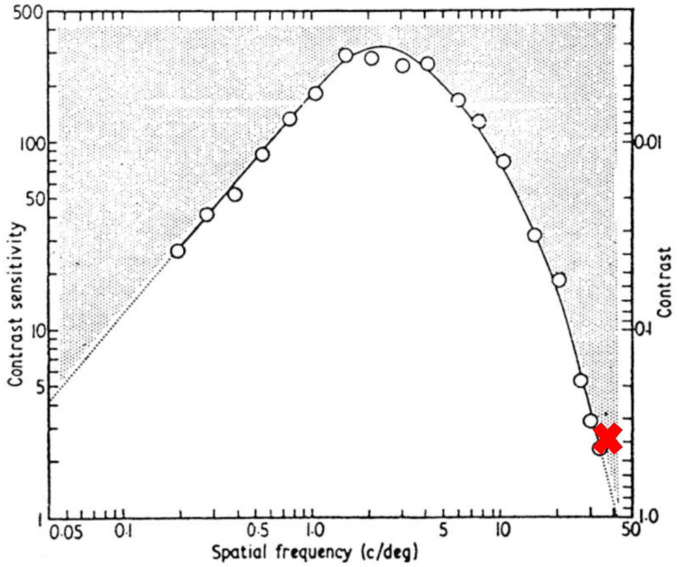


Figure 1.4. Standard contrast sensitivity function curve; visual acuity is a single point on the contrast sensitivity function (Emergent Techniques for Assessment of Visual Performance, National Research Council, 1985; Campbell & Robson, 1968). The red cross represents visual acuity.

Two examples of low and high spatial frequency gratings can be seen in Figure 1.3. Retinal ganglion cells located most centrally in the fovea are known to have the smallest receptive field sizes. Contrastingly, retinal ganglion cells with the largest receptive fields are located in the visual periphery. Receptive field sizes are of incredible importance as they ultimately govern the spatial frequency of visual input. High spatial frequencies (fine detail) stimulate small receptive fields whereas low spatial frequencies (coarse detail) stimulate large receptive fields. Typical receptive fields of retinal ganglion cells consist of two concentric areas (Figure 1.3), known as the centre and surround. These areas perform antagonistically, in simpler terms, light falling on the central area excites the neuron whereas light falling on the surround inhibits the same neuron. When the level of excitation exceeds the level of inhibition, the neuron will cause an action potential to travel down its axon. Figure 1.4 depicts a contrast sensitivity function. The highest visual sensitivity falls between the range of moderate spatial frequencies (around 1-5 cycles/per degree) and thus sensitivity for spatial frequencies above and below this range decreases. The highest spatial frequency visible indicates visual acuity. Visual acuity refers to the sharpness of retinal function in central vision (Hofstetter, Griffin, & Cline, 1997). Simply put, visual acuity describes the ability to see high contrast detail in central vision. One method of measuring visual acuity is using the Snellen chart, where individuals are required to centrally fixate on examples of stylized letters from a fixed distance (known as optotypes). An alternative measure of visual acuity could also present Landolt rings instead of letters. However, a more reliable form of measurement is using logMAR visual acuity charts as these charts typically contain more letters than a Snellen chart. Each row on a logMAR chart contains 5 letters, and each row has a step of 0.1 log units between the next row. The value assigned to each individual letter is 0.02 log units. An individual's logMAR score is thus the total of each letter correctly read. A 6/6 measurement on a Snellen chart is equivalent to a 0.0 logMAR score (Elliott & Flanagan, 2007). The main advantages of LogMAR charts is that the letter size varies logarithmically between lines so is standardised, as is the letter legibility. Additionally, as the logMAR scoring method accounts for each letter, it allows for more reliable and precise measurements of VA compared to simply deducing a score from each line alone (Bailey, Bullimore, Raasch, & Taylor, 1991).

The visual system is not passive and instead is continually adaptive to changes in sensory information. Accordingly, it adjusts to accommodate for these changes (more commonly known as neural adaptation). Demonstrations of such adaptive mechanisms include motion after-effects, orientation after-effects, and negative afterimages (Barlow & Hill, 1963). The motion after-effect is thought to be a result of motion adaptation; whereby after viewing a moving visual stimulus with stationary eyes, fixating on a stationary image results in the perception of motion in the opposite direction, (with respect to the direction of the initial stimulus presentation). Visual adaptation more specifically occurs as responsiveness to a constant visual stimulus changes over time within sensory systems. Notably, visual adaptation can occur for a variety of visual features, such as orientation, motion and spatial frequency; and is thought to occur to establish coherence of the sensory world, and maintain perceptual constancy (Webster, 2015). Neural adaptation will be expanded upon and explored further in Chapter 2.

1.2 Auditory Perception

The ability to perceive sounds is known as auditory perception. This occurs through a detection of vibrations and changes in pressure of the surrounding medium (for example, air or water, through time).

In humans, hearing is performed by the auditory system where vibrations are detected and transduced into nerve impulses by the ear. These nerve impulses are then translated by the temporal lobe and communicated to other areas of the brain. To elaborate on the precise mechanisms underlying hearing and auditory perception, it is essential to understand the three components of the human ear; the outer, middle and inner ear(s) (Figure 1.5).

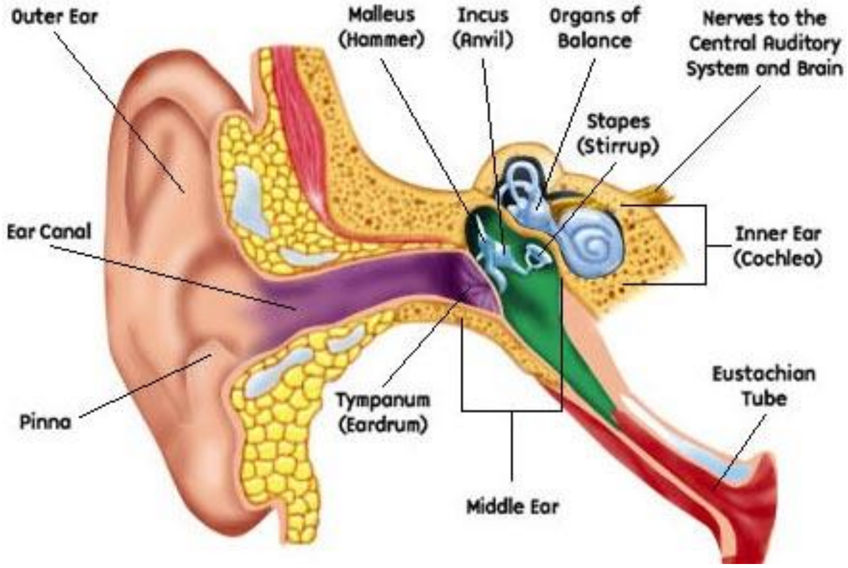


Figure 1.5. Key structures of the outer, middle and inner human ear. (*Hearing Haven website*, retrieved 21 September 2016, from http://www.hearinghaven.com/how-do-we-hear/ear-diagram/).

The outer ear corresponds to the visible part of the ear and the ear canal terminating at the ear drum. This part of the ear also includes the pinna – a structure that helps focus sound waves through the ear canal and in the direction of the ear drum. The structure of the ear drum is that of an airtight membrane and therefore, as sounds arrive in this location, they cause the membrane to vibrate. It is also noteworthy to consider that due to the asymmetrical nature of the outer ear, the location the sound is arriving from will dictate how the sound is filtered. The middle ear contains an air-filled chamber within which are three of the smallest bones in the body (known collectively as the ossicles) and individually as the malleus, incus and stapes. These structures help transmit vibrations from the ear drum towards the inner ear. Lastly, the inner ear contains the cochlea, a spiral-shaped, tube-like structure which contains the organ of corti. This incredible receptor organ allows for the translation of auditory signals into action potentials. Specifically, this occurs when vibrations to the structures of the inner ear cause cochlear fluid to displace and create movement of the hair cells at the organ of corti to produce electrochemical signals. In this way, the organ of corti is essential to allow mechanotransduction in the inner ear and thus, allows for the cortical and cognitive understanding of sound.

Information from the cochlea then travels through the auditory nerve towards the cochlear nucleus in the brainstem. These signals are then projected to other areas of the brain, such as the inferior colliculus, which then integrates this sound information with input from other areas of the brain and allows for subconscious reflexes. The inferior colliculus also projects to parts of the thalamus such as, the medial geniculate nucleus where this sound information is further communicated to the primary auditory cortex located in the temporal lobe. The primary auditory cortex also holds Wernicke's area, an area believed to help interpret the sounds necessary to identify and comprehend spoken words.

As experienced with other senses, the auditory system is not immune to limits; and understandably, there exist some general limitations to human audition. The first of these is in what can actually be heard by humans. Whilst the threshold for detection of frequencies substantially increases with age (meaning lower frequencies can often go unheard), most young and healthy adults can detect frequencies between 20-20,000Hz (Figure 1.6). In terms of an absolute threshold of hearing (meaning the lowest energy physically detectable), it has been found that this absolute threshold of detection depends greatly on the frequency of noise being perceived. A comparative analysis conducted in 1979 suggests that the lower limit of perception lies at -5dB rather than 0dB. However, the authors note that whilst this threshold has been documented, it is incredibly rare and for the majority of people the threshold lies between 0 and 5dB (Robinson & Sutton, 1979).

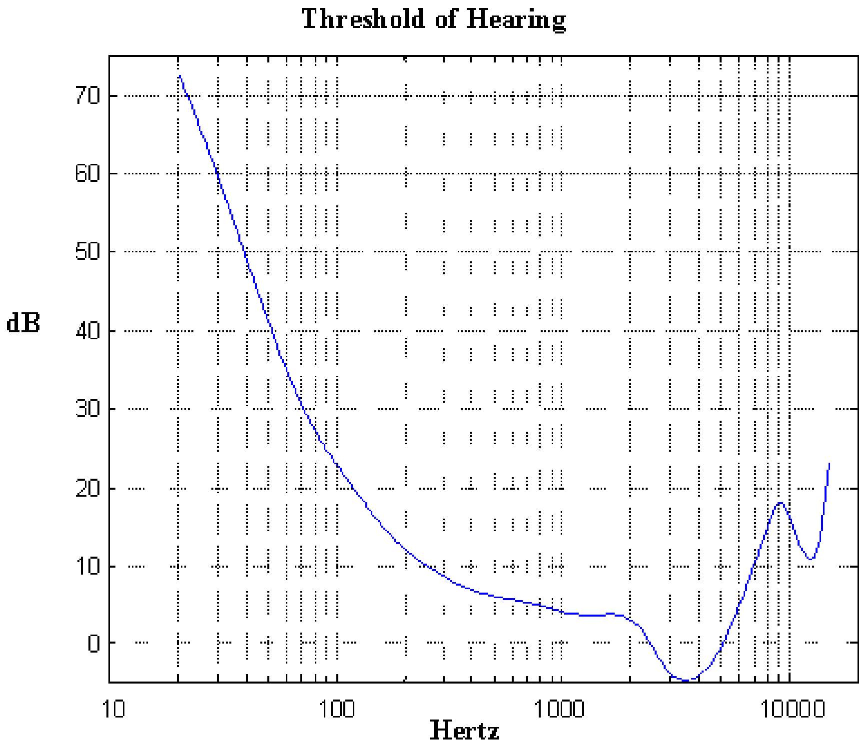


Figure 1.6. Threshold of normal human hearing plot. As auditory perception is influenced by the frequency of signals, the y-axis represents the auditory intensity required for hearing (in decibels), whereas the x-axis represents the frequency of signal presentation (in Hz). From ISO, R. (1987). 226: Normal equal-loudness contours for pure tones and normal threshold of hearing under free field listening conditions. International Organization for Standardization, Geneva.) (*Acoustics—Normal Equal-Loudness Contours [ISO 226:1987]*, 1987).

Even in individuals whose hearing falls within ‘normal’ clinical thresholds, differences in temporal perception bound to intrinsic individual differences can exhibit themselves. Furthermore, despite neurotypical cochlear neuropathy (Shinn-Cunningham, Varghese, Wang, & Bharadwaj, 2017), differences manifest through physiology and behaviour in individuals with “normal” hearing can present as a result of auditory nerve fibre differences that typically respond to sound (Shinn-Cunningham et al., 2017).

1.3 Tactile Perception

Tactile perception refers to the sense of touch. This ability depends on the somatosensory system – an incredibly complicated network of nerve cells, which respond selectively to particular changes - to both the surface being touched and also the internal state of the body. A collection of nerve cells also known as sensory receptors communicate signals along to the spinal cord where these signals can be processed by other nerve cells and later, sent to the brain for extended processing. Such sensory receptors are found all along the surface of the body even in internal tissue, such as the epithelial tissue, skeletal muscles and the cardiovascular system. Thus, the somatosensory system is composed of both sensory receptors and afferent neurons that send signals towards neurons located in the central nervous system.

Broadly, the somatosensory system is a three-order neuronal system that communicates detected sensations peripherally and, using pathways through the spinal cord, brainstem and thalamic relay nuclei, conveys sensory information to the sensory cortex located in the parietal lobe. Receptors carry sensory impulses through sensory afferents to the dorsal root ganglia, where cell bodies of the first order neurons are located. These then travel through to the spinal cord, either ipsilaterally or contralaterally. The spinal cord contains neurons of the second-order fibres containing information regarding pain, touch and temperature sensations. Fibres for second-order neurons containing information regarding touch, position and vibratory sensations are held within the medulla. These fibres are then conveyed either to the thalamus or the cerebellum. The thalamus is the location of third-order neurons. The thalamic nucleus then transports sensory afferents to cortical sensory areas where this information is organised and analysed in an incredibly sophisticated manner.

Whilst sensory receptors are characterised by their ability to identify changes in their immediate periphery, these receptors are also crucially able to adapt to a variety of stimulus features. Specifically, this means they are able to reduce and control their rate of discharge resulting from continuous or repetitive stimulation. Receptors possessing this quality initially respond maximally as the stimulus is experienced, however as continuous stimulation is experienced, this response begins to fade, resulting in effects of adaptation as experienced in other senses such as vision and audition also. It is important to note here that not every sensory receptor holds the ability to adapt to evolving stimulus features and therefore, nonadaptive sensory receptors actually respond continuously for the duration of the stimulus being responded to.

Further specialisation within the somatosensory system occurs depending on the exact type of touch being experienced, explicitly whether this is fine or crude touch. Fine touch, also known as discriminative touch, allows for the identification of the location of the touch. Crude touch on the other hand, is where identification of touch exists however awareness of the exact location of touch is unavailable. Processing of fine touch typically occurs in the posterior (dorsal) column-medial lemniscus pathway which then sends information regarding the fine touch to the cerebral cortex. Processing of crude touch information on the other hand, occurs in fibres located in the spinothalamic tract. A subject will be able to discriminate fine touch so long as the fibres in the posterior column-medial lemniscus pathway operate as normal. As soon as these fibres are severed or disrupted, whilst the subject will still be able to discriminate touch, they will not be able to gauge the precise location of this touch and therefore, will be reduced to experiencing crude touch only.

A classic task used to investigate thresholds for tactile perception is the two-point discrimination task. This task asses the ability to gauge that two closely placed objects are touching the skin at two different points rather than confusing them for one. The test is usually conducted on a range of areas on the surface of the skin to better understand how densely innervated that particular location of skin is. Whilst this is an incredibly traditional task, it has been criticised on occasions for poor resolution of spatial-tactile acuity. Demonstrations of the task showing low sensitivity and understating sensory deficits or even failing to detect them have all been documented (Van Boven & Johnson, 1994; van Nes et al., 2008). In response to these criticisms, a number of alternative tasks have been implemented to test 'pure' spatial-tactile acuity – examples of these include the grating orientation task, the raised letter task and the two-point orientation discrimination task (Craig, 1999; Vega-Bermudez & Johnson, 2001; Tong, Mao, & Goldreich, 2013). As these tasks require the subject to identify the spatial nature of the perceived sensations in an absence of response magnitude clues, for example, identify the exact spacing of the two-points and their orientation rather than simply stating whether they were felt or not, researchers have begun to implement them more often in tactile research (Johnson & Phillips, 1981; Tong et al., 2013).

Wilder Penfield has created a cortical map of body surfaces in the brain (called the 'homunculus' depicted below (Figure 1.7)). It is important to note that whilst this map presents an incredibly useful understanding of the representation of bodily areas cortically, it is still susceptible to change and reports of substantial plasticity exist in subjects who have experienced significant injury or stroke (Borsook et al., 1998).

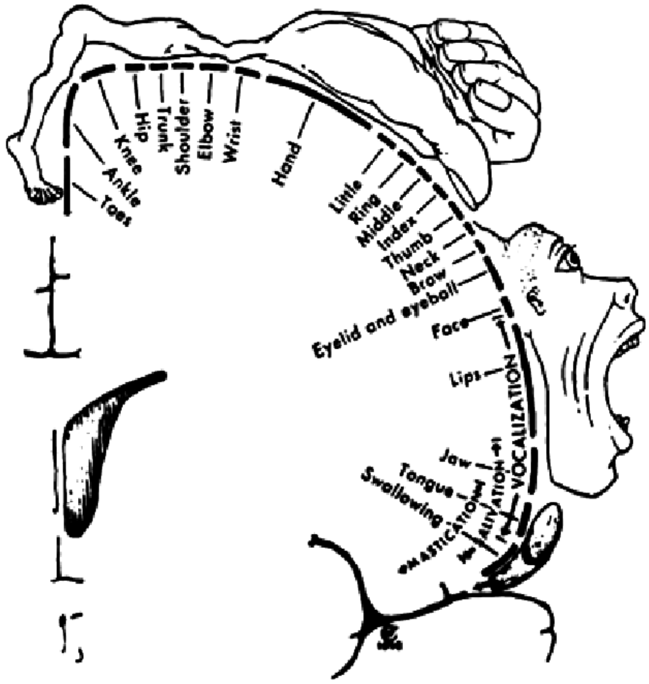


Figure 1.7. Human homunculus indicating physical bodily areas and areas they correspond to cortically (Penfield & Rasmussen, 1950).

On a final note, the somatosensory system is not immune to limits of perception either. Sensory neurons are not distributed uniformly amongst the surface of the skin; in other words, some areas such as the fingers are more densely populated with sensory neurons compared to others such as the back. It is due to this feature that discrimination of different tactile sensation is different depending on the area stimulated. For a minority of individuals, deficits in being able to perceive objects through touch also exists. These individuals are unable to categorize objects in a tactile sense, and this deficit is thought to result from damage or lesions to the somatosensory cortex, termed astereognosis.

1.4 Principles underlying multisensory perception

The purpose of sensory systems is to guide adaptive behaviour. Multisensory processing is the function that deals with how sensory modalities interact, combine and influence processing. It is wholly reliant upon abilities of the nervous system to incorporate and integrate information from numerous modalities. Crucially, not only does multisensory integration allow us to have meaningful perceptual experiences but also maintains perceptual constancy, and is therefore central to adaptive behaviours. Whether groups of temporally coincident sensory signals are to be integrated or segregated is based on the congruence of those sensory signals. Sensory signals can be categorised by their modality (for example, is the stimulus visual or auditory?), the spatial location of the sensory signals (for example, are the signals arriving from the same physical location or separate locations?), and duration of the signals (for example, are they present for the same amount of time).

Stein and Meredith have postulated three principles which multisensory integration appears to observe (Stein & Meredith, 1993):

1) The Spatial Rule: Successful multisensory integration is more likely when unimodal sensory stimuli arise from approximately the same spatial location.
2) The Temporal Rule: Successful multisensory integration is more likely when unimodal sensory stimuli occur at approximately the same time.
3) The Principle of Inverse Effectiveness: Multisensory stimuli are more successfully and effectively integrated when the alternative, unisensory response is comparatively weak.

In support of these principles, data from experimental studies show that subjects typically respond faster to multisensory stimuli compared to the same stimulus presented in isolation (Hershenson, 1962), and to double targets where two (unrelated) targets are presented simultaneously compared to being presented in isolation (Ridgway, Milders, & Sahraie, 2008).

Numerous studies have also indicated the dependence of integration upon several low-level and high-level factors (Radeau & Bertelson, 1977; Welch & Warren, 1980; Welch, 1999; Spence, 2007; Vatakis & Spence, 2007, 2008, 2010). Low level factors refer to temporal information such as temporal synchrony and temporal correlations between modalities and also information regarding spatial locations (Chen & Vroomen, 2013), whereas high-level factors refer to prior knowledge and semantic congruency (Doehrmann & Naumer, 2008). A well-established view amongst researchers in this field of a theory that combines both factors is that of the “assumption of unity” – as information from multiple modalities share more (amodal) properties, the more likely it is that they will be integrated as the brain understands them as originating from a common source (Bedford, 1989). The most crucial amodal property then, is that of temporal coincidence (Radeau, 1994). Only when information from two different sense organs arrives in the brain at the same time are events thought to be considered as multisensory in nature, otherwise two separate amodal events are thought to occur rather than one multisensory event (Keetels & Vroomen, 2012). Constructing a clear and definitive picture of multisensory temporal processing is indeed problematic as the brain has no distinct sense organ that registers time on an absolute scale. In addition, to successfully perceive synchrony the brain has to process differences in physical and neural transmission times, in other words, naturally occurring lags in arrival times and processing times of the different information streams. It is noteworthy that these times also differ for different senses. Ultimately, intersensory timing is flexible and adaptive, however in efforts to deal with various lags between the senses the brain employs a variety of different methods:

1) Manipulating a window of temporal integration
2) Compensation for external factors
3) Temporal recalibration
4) Temporal ventriloquism

Specifically, the first hypothesis suggests that processing systems may be dismissive of small temporal delays between stimulus presentations and therefore, manipulate this window of temporal integration by increasing its duration. The second hypothesis suggests that an incorporation of external factors and previous world experiences may help form coherency in asynchronous presentations of multisensory stimuli, for example, even though in conversation, our ability to observe lip movements and their corresponding sounds may not coincide, our understanding that the two correspond to the same social event may enable us to maintain perceptual constancy. Temporal recalibration refers to the ability of sensory systems to actively alter subjective simultaneity in multisensory signals to perceptually reduce discrepancies between the two and maintain the perception of simultaneity. The final hypothesis refers to an ability to shift the perceived timing of a specific sensory stimulus towards that of another modality, such that they may be perceived as having occurred together.

Multisensory integration is thought to occur in a statistically optimal fashion (Hartcher-O'Brien, Di Luca, & Ernst, 2014) where each sensory signal is weighted according to its precision. As an unbiased estimate is produced with the highest possible precision, this results in a statistically optimum approach (Hartcher-O'Brien et al., 2014). A note should also be made of the ‘Modality Appropriateness Hypothesis’ (Welch & Warren, 1980). Welch and Warren have used this theory to explain how in situations of sensory conflict and uncertainty, the modality most reliable and fitting for the occasion becomes the one to dominate perception; thus, different senses contribute differentially to sensory integration depending on how reliable and appropriate they are, given the task at hand (Welch & Warren, 1980). The Modality Appropriateness Hypothesis explains that in situations dealing with the spatial localisation of stimulus, vision has a greater influence than audition and similarly, in situations dealing with explicit timing, audition has a presence that overrides vision (Welch, DutionHurt, & Warren, 1986). The critical importance of reliability in multisensory perception has also been extended to Bayesian Integration (Ernst & Banks, 2002), who suggest cue combination occurs by utilising the maximum likelihood estimation (MLE) principle. Each sensory signal is associated with its own noise value (due to noise in physical environments and internal conditions, such as inherent noise in internal transmission for example, from spontaneous neural firing) (Ernst & Bulthoff, 2004). The MLE principle deals with minimising uncertainty by combining multiple observations and the noise associated with each observation and associating heavier weightings to the more ‘reliable’ signals. The perceptual decision is then dominated by the estimate with the lowest variance. Sensory information being integrated in this way increases the reliability of the estimates and delivers the “most reliable unbiased estimate” available (Ernst & Bulthoff, 2004).

In summary, sensory systems differ vastly on how they perceive temporal events depending on the time taken to receive and process the range and variety of sensory signals. Our brains intuitively combine signals from multiple senses when those signals are presented in close temporal or spatial proximity (Deroy & Spence, 2016), and it has been noted that multiple senses increase the likelihood of veridical perception of the real world (Stein & Meredith, 1993; Ernst & Bulthoff, 2004). Indeed, the causal origin of cross-modal signals is perhaps one of the most important determinants of multisensory binding (Deroy & Spence, 2016). Other notable features of multisensory binding include bi-directionality, for instance, if vision can influence audition then audition can also influence vision. Many of these cross-modal relationships are shared across cultures (Athanasopoulos & Moran, 2013) and some are even considered universal (Deroy & Spence, 2016). A question currently undergoing intense study is how temporal signals from each modality are weighted (Hass, Blaschke, & Herrmann, 2012). The Modality Appropriateness Hypothesis suggests that only the most reliable modality would be picked to contribute an estimate of time, and consequently, estimates from less reliable modalities would not be used. The Bayesian Integration alternative however, posits that weights are assigned according to the reliability of each modality, thus incorporating all available sources of information to combine an estimate of time (Deneve & Pouget, 2004).

1.5 Neural correlates of multisensory integration

Results from brain imaging studies implicate the posterior medial frontal and insular cortex to be importantly activated in the timing of visual and auditory stimuli, whereas the MT/V5 has been suggested to be necessary for the timing of visual events only. It appears plausible that because multisensory perception and integration involve a multitude of cortical areas, the neural correlates of such abilities will also be (predictably) distributed across multiple neural areas. Future work should aim to further classify the distinct areas responsible for processing multisensory input under a comprehensive and extensive range of environments.

2. Introduction to Time Perception and Timing Models

Time is a dynamic quality fundamental for existence. In demonstrating its importance for survival it is known that almost all plants and animals – even unicellulars, have been documented to express circadian rhythms (Arechiga, 1996). Of note is also the fact that the ability to perceive time does not stand alone – whereas space has high perceptual availability, time, on the other hand, is transient and constantly fleeting - we can go back and check a map – this act, however, is impossible to do with time. Furthermore, as we lack a specialised organ to process time, our awareness of time is solely derived from our sensory systems. These too however, are not without their complications. For example, very few instances in life are purely unisensory (for example, rainbows) – most are multi-modal (for example, speech) requiring an integration of multiple sensory signals. Collectively, this means that our sense and processing of time is dependent upon a range of sub-systems processing sensory input intertwined with temporal signals. These influencing factors and intricacies are what make the study of time incredibly exciting, but complex. For example, temporal processing capacity has been thought to be influenced by a range of different factors, including (but not limited to), sensory modality, stimulus complexity, linguistic demands and combinations of various intensities of these. Deficits in visual temporal processing have been hypothesised to underlie impairments in dyslexic adults (Meyler & Breznitz, 2005). Despite being a fundamental component to physics and philosophy enthralling scientists and philosophers alike for millennia (Muller & Nobre, 2014), we understand very little about this feature, and have made only incremental progress on the subject of temporal processing in humans.

2.1 Models of Time Perception

Interestingly, despite our sense of time providing a foundation to other abilities such as motion perception and action, our sense of time is regularly far from veridical (Shi

& Burr, 2016). Shi and Burr suggest that a combination of adaptive recalibration and minimised predictive errors constitute the human sense of subjective time, suggesting that perception is an inference of sensory stimulation (Shi & Burr, 2016). They further suggest that our sense of time is unique in that it does not arise from a specific or even physical organ, and that all sensory signals contain temporal cues, irrespective of the modality they are presented to. The “heterogeneous” manner of processing is what creates disparities for subjective time across the range of sensory modalities that humans possess. Several lines of evidence support the idea that time is processed differentially depending on the combination of durations and modalities, namely:

- Psychophysical and psychopharmacological experiments both postulate the presence of distinct mechanisms underlying temporal measurements, for instance, Weber’s Ratio – the coefficient of variation, is different for durations shorter and longer than 2 seconds (Gibbon, Malapani, Dale, & Gallistel, 1997).
- Dopaminergic and cholinergic antagonists differentially affect the temporal processing of short (<1 second) and long (>1 second) durations (Rammsayer, 1999).
- Interval discrimination is significantly worse between modalities, compared to within modalities (Grondin & Rousseau, 1991).
- In perceptions of duration, sounds are consistently perceived as longer in duration compared to perceptions of lights (Wearden, Edwards, Fakhri, & Percival, 1998).
- It has been recorded that the auditory cortex appears to have a more profound effect on temporal discrimination on not only auditory stimuli but also visual stimuli. The asymmetric contributions of visual and auditory cortices in time perception have been explained by the remarkable aptitude of the auditory system in timing (Kanai, Lloyd, Bueti, & Walsh, 2011).

Broadly, the perception of time has been split into two schools of models – dedicated and intrinsic models. The former deals with theories presenting mechanisms where time is explicitly and deliberately coded by cortical systems. On the other hand, the latter refers to theories suggesting that time is encoded as an emergent property of neural dynamics (for example, state-dependent networks, which will be expanded on below) (Spencer, Karmarkar, & Ivry, 2009). The following subsections will elaborate on the most comprehensively developed models of time perception.

2.2 Pacemaker-Accumulator Model and Scalar Expectancy Theory

It has been postulated that there is one internal clock that underlies all human timing judgements (Treisman, 1963). Specifically, it has been suggested that this internal clock primarily deals with the function of transforming a period of objective time into subjective time (Allman et al., 2014). Whilst the neural bases for either the pacemaker or accumulator are unknown, they are suggested to have a link with cerebral oscillations (Nagarajan, Blake, Wright, Byl, & Merzenich, 1998).

Within this internal clock model, it is suggested that a pacemaker mechanism exists which emits a series of crucial pulses. When an interval is to be calculated, a trigger switch is activated by the onset of that interval which then allows the counting process to begin, allowing the accumulator to count the total pulses during the interval (Zakay & Block, 1997) and the duration to be estimated from the total count of pulses. The number of pulses emitted during a certain time frame are counted by an ‘accumulator’ which then determines temporal frames. The pacemaker-accumulator model suggests a separate pacemaker for each modality (Hass, Blaschke, Rammsayer, & Herrmann, 2008), these pacemakers emit pulses at particular frequencies which are then modulated by events in that modality (Brown, 1995; Kanai, Paffen, Hogendoorn, & Verstraten, 2006; Eagleman, 2008). A centralised temporal hub then counts these pulses. The accumulator hub and the trigger switch are both centralised. When disparity exists between modalities and their independent estimates of time (Gamache & Grondin, 2010), the final estimate can only be contributed to by modalities that contain both the onset and offset estimates. The pacemaker-accumulator model therefore implies that the same clock times signals from multiple modalities. Evidence suggesting asymmetrical influence of multiple modalities on time perception (Hass, Blaschke, & Herrmann, 2012), would need to be addressed and modified by supporters of the model for it to still be considered an appropriate and relevant explanation of human temporal processing.

Akin to pacemaker theories, the Scalar Expectancy Theory also posits that an internal clock dominates human (and animal) timing behaviour. Specifically, that a pacemaker (the internal clock), accumulator and a connecting switch modulate this internal clock. The theory also proposes memory stores and decision mechanisms that help construct timing behaviour. Further, it has been suggested that this pacemaker does not operate on a fixed rate and can modulate its speed bi-directionally. This means the pacemaker can both, accelerate and decelerate for example, in duration adaptation experiments inducing duration overestimation and compression, respectively (Yuasa & Yotsumoto, 2015). In the understanding of an internal clock model, differences in clock speeds for specific modalities can be as a result of differences in the pacemakers for those modalities thus explaining perceptual differences (Yuasa & Yotsumoto, 2015). The scalar property of timing (also called Weber’s Law) refers to the observation that interval timing errors emerge in a linear manner with the interval’s estimated size. This observation has been documented in a number of animals including humans, rodents and pigeons (Gibbon, Malapani, Dale, & Gallistel, 1997; Malapani & Fairhurst, 2002; Buhusi et al., 2009). Despite the support for this theory from animal studies, properties of human timing are undoubtedly more complex and a key reason behind this is due to attentional allocation (Hallez & Droit-Volet, 2017). Moreover, the Scalar Expectancy Theory has dominated the field for decades positing a single centralised and modality-independent clock. This has recently become challenged by the hypothesis of distributed sensory timing mechanisms across several brain areas/circuits and that the recruitment of these mechanisms depends on the psychophysical task at hand, length of temporal intervals and sensory modality (Ivry & Schlerf, 2008; Vicario, Martino, & Koch, 2013; Mioni, Grondin, Mapelli, & Stablum, 2018).

2.3 Interval Timing Models

Some of the earliest scientific works using time and specifically, duration reproduction, were conducted by Karl von Vierordt who asked subjects to reproduce an interval between two taps by tapping themselves (Vierordt, 1868) which informed us of principles of perceived duration in relation to physical duration (Lejeune & Wearden, 2009). The intrinsic timing model suggests time is an inherent and largely generalized feature of neural dynamics (Bueti, 2011). This suggests that principally, any area in the brain should, and indeed is, able to process and encode time. A great advantage of these models is that because they assume time is encoded the same way as other stimulus properties are such as motion or colour, they allow for an explanation of the functional organisation of sensory timing mechanisms. However, much of the evidence in support of intrinsic timing models relies on much shorter durations of less than 500ms (Buonomano & Maass, 2009; Spencer et al., 2009). And so, for intrinsic timing models to fully explain sensory timing mechanisms, much larger testing durations are needed (Bueti, 2011).

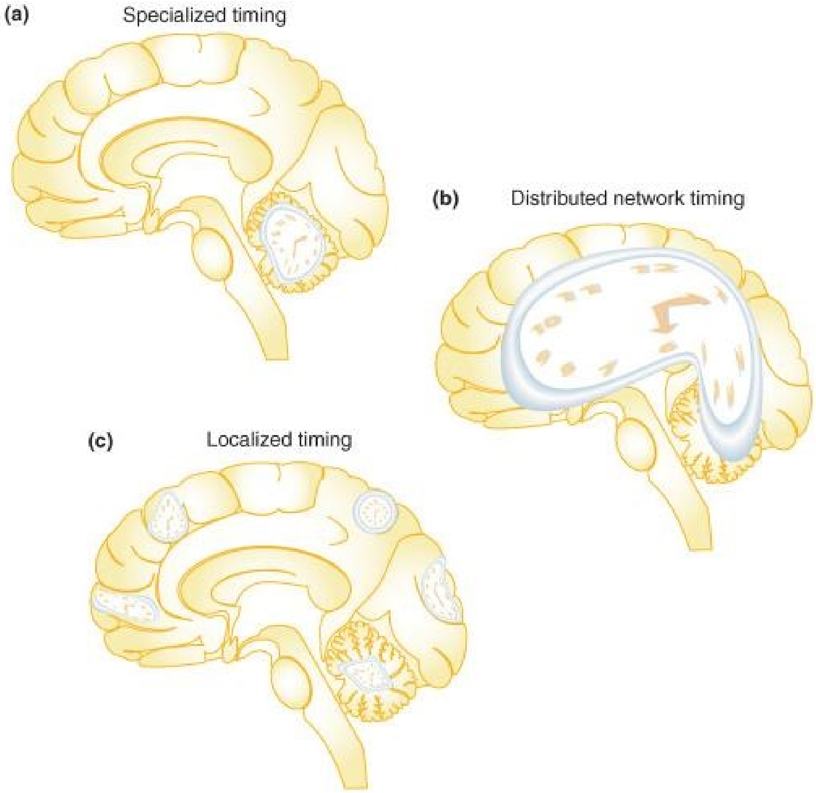


Figure 2.1. Schematic demonstrating the neural mechanisms of timing. a) A specialised timing model posits a specific neural region that is dedicated towards the representation of temporal information. This system is thus recruited when temporal processing is required. The cerebellum is presented as a specialised timing structure in this schematic. b) The distributed timing model posits that temporal information is processed by a symphony of neural structures. c) The local timing model posits that instead of being processed by a dedicated timing mechanism, temporal information is processed by the neural structures recruited by the particular task at hand. Image acquired from Ivry & Spencer, 2004.

It has been proposed that interval-based timers rely on non-oscillatory mechanisms (Wittmann, 1999) where different durations are then processed by dedicated timers specific to those durations localised in the cerebellum (Ivry, 1996). Contrastingly, oscillatory clock-counter systems have been proposed to be localised to the basal ganglia (Wittmann, 1999).

2.4 Oscillatory and Neural Models

2.4.1 Striatal Beat-Frequency Model

Despite having no physical organ, such as those that relate to our perception of colour or sound, our perception of time is no less perceptually salient (Ivry & Spencer, 2004). The physical and biological worlds both provide us with multitudes of oscillatory events. Physical examples include planetary motion in the form of years, seasons and days, whereas biological examples include breathing cycles and heartbeats. Biological clocks can also be entrained to the physical time keepers tracking days and seasons and are present in some of the simplest lifeforms we can examine, such as bacteria, algae and yeast (Fitch, 2012). Entrainment refers to when two or more oscillators become coupled in their activity. Neural entrainment refers to temporal calibration of oscillators within the brain (van Wassenhove, 2016).

Cortico-striatal circuits are believed to subserve interval timing under the Striatal Beat-Frequency (SBF) model. The model proposes a bundle of cortical neurons that constantly oscillate at various frequencies, these are accompanied by striatal spiny neurons responsible for detecting patterns of phases within the cortical oscillating neurons. At the onset of an event, cortical oscillating neurons are reset and begin a new cycle of oscillation. These oscillators are linked to different frequencies which then project this information to medium spiny neurons (A and B) located in the striatum, activating if particular patterns of phases are evidenced amongst the oscillators (Figure 2.2). These medium spiny neurons (A and B) then detect oscillating patterns amongst the cortical oscillators – as different oscillators oscillate to different frequencies, by detecting specific coincidental patterns, the spiny neurons are able to code multiple durations (Buhusi & Meck, 2005; Murai, Whitaker, & Yotsumoto, 2016).

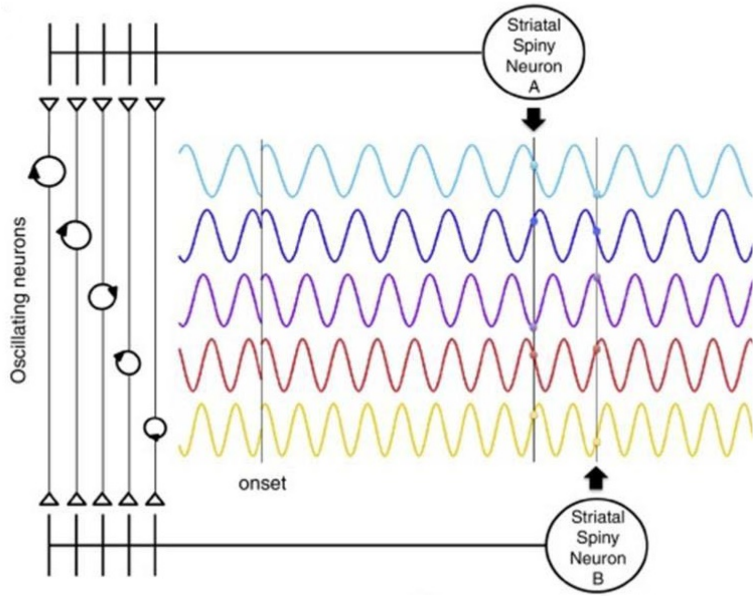


Figure 2.2. The striatal-beat frequency model. The top right of the image shows lower frequencies whereas the bottom right of the image shows higher frequencies. The model posits cortico-striatal circuits that allow a neural construction of interval timing. Cortical neurons on the left side of the image oscillate at a number of different frequencies and the neurons on the right of the image (Striatal Spiny Neuron A & B) detect patterns of oscillations amongst the oscillating neurons. As the oscillating

neurons have different frequencies, this allows the system to code for different durations (Image courtesy of Murai et al., 2016).

Importantly, the SBF model addresses modality-specific and modality-independent features of timing. This is due to the oscillators being located in multiple areas of the cortex, and this assumption is based on the fact that medium spiny neurons actually receive input from all over the cortex (Cowan & Wilson, 1994). Consequently, these oscillators distributed across the cortex can be modulated within a modality-dependent framework. The coincidence detectors in the form of medium spiny neurons are centrally based in the striatum and therefore modality-independent. This hypothesis is further strengthened by evidence promoting the role of the striatum in multisensory integration (Reig & Silberberg, 2014).

Temporal computations that are state-dependent propose that performance of neural dynamics increasingly depend on the sensory modalities demarcating the temporal intervals, supporting the hypothesis that intrinsic neural mechanisms are modality-specific, at least for interval timing (Fornaciai, Markouli, & Di Luca, 2018). A key shortfall of this hypothesis however, is that it fails to account for the processing of intervals presented immediately after one another, as the activity within these networks cannot immediately return to their default resting state. Whilst this has been documented in experiments employing short-intervals of around 100ms, (where performance was markedly improved for an interval presented in rapid succession of another). A recent replication using larger intervals of 300ms found further compelling results. An interval between two auditory stimuli was not found to influence discrimination of those stimuli. Whereas the same trial in the visual modality was found to significantly impair duration discrimination, a result that is consistent with the modality-specific understanding of state-dependent networks (Fornaciai et al., 2018). It has also been documented that coincidental activation of cortico-striatal neurons mediates the representation of time in a distributed manner (Buhusi & Meck, 2005). Oscillator-based explanations of temporal discrimination have also gained support from studies using a range of isochronous auditory sequences deviating from temporal expectations (McAuley & Kidd, 1998).

Additionally, the encoding of duration in state-dependent networks is exclusively context dependent. This is due to the fact that an interval's representation is varied as a function of the networks' initial state. State-dependent networks however, are also not without their limitations. The first requires that the network be in a specific 'regime' that allows the state-dependent change expression. At times these can be nontrivial as excitation and inhibition need to be balanced for this. To elaborate, this means inhibition must allow excitatory neurons to fire without firing in excess and 'preventing runaway excitation'. The second limitation concerns the fact that these networks encode time retrospectively and therefore, would be the least effective at noting particular intervals within a sequence (Bruno & Buonomano, 2004). Moreover, evidence suggests that attentional processes are capable of influencing brief duration perception (Spencer et al., 2009). However, state-dependent networks are yet to include and adapt their model to account for additional cognitive influencers demonstrating an important constraint for intrinsic models.

Oscillator-based models also encounter problems when trying to describe the processing of sequences as this would require oscillators to constantly reset, a feat that is realistically unlikely (Mauk & Buonomano, 2004). Furthermore, oscillatory models are considered to have exceptional robustness to external influences. However, experimental evidence suggests timing systems are considerably plastic in terms of perceptual learning and adaptation in interval discrimination paradigms. There exists a clear need to modify these models if they are to account for such plasticity in neural systems representing temporal information (Nagarajan et al., 1998).

Both pacemaker-accumulator and multiple-oscillator models base their function on dedicated mechanisms that are activated with the initiation of a particular event. Whilst these models can reasonably address the measurement of prospective timing, a problem is posed for retrospective timing. Specifically, this is due to the fact that all perceived events could potentially require timing judgements and therefore, each event would by default, require its own dedicated timer (Addyman, French, & Thomas, 2016).

2.4.2 Synfire Chains

Synfire Chains have been presented to explain the neural architecture representing time (Hardy & Buonomano, 2016). Specifically, this theory posits that a large pool of neurons exist, groups of which are separated into smaller pools and organised in a feed-forward manner (Hass et al., 2008). In a practical sense, this means that activation happens in a chain-like manner where only one sub-pool is activated at any one time. It therefore, becomes possible for circuits lower down in the pool to calculate elapsed time by identifying which pool is currently active. Whilst this model is biologically appropriate in explaining interval timing, we cannot neglect its inability to explain the recurring connectivity in pyramidal neurons (Song, Sjöström, Reigl, Nelson, & Chklovskii, 2005) as pyramidal cells have forward and backward connections throughout the brain. Moreover, this model fails to account for the general cortical reconnectivity due to its feed-forward nature, for instance, in many cases, signals are sent forward but are also reinterpreted retrospectively, an example being in the case of temporal recalibration in speech. To be able to confidently assert this model, issues regarding its limited capacity need to be addressed in future work (Hardy & Buonomano, 2016).

2.5 Temporal Channels Model

Early work by Blakemore and Campbell suggested the human visual system to have numerous channels to filter various spatial frequencies. They suggested each channel was narrowly tuned to a specific range of spatial frequencies (Blakemore & Campbell, 1969). Similar results were found in the macaque visual cortex (De Valois, Albrecht, & Thorell, 1982), and channels specific to orientation in the human visual system have also similarly been suggested (Thomas & Gille, 1979). The channels model in timing suggests that a population of neurons are tuned to specific durations and elapsed time is represented by the firing of these neurons in response to the durations they are exposed to (Desmond & Moore, 1988). Behavioural evidence for such channels processing multitudes of time intervals has been presented by Heron and colleagues (Heron et al., 2012) who suggest duration channels mediate human time perception. Furthermore, an influence of specific duration lengths even in cross-modal stimuli has also been suggested by Filippopoulos and colleagues who present a “common code” for durations irrespective of modalities (Filippopoulos, Hallworth, Lee, & Wearden, 2013). The idea of timing channels existing in neural circuitry is one currently gaining momentum in timing literature, and future work defining the explicit sensitivity of timing within these channels will provide exciting insights into the mechanisms underlying time perception.

It has been suggested that to further develop mature computational models of interval temporal perception, future models should comprehensively address the scalar property of interval time, timing from prospective and retrospective angles and also the effects of additional cognitive and neuropharmacological influencers (Addyman et al., 2016). Furthermore, future models should acknowledge and fit individual data, as well as group data, and explain both similarities and differences in findings from timing experiments on animal and humans (Addyman et al., 2016).

2.6 Centralised versus Distributed Timing

The key contention here is between a central timing mechanism and more segregated mechanisms of timing. The former suggests that timing is centralised and that the system used to determine the duration of a tone is the same which is used for determining the duration of a visual flash. Alternatively, the latter suggests human time perception is distributed across multiple brain networks all capable of temporal processing and that depending on the task, modality and lengths of durations used, different areas will be recruited (Ivry, 1996; Mauk & Buonomano, 2004).

The central tendency of time perception refers to the observation that in a presentation of various temporal intervals, the subjective duration of an interval is regressed to the mean of the various presentations (Hollingworth, 1910; Murai & Yotsumoto, 2016). Murai and Yotsumoto (2016) employed the central tendency effect across the visual and auditory sensory modalities with sub and supra-second durations to assess centralised theories of sensory timing. They reported that when participants were required to reproduce intervals, sub-second interval reproduction resulted in a larger magnitude of central tendency for the visual compared to the auditory modality. In the supra-second range however, both auditory and visual modalities resulted in comparable and correlated central tendency magnitudes. Murai and Yotsumoto then used an interval discrimination task to further assess underlying mechanisms in sub-second timing. A similar pattern of results to the first experiment was found in that a larger central tendency effect was observed for intervals demarcated by the visual modality compared to the same interval demarcated in the auditory modality. A strong correlation however, was found for magnitude of auditory and visual central tendency effects in the sub-second range. The authors report these findings as evidence for both centralised and distributed timing mechanisms, specifically, they suggest that a modality-independent mechanism mediates the central tendency effect in the supra-second range that is irrespective of the sensory modality. The central tendency effect for sub-second durations is however, mediated by both modality-dependent and amodal timing mechanisms (Murai & Yotsumoto, 2016).

In efforts to explore whether a common timing mechanism exists overlooking sub-second time perception, evidence was gathered from multiple interval timing tasks (Merchant, Zarco, & Prado, 2008). Using a range of tasks manipulating the type of processing strategy (either perception or production), and the modality (auditory or visual) and the number of intervals (one or four), results showed that performance variability increased linearly as a function of interval duration across each task (Weber's Law). Performance variability was also larger in perceptual rather than production tasks, and for auditory compared to visual stimuli; it was also found that as the number of intervals increased, so did performance variability. The authors hypothesize distributed mechanisms that are partially overlapped that oversee the temporal processing within different contexts. The common timing hypothesis has been supported by studies where temporal learning has occurred, for instance in cases where training on a timing task can be generalised and utilised to other timing behaviours (Karmarkar & Buonomano, 2003) such as learning to discriminate timing intervals causing a subsequent improvement in motor timing (Meegan, Aslin, &

Jacobs, 2000). Thus, a common timing mechanism such as an internal clock would act independently of modality, tasks and context (Karmarkar & Buonomano, 2003). Conflicting evidence is provided by the fact that our subjective experience of time can be modified by additional factors. Examples include improved performance for filled versus empty intervals (Rammsayer & Lima, 1991), and for auditory compared to visual intervals (Goldstone & Lhamon, 1974) and finally, for multiple consecutive intervals rather than a single interval (Ivry & Hazeltine, 1995). It has previously been argued that temporal judgements acting independently of sensory features supports a centralised timing theory. Merchant et al.'s data however, suggest that several factors (such as processing strategy, sensory modality and structure of stimuli) influenced an subjects temporal accuracy, thereby supporting a distributed sensory timing hypothesis (Merchant et al., 2008).

Sub-second discrimination training of temporal intervals results in specific improvements of the trained interval in terms of temporal discrimination (Wright, Buonomano, Mahncke, & Merzenich, 1997). Evidence suggests this interval-specific training effect can transfer cross-modally and across skin location and hemispheres, despite being temporally-specific (Nagarajan et al., 1998), and even through the perceptual to motor mediums (Meegan, Aslin, & Jacobs, 2000). Moreover, the ease with which we are able to compare time across modalities is often taken as evidence for the presence of a single, supramodal clock. Evidence of the brain using a single timing circuit may however come from the simple nature of many tasks used in human timing experiments, therefore the existence of multiple, more sophisticated temporal circuits cannot be discounted (Mauk & Buonomano, 2004).

Supporting evidence towards distributed timing mechanisms have been presented using more cognitive explorations of duration perception (Takahashi & Watanabe, 2012). Their study presented subjects with a sample stimulus that could be either auditory or visual, the subject was then presented with a variable delay lasting between 0.5-5 seconds and was then presented with a comparison stimulus, again to either the auditory or visual modality. Subjects then responded to which stimulus was longer, the sample or comparison. The most notable finding was that the memory process for stimulus duration was modality-specific and that the perceived duration during the task was dependent on the sensory modality of the sample stimuli that was then compared to the comparison stimulus (Takahashi & Watanabe, 2012).

Behavioural evidence of unimodal auditory and visual rhythm perception has also demonstrated differences between the modalities, specifically, that auditory rhythms express an advantage over visual rhythms in the encoding of beat-based stimuli. The authors further suggest that the encoding of visual rhythms by auditory rhythms can occur through prior auditory exposure, however this only occurs for stimuli that fall within a very narrow temporal range. Moreover, this interaction is neither obligatory nor automatic (McAuley & Henry, 2010). These findings propose further evidence that differences exist between human sensory systems.

Notably, variability even exists in temporal processing within a modality, for example with the auditory time shrinking illusion (Nakajima, ten Hoopen, Hilkhuysen, & Sasaki, 1992). It was found that empty auditory intervals shorter than 200ms were perceived to last even less when they were immediately preceded by a shorter interval. In their behavioural investigation, the length of the first interval was kept constant at 50ms, whereas the duration of the following interval varied from 40-280ms. The authors report that for durations up to 100ms, the perceived duration increased minimally compared to the objective duration. After 120ms, the perceived duration increased swiftly and reached veridical perception at 160ms. The objective duration increased gradually, however subjective perception changed suddenly, suggesting typical categorical perception (Nakajima et al., 1992). The 200ms duration point has been noted as the trade-off point where our timing system shifts from one processing mechanism to another in the processing of auditory durations (Rammsayer & Leutner, 1996). Evidence of the time-shrinking illusion has been replicated within the tactile modality also (Hasuo, Kuroda, & Grondin, 2014) showing that a time interval was underestimated when the interval before lasted a shorter duration.

Auditory dominance has been documented previously (Hass et al., 2012). Recently, it has also been found that brief adaptation to auditory time intervals can modify the perception of motion (Zhang, Chen, & Zhou, 2012; Kaya, Yildirim, & Kafaligonul, 2017). Studies with auditory and tactile unimodal and cross-modal signals in music meter perception have also found that auditory input dominates temporal perception (Huang, Gamble, Sarnlertsophon, Wang, & Hsiao, 2012). A substantial body of evidence (Recanzone, 2003; Burr, Banks, & Morrone, 2009; Chen & Yeh, 2009; Klink, Montijn, & van Wezel, 2011) suggests the asymmetrical dominance of temporal stimuli in the auditory modality on the subjective duration of visual intervals. This evidence has supported the case for the "Modality Appropriateness Hypothesis" (Welch & Warren, 1980), suggesting that the auditory system is significantly superior to the visual system in temporal precision asserting the dominance of auditory perception over vision in the temporal domain. Others, have however, found opposing evidence (van Wassenhove, Buonomano, Shimojo, & Shams, 2008) and even evidence for a symmetrical relationship across audition and vision in time perception (Wada, Kitagawa, & Noguchi, 2003).

Recent psychophysical evidence has provided causal support that time perception and continual motor timing rely on dissociated mechanisms, which hitherto was only support by correlational evidence (Hass et al., 2012). This hypothesis counteracts the modality appropriateness hypothesis and the authors suggest instead, their findings can be explained by a Bayesian account of integration of modality-specific timing information organised by a "central temporal hub". Additionally, in unimodal studies, motion presented visually at a speed similar to that used in Hass et al. (2012) altered subjective duration by up to 400ms, even for intervals in the sub-second range (Brown, 1995; Ryota Kanai et al., 2006). Hass et al. posit that if distortions in subjective duration were truly overlooked by a centralised clock, there should be no difference in effect across auditory and visual modalities. Contrastingly, Hass et al. found evidence that effect sizes in auditory conditions were 2 orders of magnitude smaller than those found for visual conditions. The authors assert that discounting a centralised clock mechanism does not rule out the possibility of amodal clocks, but that further research will be needed to visualise and construct these models (Hass et al., 2012). These contentious findings prohibit the development of a straightforward model aiming to describe the organisation of cross-modal input in temporal perception (Hass et al., 2012).

The processing capabilities for either mechanism, be it centralised or distributed are astonishing considering neural transduction times and the time-sensitive nature of judgements that need to be made in everyday life. Timing is a fundamental and regularly neglected aspect of the majority of human behaviours. While a longstanding contention has questioned the presence of a central, sensory timing mechanism, against more distributed mechanisms between each sense, it has also been suggested that the timing mechanism or timing 'clock' that regulates behaviour may be a function of the task at hand rather than a universal principle overlooking all behaviours (Bueti, Bahrami, & Walsh, 2008).

2.7 Neural correlates of time perception

The problems neural systems face in encoding time can be investigated under three themes (Miall, 1996):

1) Processing continuously variable temporal signals and having to extract information regarding their temporal structure.
2) Detecting, storing and recalling various time intervals.
3) Producing time-sensitive responses.

Notably, the fact that there are no patient groups fundamentally unable to process time (akin to how individuals with amnesia are unable to process and recover memories), perhaps indicates just how crucial timing and time perception is for survival. It is clear then that in exploring a neural substrate for time we are left with two distinct challenges - 1) patient groups who exhibit timing deficits often also present other neural and/or other cognitive confounds, and 2) unlike vision or audition, there is no clear physical organ to trace temporal connections. Furthermore, analysis using fMRI during timing tasks shows activation in multiple areas simultaneously (Ferrandez et al., 2003; Ivry & Schlerf, 2008) namely the more posterior parts of the medial frontal and insular cortex (Wittmann, Simmons, Aron, & Paulus, 2010). These support predictions from brain imaging studies in that the parietal cortex is key in human time perception (Bueti et al., 2008). The role of the parietal cortex, specifically the left side, has been heavily implicated in temporal processing. This has been founded by evidence presented using various neuroimaging techniques, such as fMRI (Pouthas et al., 2005; Rao, Mayer, & Harrington, 2001), electrophysiological studies (Gontier, Hasuo, Mitsudo, & Grondin, 2013) and non-invasive brain stimulation such as TMS (N'Diaye, Ragot, Garnero, & Pouthas, 2004; Wiener, Turkeltaub, & Coslett, 2010) and also recently, transcranial random noise stimulation (tRNS) (Mioni et al., 2018). The recruitment and computational role of the cerebellum in a task-dependent manner has also been emphasised (Ivry & Spencer, 2004). Moreover, the reproduction of both short and long intervals have been suggested to be mediated by the basal ganglia and the cerebellum (Figure 2.3) (Jahanshahi, Jones, Dirnberger, & Frith, 2006). Transcranial magnetic stimulation has presented asymmetric functionality within the auditory and visual cortices in time perception. Auditory cortex disruption impaired visual *and* auditory duration perception whereas visual cortex disruption impaired visual stimuli duration perception *only* (Kanai et al., 2011). Substantial discord however, still remains on the exact cortical structures relevant to the processing of time (Wiener, Turkeltaub, & Coslett, 2010) and to the exact perceptual underpinnings of time. The fact that time cannot exclusively be localised to any one single physical area of the brain results in the study of time perception struggling to define itself and equally, remain segregated from other cognitive features such as attention and memory (Meck, 2005b). The answer to this issue however, can be found within the realm of sub-second timing studies – whereby the range of temporal precision studied can be considered to be automatic and considerably ignorant of other cognitive functions.

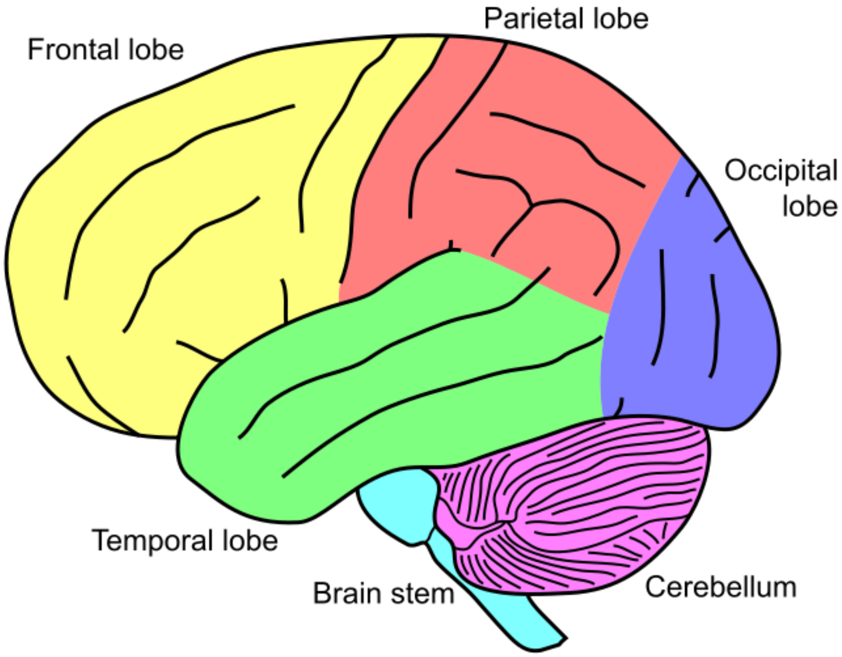


Figure 2.3. Human brain from a lateral view (NeuroTiker, 2007).

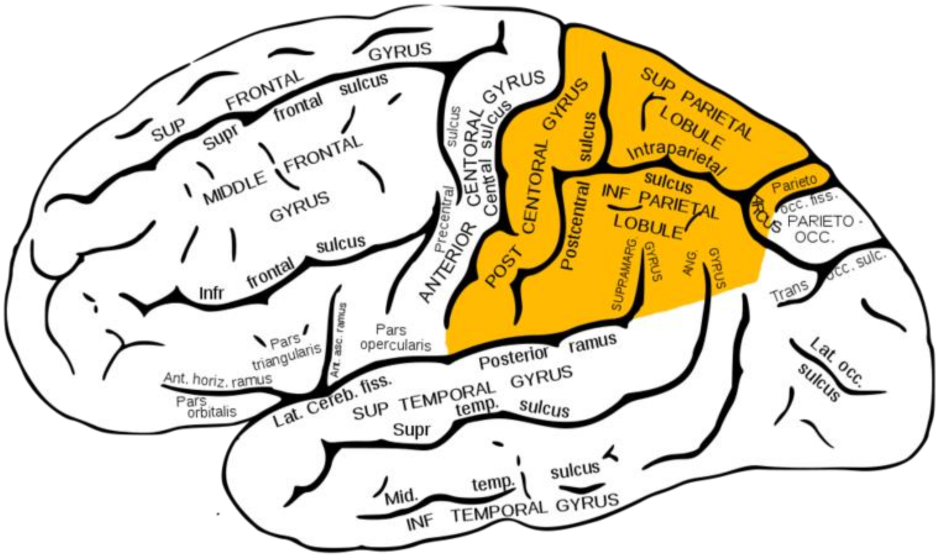


Figure 2.4. Schematic of the human brain with a focus on the different structures within the parietal lobe (Mysid; & was_a_bee, 2010).

Nevertheless, many brain regions have been found to be important in the measurement and processing of sub and supra second temporal durations (Lewis & Miall, 2003). Using fMRI, increased activity was reported in the bilateral insula and dorsolateral prefrontal cortex, right hemisphere pre-supplementary motor area, frontal pole, and inferior parietal cortex during both sub (0.6s) and supra (3s) durations (Figure 2.4). Importantly, differences were also distinguished for both intervals. For the sub-second interval, these were the frontal operculum, left cerebellar hemisphere and middle and superior temporal gyri, suggesting the recruitment of the motor system for shorter auditory durations. The left posterior cingulate and inferior parietal lobule are more active in the supra-second interval presentation conditions (Lewis & Miall, 2003).

As previously stated, along with the posterior regions of the medial frontal cortex and insular cortex, the inferior parietal lobule (IPL) corresponding to the right supramarginal gyrus has also been implicated in cortical time-representation mechanisms (Hayashi et al., 2015). In an experimental set up utilising fMRI and unimodal sensory adaptation, it was found that post-adaptation, a reduction in neural activity was evidenced when a visual stimulus of the same duration was presented repeatedly. Moreover, these adaptation effects gradually decreased as the difference in duration increased between the reference and test durations. Effects were observed across a range of durations (300-600ms) but, more notable is the finding that these effects prevailed regardless of attention to time. The authors also report that the right supramarginal gyrus (SMG) acts as a locus for duration encoding and that neural populations tuned to duration in the right SMG represent time intervals (Hayashi et al., 2015). The authors conclude that populations of neurons specifically tuned to duration represent time intervals in the right IPL, which is then transferred for task-specific processing to the SMA (supplementary motor area) (Hayashi et al., 2015) and that these effects are present irrespective of whether the task instructed subjects to make a same versus different judgement or shorter versus longer judgement in the response task. Further support for this neural structure comes from studies using transcranial magnetic stimulation (Wiener, Hamilton, Turkeltaub, Matell, & Coslett, 2010). Whilst the IPL's complete role in multisensory time perception remains yet to be deciphered, evidence strongly supports its current role in human time estimation.

The role of the cerebellum has also been implicated in time perception and its role in generalising the results of perceptual learning to motor reflexes has previously been suggested (Meegan et al., 2000) (however see Hayashi et al., 2015 for contradicting evidence). It has also been suggested that timing abilities are processed differently depending on the specific task requirements (Bradshaw & Watt, 2002) and an fMRI investigation into the reproduction and estimation of time found that whilst there were common areas activated for both tasks such as the basal ganglia and cerebellum, key differences were also noted, particularly during the temporal reproduction task where a wider cortical network including the right pre-SMA, the left middle frontal gyrus, the left premotor cortex, the fusiform gyrus and also higher visual areas V5/MT were activated (Bueti et al., 2008). The authors suggest that the commonalities in neural substrates indicate that both the perception and (re)production of time are “sustained on the beat of the same clock” further suggesting a centralised timing network (Bueti et al., 2008). The cerebellum is understood to underlie certain aspects of motor timing, believed to rely on distributed mechanisms instead of a single dedicated clock (Mauk & Buonomano, 2004). The role of the cerebellum and the basal ganglia have been heavily postulated in time perception (Wittmann, 1999). To differentiate the functional role of the cerebellum and basal ganglia in timing, it has been hypothesized that both relate to the timing of different ranges. The cerebellum has been suggested to organise timing signals of a brief nature, specifically to the 3 second integration level (Clarke, Ivry, Grinband, Roberts, & Shimizu, 1996). Whereas timing signals ranging from durations of seconds to minutes has been assigned to the dopamine function in the basal ganglia (Meck, 1996; Hazeltine, Helmuth, & Ivry, 1997).

In vivo calcium imaging of hippocampal neurons in mice has shown evidence of “time cells” responding to stimuli in the tens of seconds range. Further investigation of these cells has found that populations of neurons responsible for the encoding of temporal information over the tens of seconds range are also recruited to distinguish temporal periods over significantly longer time scales (Mau et al., 2018). In a recently published review (Allman et al., 2014), a compilation of papers implicating different neural regions in interval timing functions has been presented. Some of the earliest papers date back to 1987 (Meck, 1987) where an aspiration lesion to the lateral agranular frontal cortex in rats was shown to create a shift of timing functions, to more recent work where human fMRI indicated the supplementary motor area, left premotor cortex and the left insula to show greater activation in beat-based subjects compared to subjects who had perceived groups of particular durations (as opposed to beats) (Teki, Grube, Kumar, & Griffiths, 2011). The authors of the review note 18 different structures (some encompassing other areas) in their efforts to highlight the different neural structures that have been implicated in interval timing to date (Allman et al., 2014). The fact that so many areas have been postulated in interval timing alone, suggests that it relies on complicated and considerably distributed mechanisms of the brain.

Many studies exploring the role of the basal ganglia in mediating time perception have involved supra-second intervals (Maricq & Church, 1983; Pastor, Artieda, Jahanshahi, & Obeso, 1992; Meck, 1996) with the exception of (O'Boyle, Freeman, & Cody, 1996), whereas those examining the role of the cerebellum have typically used sub-second intervals (Hore, Wild, & Diener, 1991; Ivry & Gopal, 1992). It is important to investigate how exactly, the cerebellum, a structure heavily implicated in time perception would serve this function. Cerebellar Purkinje cells are understood to be a site of substantial informational convergence (Hazeltine et al., 1997). As many as 200,000 parallel fibres are thought to provide input to a single Purkinje cell, promoting efficient pattern recognition (Albus, 1971; Marr, 1969). Purkinje cells may realise patterns of activity that are duration-dependent along parallel fibres that then signal when anticipated temporal events could occur. Duration could also be coded in the input to Purkinje cells in numerous ways, for instance, granule, stellate and basket cells could operate akin to neural networks converting temporal to spatial information (Buonomano & Mauk, 1994; Buonomano & Merzenich, 1995). In this way, slow pre-synaptic and post-synaptic mechanisms could still enable neurons to locally represent temporal information (Hazeltine et al., 1997).

Temporal learning and the resulting cortical plasticity were investigated on modulation rate discrimination on 9 participants using magnetoencephalography

(MEG) (van Wassenhove & Nagarajan, 2007). Specifically, the authors investigated whether three hours of training over a period of three days were enough to substantially improve temporally modulated tone train discrimination. Furthermore, whether such improvements were linked to systematic plasticity within the auditory cortex. The authors report that systematic increases in amplitudes of early auditory evoked responses were paired with discrimination learning of auditory temporal modulation, as a response to trained stimuli. Interestingly, plasticity within the auditory cortex and learning in part generalised to tasks of interval discrimination but not to frequency discrimination suggesting that these two processes are dissociable (van Wassenhove & Nagarajan, 2007). These findings suggest the constant level of dynamic updating that continues throughout life to maintain perceptual efficiency in order to live in a constantly chaotic world.

Approximately 4% of the general population has amusia (tone deafness), and research on these cohorts of individuals suggest a dissociation between detecting pitch changes compared to detecting changes in time (Hyde & Peretz, 2004). Specifically, all individuals with amusia showed poorer performance in detecting pitch changes, this result did not change with practise. Contrastingly, time changes were detected on similar performance levels as controls and also demonstrated improvements with practise. The authors suggest a possible congenital neural anomaly that selectively diminishes pitch processing. In relation to this, case studies have also identified individuals who can move to a beat but fail to perceive the rhythm of music (Begel et al., 2017). These individuals are able to process regularity in time when not paying explicit attention to the rhythm. The authors assert that motor synchronization may perhaps be supporting this implicit perception of rhythm and also that if despite poor perception, synchronization to a beat can still occur, then perhaps perception and action can be dissociated in explicit timing tasks (Begel et al., 2017).

After investigating sensitivity to temporal structure on continuous sound sequences using MEG and the latency of offset responses within those sound sequences, it was found that when sequences are ignored, the temporal structure of even simple sequences is imprecise (Andreou, Griffiths, & Chait, 2015). Furthermore, pattern coding was substantially improved when the sensory signals had been made behaviourally salient. The authors report that the learning of structure in temporally defined stimuli is not “automatic” and in fact, is modulated by the relevance of those signals to behaviour (Andreou et al., 2015).

A substantial body of work using a range of participant groups, timescales and stimuli contexts now implicates the integration of cortical circuits with the basal ganglia, cerebellum and hippocampus to support temporal and motor processing in at least one or more dedicated timekeeping mechanisms (Gibbon et al., 1997; Doyon, Penhune, & Ungerleider, 2003; Meck, 2005a; Jin, Fujii, & Graybiel, 2009). It is also clear that the posterior parietal cortex alongside the cerebellum and basal ganglia plays an important role in timing and time perception. However, interpretation difficulties still exist in neuroimaging studies investigating time perception. This is due to the magnitude and latency of time-related brain activity, as stimulus duration can at times overlap with the size of sensory input and/or the latency of neuronal activity elicited by stimulus presentation (Murai et al., 2016). Carefully manipulating temporal parameters to identify the neural underpinnings of temporal adaptation will be a challenge that future research needs to address.

2.8 Sensory Adaption

Sensory adaptation refers to the process of a sensory system adjusting its neuronal response as a result of a change in the external environment. In other words, changes in the external environment can create shifts in sensory status or the sensory resting level. These shifts subsequently impact sensory perception, which in turn, is able to influence and adjust behaviour. Neurally, adaptation refers to a change in responsiveness of sensory systems to an external stimulus. The sensory neurons stimulated activate and respond immediately, they respond progressively less and less until they may even cease to respond altogether (Webster, 2012). A common everyday example is of retinal light/dark adaptation in the visual systemoccurring when leaving a brightly lit area and entering a dimly-lit room for instance, at the cinema. Adaptation is important as it can reveal the existence of dedicated processing systems.

Visual adaptation has been reported as early as approximately 350BC (Aristotle, 350BC). Notable investigations of one form of visual adaptation in the form of orientation were made by Blakemore and Campbell (1969). Using human subjects and grating patterns, they showed that after being exposed to a particular orientation of this grating, repulsive after-effects during the test period were found when a grating that was differently orientated (e.g. vertical) now appeared to be angled directly opposite to what it was previously. To further express the range of adaptable features, a second example of spatial frequency adaptation is presented below in Figure 2.5.

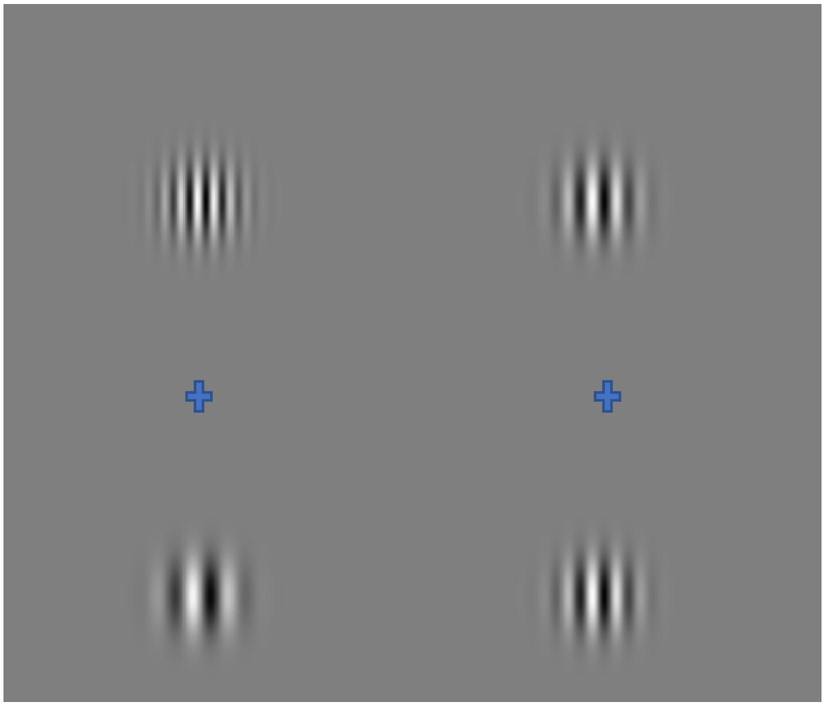


Figure 2.5. Example of spatial frequency adaptation. If we fixate on the left hand side for 20 seconds, it creates a distortion of the spatial frequencies demonstrated on the right. This is an example of a repulsive/rebound after-effect typical of many adaptation experiments.

As the human sense of time can only be communicated through sensory systems, adaptation is an instrumental tool in allowing the experimenter to rapidly and readily manipulate and adjust perceptual sensations. Moreover, temporal adaptation is considered a favourable method to study time perception as it allows the experimenter to investigate similarities and differences between temporal and sensory perception. It specifically allows the experimenter to isolate aspects of temporal processing, such as duration perception, which may otherwise be more difficult to study (Murai et al., 2016). Repeated stimulus presentation results in reduced neuronal activation in neurons typically responsive to that stimulus (Sawamura, Orban, & Vogels, 2006). Evidence of spatially-repulsive adaptation after-effects have also been found with tactile spatial perception (Li, Chan, Iqbal, & Goldreich, 2017). As similar after-effects have also been evidenced in the visual and auditory modalities, it has been suggested that in order to code dynamic sensory input, sensory systems may utilise fundamentally similar approaches, at least in the spatial sense (Li et al., 2017). The neural adaptation model suggests the existence of duration-tuned cortical neurons, which each selectively respond to a conservative range of stimuli durations and centre on the neurons preferred duration. Repeated exposure to a set duration therefore decreases the activation in corresponding neurons, and thereby modulates the relative activation of the neural population resulting in a repulsive after-effect (Li, Xiao, Yin, Liu, & Huang, 2017).

3. Methodology

The relationship between the physical and perceptual worlds can be quantified using psychophysics. This is done through meticulous control of stimulus presentation and the gathering of a subject's response to this stimulus. Data from the subject's responses then allows the experimenter to extrapolate the processing and perceptual rules that the nervous system observes (Fechner, 1860). Some of the more commonly used psychophysical methods are elaborated on below.

3.1. Signal Detection Theory (SDT)

Each perceptual decision is accompanied by a level of uncertainty. Alongside extraneous noise that exists within the stimulus itself, there are also several other sources of noise, two of which include internal noise and cognitive noise. Internal noise refers to the low-level random firing of neurons, whereas high-level cognitive noise refers to the noise within a decision process. The contributions of internal noise are fixed, as the random firing of low-level neurons is fixed. Higher level noise however, can be influenced through the subject's cognitive state, for instance, if they are tired and unable to fully attend to the task at hand. SDT captures the relationship between this noise and its influence on thresholds of perceptual detection (Figure 3.1).

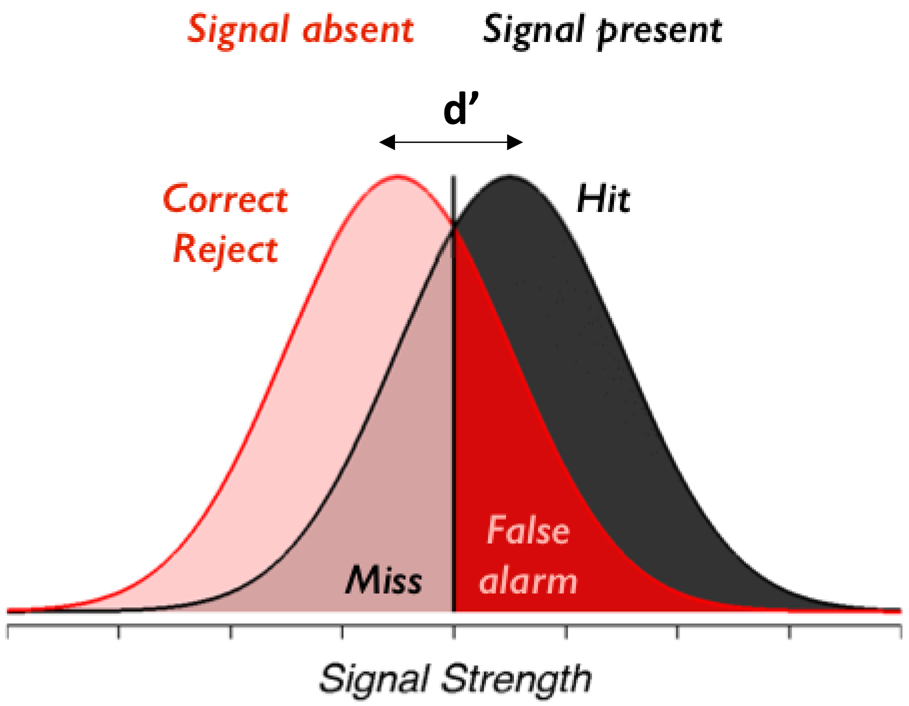


Figure 3.1. Schematic of probability density functions for the signal detection theory. The red curve demonstrates noise only, whereas the black curve indicates the signal and its accompanying noise. The ‘hit’ rate refers to an event where the signal was present on a trial and the subject correctly responds as ‘yes, the signal was present’, whereas the ‘correct reject’ refers to all trials where the subject correctly responds as ‘no, the signal was not present’ on that trial. The ‘miss’ section refers to those trials where the signal was present however, due to the internal criterion level of the subject, the subject failed to identify the signal. The ‘false alarm’ section reflects the opposite event – where the subject incorrectly identifies the signal as being present (despite it not actually being present). The criterion represented by the black arrow demonstrates the d’ (d-prime) representing the discriminability of the signal and relating to the degree of overlap between each function (Gardner, 2019).

Figure 3.1 demonstrates that on any given trial there may be four possible outcomes; the subject correctly identifies the signal (a ‘hit’), the subject correctly identifies the lack of a signal (a ‘correct rejection’), the subject identifies a signal when one is not present (a ‘false alarm’), or the subject fails to identify a signal when one was present (a ‘miss’). A more tolerant subject may set a low criterion value, resulting in an increase in correct hits, but also in false alarms. Contrastingly, a more conservative subject may set a high criterion value resulting in the opposite – more correct rejections but also more misses. The criterion value can therefore directly influence data derived from psychophysical tasks.

The ability to discriminate a signal from noise is referred to as the d-prime value (d’), which explicitly represents the degree of overlap between the two functions of the signal and the noise. A weak signal corresponds to a small d’, and the likelihood of sensory experience resulting from either noise, or a combination of the signal and noise is roughly equal. Contrastingly, when the signal strength is increased, d’ is also increased, resulting in an increased likelihood of the sensory experience arising from the signal. Discriminability varies with the spread of each function, and also the horizontal separation of the two functions (Figure 3.1).

A key assumption of SDT is the criterion. This criterion is a value internally selected by the subject, and is specifically placed at a particular point based on the likelihood ratio. This value acts as the subject’s decision-making threshold. Above this threshold they would respond positively as ‘yes, a signal was detected’ and below which they would respond negatively as ‘no, a signal was not detected’. The setting of this criterion is dependent upon the subjects’ goals and also the impact and consequences of their decision. If the subject asserts a high criterion for a positive response, the risk of a false negative is increased (responding as a ‘no’ despite the signal being present). Similarly, if the criterion is set too low, the subject exposes a higher risk of a false positive (responding ‘yes’ when the signal was not present). This criterion is variable and can be updated from trial to trial as the subject incorporates consequences from false positives and false negatives.

3.2 Psychophysical Methods

3.2.1The method of limits

This is the typical method used to calculate absolute thresholds. Specifically, calculating the absolute threshold requires the experimenter to present grossly supra-threshold stimuli (stimuli that are very obviously detectable) or stimuli that are grossly sub-threshold (stimuli that are highly unlikely to be detected). A stimulus that is either sub or supra-threshold is presented, after which the subject is required to respond to whether they detected the stimulus in a yes/no response. Should the first stimulus be sub-threshold, an ascending series will become initiated whereby the resulting stimulus presentations will increase the stimulus value in incremental steps. The value of this particular increment is pre-determined and is referred to as the ‘step size’. The subject will once again be required to make a judgement on the particular stimulus at each presentation, for example, ‘did you detect the stimulus on this trial, yes or no?’. After reaching a specific threshold, ‘no’ responses will gradually become ‘yes’ responses. The inverse of this process is known as a descending series and is begun with a supra-threshold presentation of the first stimulus. In this case, after reaching a specific threshold, ‘yes’ responses will gradually become ‘no’ responses. The absolute threshold is then calculated by averaging the several ascending and descending threshold values.

An alternative version of the method of limits may present the subject with two different stimuli (stimulus A and stimulus B). In this version, the presentation of stimulus A would remain constant while stimulus B’s magnitude or intensity is adjusted in relation to stimulus A in either ascending or descending order. In an ascending version of this task, stimulus A and stimulus B may be physically identical at the onset of the task. Stimulus B will then be adjusted on each trial until the subject reports a subjective difference between the two. This point is known as the Just Noticeable Difference, or JND. In the descending version of this task the opposite happens, where the task is begun at a point where stimulus A and stimulus B are grossly different. The magnitude or intensity of stimulus B will then be adjusted to a point where the subject can no longer distinguish between the two (Figure 3.2).

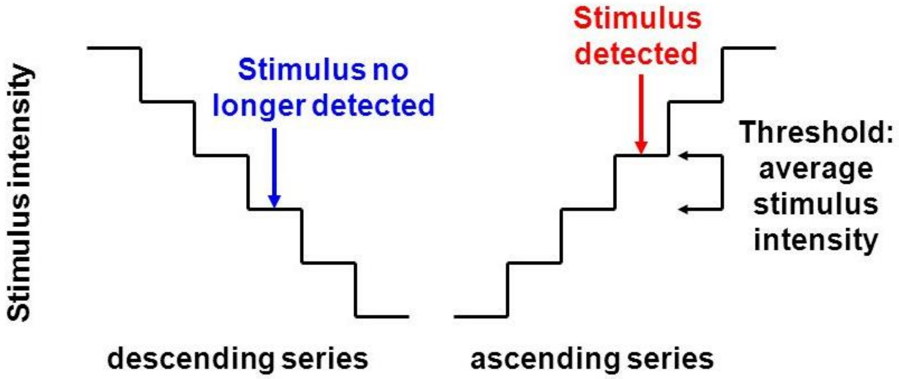


Figure 3.2. Examples of ascending and descending series in the method of limits (Larsson, 2019).

The method of limits is known to be affected by a number of cognitive factors presenting response bias, particularly errors of expectation and errors of habituation. Expectation errors arise from a keen subject who decides 'hastily' and specifically, prior to reaching their subjective threshold value. This type of subject will respond 'yes' to subthreshold stimuli when uncertain. On the other hand, errors of habituation refer to a subject who has become habituated with responding as 'yes' (on a descending series) or 'no' (on an ascending series). As both types of errors arise partially from inadequate experimental procedures, as opposed to the actual stimulus parameters, both types of errors are termed 'response biases'. One approach to overcome these is randomising the starting point of the stimulus presentation to minimise the magnitude of habituation and expectation errors. Providing the subject with thorough and clear instructions can also work to reduce response bias.

3.2.2. The staircase method

An adaptive procedure (commonly known as an adaptive staircase), is considered to be a variation on the method of limits. Upon reaching the threshold at which the subject's responses would transition (for instance, from 'yes' to 'no'), the experiment would continue and the subject would be presented with further presentations in a reverse stimulus sequence (Figure 3.3.). For example, in the first instance the subject could be presented with an ascending sequence, once the responses gradually transition from 'no' to 'yes', the stimulus value corresponding to this threshold is recorded and the subject is now presented with a descending sequence. As before, once the responses gradually begin to transition from 'yes' to 'no', the threshold value is recorded and the subject is again, presented with an ascending sequence. This pattern continues until a number of reversals (that are usually pre-determined, akin to the step size) have been made. The last few pre-determined reversals are then averaged to provide a value for the subject's threshold.

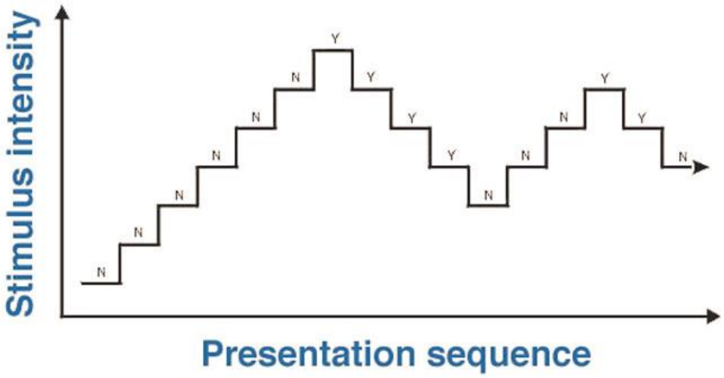


Figure 3.3. An example of the staircase method. Y refers to the response 'yes, the stimulus can be seen' whereas N refers to 'no, the stimulus cannot be seen' (Kalloniatis & Luu, 2011).

3.2.3 The method of constant stimuli

Sensory thresholds are determined using the method of constant stimuli by randomly presenting the subject with stimuli of which some are sub-threshold and others are supra-threshold. Exposing the subject to stimuli typically spanning a comprehensive range (for instance, from 'always seen' to 'never seen'), can then allow the experimenter to be confident of identifying the point of interest.

When a stimulus is presented over a range of different strengths and intensities, it enables the experimenter to generate a ‘percentage correct’ value for the subjects’ responses. This ‘percentage correct’ value can span the range of ‘always detected’ to ‘never detected’ for each stimulus intensity/strength. A psychometric function can then be formed from these values when each stimulus strength has been presented to the subject an equal and sufficient number of times. This is done by plotting the resulting values (of ‘percentage correct’) against stimulus intensity thereby forming the psychometric function (Figure 3.4.). The threshold using this method is typically considered to be the point at which the stimulus is detected 50% of the time.

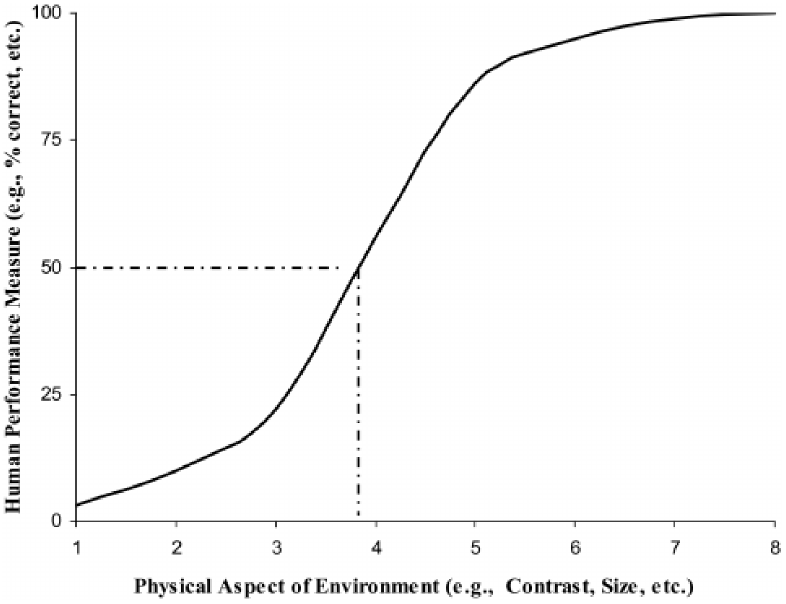


Figure 3.4. Standard psychometric function. Plotting the range for ‘percentage correct’ against the range of stimulus intensities allows the experimenter to present this psychometric function; x-axis demonstrates a specific feature of the stimulus, y-axis represents the value of correct responses on the task as a percentage (%) (Fidopiastis, Fuhrman, Meyer, & Rolland, 2005).

A key difference between the method of constant stimuli and the method of limits is the randomised presentation of stimuli. When the method of constant stimuli is employed, the subject is equally likely to be exposed to any stimulus from the range of stimuli that could be presented. In contrast to the method of limits which presents subjects with stimuli in a fixed order, the method of constant stimuli prohibits the subject from predicting and anticipating where they are in terms of their threshold, and also eliminates the aforementioned errors of expectation and habituation. A notable disadvantage of this method however, is that it can be considerably time-consuming to gather reliable data on all possible stimulus features for an equal number of observations.

3.3. Psychophysical decision types

A detection task requiring a yes/no response is considered to be the simplest psychophysical decision type. As described in the method of limits, the subject would be presented with a stimulus and then typically be asked ‘did you detect the stimulus on this trial, yes or no?’, after which they would respond as either ‘yes’ or ‘no’. The ‘yes’ and ‘no’ responses generated from this task can then be used to form a psychometric function signified by a characteristic ogive as shown in Figure 3.5. Moreover, the stimulus level at which the subject responds with ‘yes’ on 50% of trials is referred to as the threshold.

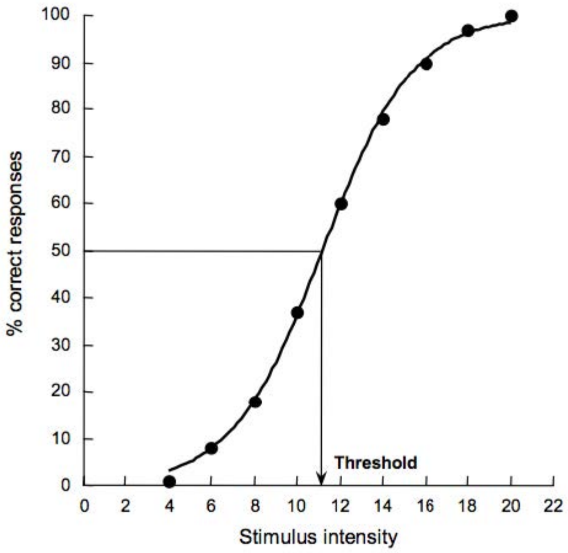


Figure 3.5. A typical psychometric function as generated from a yes/no detection task. The black arrow indicates this subject's threshold at 50% (Fulcher, 2017).

Another very different decision type is the 'two-alternative forced choice' (2AFC) decision. In this procedure, a subject is exposed to two stimuli with only one varying on a specific stimulus feature, for example, one of the two stimuli may be longer or shorter than the other, or one of the two stimuli could be dimmer/brighter than the other. The subject is then required to respond to whether the stimulus feature was present in the first or second stimulus. The subject has two choices, either 'the first' or 'the second'. The subject is prohibited from abstaining from responding, and they are also not permitted to respond as 'I don't know'. In this way, the subject is forced to respond with either 'first' or 'second'. This varies considerably from the yes/no decisional response associated with measurements of absolute threshold.

The responses gathered from yes/no decisional responses are more likely to be influenced by criterion shifts. This is because a more conservative subject may set a very high or low criterion to respond as 'yes' or 'no' (respectively). If the former is the case for example, the subject is likely to respond as 'no' if they have some level of uncertainty regarding the stimulus presence. Similarly, a relatively low threshold for 'yes' responses would result in the subject responding in the opposite manner and result in consequent shifts of the psychometric functions. In contrast, the 2AFC procedure is less susceptible to internal criterion shifts. This is due to the fact that the subject is making a judgement based on two stimuli relative to one another, rather than on the absolute presence of a stimulus. As there is no advantage of employing a response bias towards the first or second interval when both present uncertainty, criterion shifts become largely redundant. Additionally, on each trial and in each response period of a 2AFC task, the subject is presented with noise *and* the signal (to be responded to), resulting in both types of information being employed when making the final judgement. Due of these features, the 2AFC procedure is considered by many psychophysicists to be the ‘gold standard’ (Heron, 2006) as it provides an overall more robust and thorough measure of performance.

3.4. Psychophysical tasks

3.4.1. Reaction time tasks

At their most basic demonstration, reaction time tasks are perhaps the simplest of psychophysical tasks available. In essence, a subject is presented with a stimulus, and is required to respond as rapidly as possible upon its presence. The response is usually made by pressing a particular key on a keyboard.

3.4.2. Simultaneity Judgement tasks

Similar to other psychophysical tests, simultaneity judgement tasks present subjects with two stimuli in rapid succession. The delay between these stimuli is known as the stimulus onset asynchrony (SOA) and is typically varied. Upon presentation of the two stimuli, the task requires the subject to respond to whether the stimuli were presented simultaneously or not (Figure 3.6.). Assessing judgements of simultaneity across a range of SOA’s allows the experimenter to gauge each subject’s point of subjective equality (PSE) – that is, the point at which the subject believes the stimuli to be presented at the same time (Figure 3.7.).

Simultaneity has been found to be influenced by the sensory modalities that the timing signals are presented within (Wittmann, 1999), and simultaneity judgement tasks are also known to suffer from criterion-dependent biases as the judgement on whether the stimuli are simultaneous or not entirely depends on a subject's internal threshold.

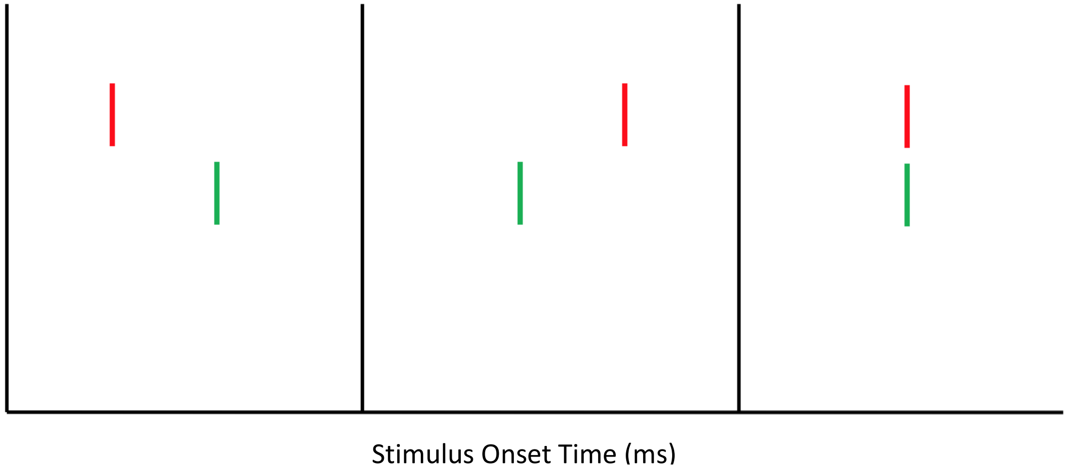


Figure 3.6. Schematic showing different presentation orders in a simultaneity judgement task. The red bars indicate stimulus A, whereas the green bars indicate stimulus B. The left side of the panel suggests stimulus A is presented before stimulus B, whereas the middle portion of the panel demonstrates the opposite. In contrast the right side of the panel demonstrates a condition where stimulus A and stimulus B have the same onset time. The distance between the red and green bars indicate the stimulus onset asynchrony (SOA).

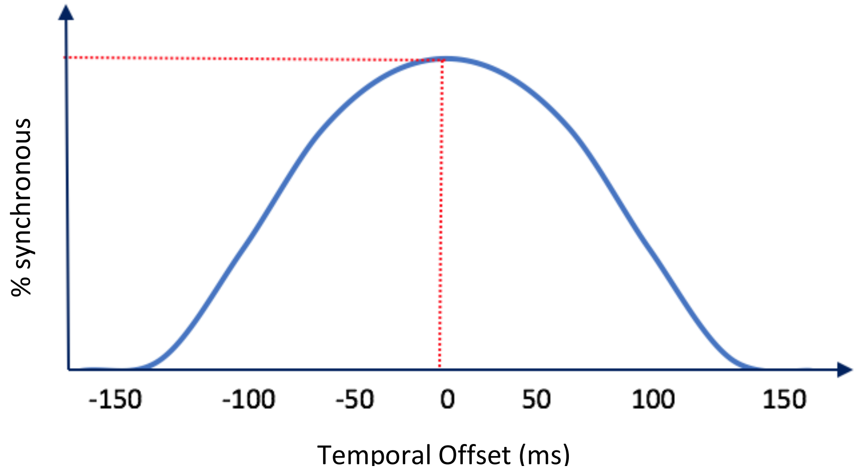


Figure 3.7. A typical simultaneity judgement task. The red dashed line indicates the point of subjective equality (PSE). The exact point at which the subject indicates two stimuli to be temporally synchronous will depend on the subject's internal criterion.

3.4.3. Temporal Order Judgement tasks

Temporal order judgement (TOJ) tasks present a subject with two (or more) stimuli in succession. The subject is then required to respond by indicating which stimulus was presented first thereby requiring unspeeded binary, forced-choice judgements. As this task involves collecting data over a range of different temporal orders, the complete and final dataset can give knowledge of both the point of subjective simultaneity (PSS) and the just noticeable difference (JND). As explained previously, the PSS indicates the point at which the subject would feel that the two stimuli have occurred at the same time. The JND refers to the minimum physical difference needed between two stimuli for the subject to consciously perceive the two stimuli as having a different temporal order.

TOJ tasks are typically thought to have a higher task difficulty than simple reaction time or stimulus judgement tasks (Vroomen & Keetels, 2010). This is because simultaneity judgement tasks require subjects to respond whether stimuli were synchronous or asynchronous, whereas TOJ tasks extend this judgement and require subjects to respond according to *which* stimulus presentation preceded the other. Therefore, TOJ tasks require greater cognitive precision compared their alternatives.

3.4.4. Temporal Reproduction tasks

This explicit method of time perception has been used as far back as 1931 (Triplett, 1931). In essence, the subject is exposed to a particular stimulus and is asked to reproduce either the duration that the stimulus was presented for, or alternatively, the temporal frequency (also known as rate) at which the stimulus was presented. Responses are made most commonly by either a keypress or by tapping on a response disk with the index finger. This task requires fewer cognitive resources (e.g. attention or memory) when responding and therefore reproduced estimates tend to demonstrate less variability in their corresponding error values (Gil & Droit-Volet, 2011; Indraccolo, Spence, Vatakis, & Harrar, 2016).

Rate reproduction is a relatively underexplored method in psychophysics. As rate refers to explicit frequency of sensory presentation, rate reproduction experiments involve subjects recreating the temporal rate of stimuli already presented – traditionally recorded through tapping on a response device. It is important to note that perception and motor action have been proposed to operate under the same timing mechanism, efficiently synchronizing the two abilities (Tomassini, Vercillo, Torricelli, & Morrone, 2018), and thereby validating the use of this response method in studies of time perception.

For the experiments described within this thesis, taps were executed on a piezoelectric disk. These responses were then extracted in Matlab using the ‘audiorecorder’ function where each tap was recorded as a spike. The frequency of responses were calculated by averaging the number of spikes by the response window - usually lasting 2 seconds (unless otherwise stated in the methodology for each chapter).

As afferent pathways carry signals towards the central nervous system, there is a chance that reproducing test frequencies by tapping may produce some afferent input. In the following experiments the method of reproduction was never singled out for one sensory modality (meaning all modalities were responded to using this method). The likely impact of this is that if the afferent nature of responding created some temporal lag, it would have applied to all conditions using the reproduction response method and consequently becomes negligible.

3.4.5. Synchronisation-Continuation tasks

Synchronisation-continuation tasks are one of the more challenging tasks in timing research. They consist of two components. The synchronisation feature of this task involves keeping pace to an external stimulus by tapping (or responding regularly in another form), over a period of time. The continuation feature of this task examines whether the same individual can maintain a motor response at the same speed, despite the termination of the external stimulus. Errors in the form of accuracy are normally measured.

The difficulty of synchronisation-continuation tasks comes due to the artificial nature of this task. For instance, many individuals in the general population (barring musicians such as drummers), do not face this type of task on a day-to-day basis. It is for this reason that measures using synchronisation-continuation tasks regularly report higher intra- and inter-subject variability (Wing & Kristofferson 1973a, 1973b).

As explained above, while there are a number of methods used to quantify temporal perception, in this thesis I will primarily utilise temporal reproduction tasks (Chapters 5, 6 and 8) and the two-alternative forced-choice task (Chapters 6 and 8). The experiments conducted in Chapter 7 will utilise a simplified form of the two-alternative forced-choice task by asking subjects whether the stimulus was regular or not (thereby presenting two options of which the subject is forced to pick one from).

3.5. Weber’s Law

Central to the foundational theory of psychophysics is Weber’s Law. Weber’s Law quantifies perceptual change in a physical stimulus. Explicitly, the law states that a change in the stimulus that is “just noticeable” is proportional to the original physical stimulus (Britannica, 2016). In simpler terms, an empty bag can be made to feel heavier by the addition of one single book. This same single book, may however, fail to make a bag of books feel significantly heavier – despite the fact that in both cases, the same amount of physical weight is being added. Weber’s Law thus explains that to increase the perceived intensity of a stimulus, its physical magnitude must be increased by a constant **proportion**, rather than a constant **absolute** amount. Gustav Fechner later adapted Weber’s Law to incorporate individual differences in perception and acknowledged that the subjective sensation of a stimulus is directly proportional to the logarithm of the stimulus intensity. Together, Weber and Fechner defined objective methods to measure the limits of sensitivity in human sensory systems. Specifically, they were able to establish detection of the weakest detectable sensations in terms of the stimulus energy necessary to produce them. As explained previously, sensitivity of perceptual systems is defined in terms of thresholds. In simple terms, a threshold represents the limits of the perceptual system. The absolute threshold refers to the smallest amount of stimulus energy necessary to produce a sensation, i.e. the smallest intensity of stimulus that can be perceived. Contrastingly, the difference threshold measures the minimal difference in two stimuli needed to elicit subjective awareness that the two stimuli are not the same (or identical). In practice, psychophysical laws are applied to results from experiments most commonly through psychometric functions. Data derived from assessments of absolute sensitivity and discrimination assessments allow us to plot thresholds and extract the point of subjective equality (PSE) and quantify the just noticeable difference (JND).

3.6. Curve fitting – The Psychometric Function

Results from psychophysical tasks are analysed by fitting a curve through the plotted data. Curve fitting using a goodness-of-fit metric is typically referred to as ‘regression’. The resulting psychometric function then allows the experimenter to extract and quantify relevant parameters such as the JND and PSE (Figure 3.8.).

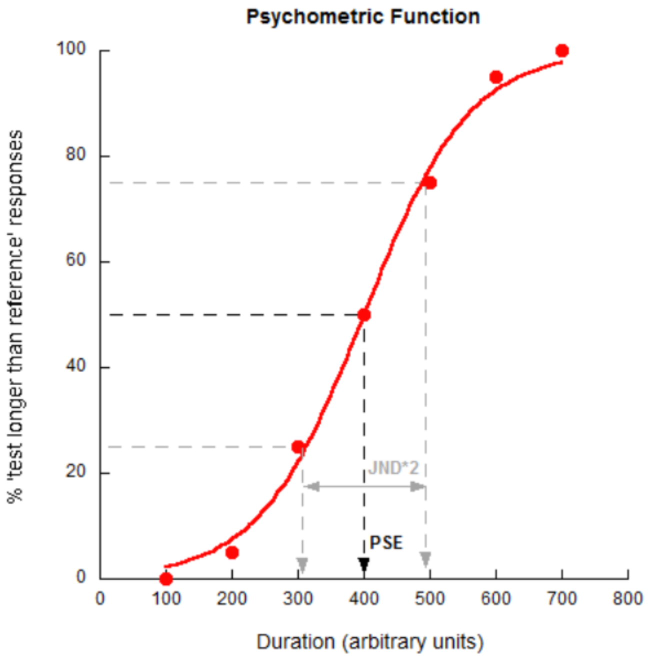


Figure 3.8. A typical psychometric function. The PSE arrow indicates the point of subject equality as the point at which both stimuli that are being compared are perceived as perceptually equal (to the subject). The difference between the performance values (with the lower end at 25% and upper end at 75%) and the PSE indicates the just noticeable difference (JND), as indicated by the grey arrow (Fulcher, 2017).

The psychometric function must aim to present the most representative parameter values utilising and acknowledging all data points and their corresponding error values. All psychometric functions presented within the experiments conducted in this thesis are fitted using the method of least squares employing the Levenberg-

Marquardt algorithm, and using the computational capabilities of the computer software “Kaleidagraph 4.5.2”. This is a widely-used modelling method within and outside the field and has great adaptability to other data sets also. Using this method, unknown parameters are estimated by minimising the sum of squared deviations between the data and the model. The effectiveness and resulting esteem towards the method of least squares is due to its effectiveness and completeness. Specifically, this is because many behaviours and processes in science are well-described by linear models – largely due to the processes being inherently linear or because at least over short ranges, most processes can be approximated by a linear model. Another advantage is that the method of least squares is able to capture trends and model the data well even despite relatively small data sense.

As the method of least squares is substantially vulnerable to outliers, alternative methods include the method of least absolute deviations, this method too however, can still leave outliers to considerably impact the model. A second alternative is the Maximum Likelihood Estimation model, which is an improvement in that it is not as sensitive to outliers in the response variable, however, may still be sensitive to outliers in other variables, in which case no specific advantage is presented (when compared to the method of least squares).

To fit the data using the method of least squares, Kaleidagraph employs an iterative method to minimise the sum of squared vertical offsets of each data point. This method is described as iterative because Kaleidagraph computes the sum of least squares repeatedly until the curve with the smallest difference from all data points is arrived at. The resulting curve will allow the experimenter to extract the values corresponding to the slope and mid-point of the curve (Weisstein, 2005).

A strength of the method of least squares is that as well as being appropriate for linear regression, it can be applied to a broader range of functions. In the following experiments, the psychometric functions are fitted with a logistic function *f(x)* of the form

$$y = \frac{100}{1 + exp - \frac{(x - \mu\,)}{\theta}}$$

In this function, μ represents the estimated mid-point of the psychometric function (PSE) and θ represents an estimate of the threshold of discrimination (JND).

The experimenter must explicitly estimate both these values prior to any computational software attempting any iterative process. This can sometimes act as a disadvantage of the method of least squares, as the subject is required to ‘eyeball’ and roughly gauge the data in efforts to estimate these values. Should the values estimated by the experimenter be far from veridical, the curve fitting may generally result in a poor fit and in misleading parameter values. Furthermore, a second disadvantage is the receptiveness to outliers. This is due to the method of least squares utilising the vertical distances of each data point from the curve, and then squaring these. As a consequence, it becomes possible for a single outlier to substantially impact the fit of the curve. A possible solution to overcome this is to present each stimulus strength a large number of times during experimentation. In doing so, as the presentation number increases any questionable measurements become normally distributed, thereby minimising the impact of any outliers.

For all psychometric functions in this thesis, mu and sigma values were output from the fit of the psychometric functions in Kaleidagraph. These were approximated also using the method of least squares and thus were the most representative values of the overall data. There was no manual altering (minimising) of the error as the parameter values provided by Kaleidagraph were those with the lowest estimates for their respective errors.

The error value listed during each iteration is a normalized Chi Square value. It is calculated by taking the current Chi Square value and dividing it by the sum of the squared values of the function. This value is multiplied by 100 to obtain a percentage. This percentage is compared to the Allowable Error to see whether or not to continue iterating. The errors that are displayed for the parameters are the standard error of the parameters. It can be read as the parameter value +/- the error.

4. Investigating Sensory Timing using Psychophysics

Having described the general methods used in psychophysics (Chapter 3), the present chapter will now go on to examine specific investigations of sensory timing concentrating on the methodology they have used.

4.1. The influence of sub and supra-second stimuli on timing

A considerable body of evidence suggests that time perception adheres to Weber's Law; specifically, that the ability to subjectively discriminate two temporal durations is dependent upon the ratio of the two stimuli's physical differences (Dale, Gratton, & Gibbon, 2001; Wearden & Bray, 2001; Brannon, Libertus, Meck, & Woldorff, 2008; Hayashi et al., 2015). Using confirmatory factor analysis, earlier work conducted by Rammsayer and Troche (2014), suggested that interval timing is mediated by a unitary mechanism for sub and supra-second intervals. More recent work however, conflicts with this view. Differences in duration discrimination judgements have been found between sub and supra-second temporal intervals of visual and auditory modalities. Rammsayer and colleagues (2018) suggest that the processing of very brief (sub-second) intervals is mediated by modality-specific and sensory-automatic processing, whereas discrimination of longer (supra-second) intervals is determined by more amodal and higher-order cognitive capacities. They assert that these processes are distinctly separate yet functionally related. Additionally, that these processing differences determine the differences in duration discrimination across the visual and auditory modalities, and what they name 'The Sensory-Automatic Timing Hypothesis' (Rammsayer, Borter, & Troche, 2015; Rammsayer & Pichelmann, 2018).

Research into time perception has more commonly utilised intervals in the range of 100ms to a few seconds and in human timing literature, intervals below a second have received considerable attention. This is in essence due to their relation to fundamental adaptive behaviours such as motor coordination and in speech perception. Specifically, the 0.1 to 1 second range has received consideration because:

1) The highest sensitivity for temporal discrimination is located between 300 – 800ms (Drake & Botte, 1993; Friberg & Sundberg, 1995).

2) The preferred rate of tapping in a task (otherwise known as a preferred tapping tempo) is known to lie at around 350ms for children, 600ms for adults and 700ms for seniors (Zelaznik, Spencer, & Doffin, 2000; McAuley, Jones, Holub, Johnston, & Miller, 2006).

3) Moreover, it has been asserted that intervals consisting of durations shorter than 100ms appear instantaneous whereas those lasting 5s or longer involve long-term episodic memory. In human timing experiments it is crucial to consider that any durations lasting longer than 1000ms may encourage humans to use chronometric counting. Despite counting still being possible during shorter durations, it is typically considered to hold less utility (than in longer durations) (Fraisse, 1984; Wearden & Lejeune, 1993; Nichelli, 1996).

In addition, short-term adaptation to stimuli possessing both sub and supra-second durations failed to elicit the same amplitude of after-effects observed with long-term adaptation (Li, Xiao, Yin, Liu, & Huang, 2017), suggesting the duration after-effect critically depends on the exposure time to the adapting stimuli.

Moreover, studies employing mono-aural (one ear) and inter-aural (both ears) anisochronous (temporally irregular) sound sequences have found differential processing strategies for different temporal frequencies (ten Hoopen et al., 1994). Particularly, they find that the processing mechanisms employed for sequences faster than 3-4Hz differ to the processing mechanisms used for slower rhythms (ten Hoopen et al., 1994).

4.2. Experimental features influencing performance

Other considerations when presenting stimuli include the localisation of stimuli, for example, auditory stimuli could be presented over headphones, via speakers, or localised with visual stimuli and presented through speakers or even presented in a spatially offset manner with respect to visual stimuli. In a large proportion of the upcoming experiments, we chose to present acoustic signals via headphones rather than localising the sound output via speakers. The spacing of our ears on either sides of our heads means that sounds arriving off the midline will have to face differences in the path lengths of the source of the sound to the physical ear as an ear farther away from the sound source will take longer to receive the sound. This difference in arrival times of sounds to each individual ear is referred to as the inter-aural time difference (ITD). The magnitude of this difference is dependent upon the precise physical architecture of the head and ears and is therefore variable between individuals (Carlile, 1996). Thus, to reduce the confounding effects of varying ITDs across our subjects we localised all sounds to the same pair of headphones for each subject.

Regarding the processing and comparing of cross-modal stimuli, explanations of a temporal cost have been considered. Explicitly, that a temporal cost is involved in switching between different senses and that this additional processing strain results in longer response times when switching between sensory modalities in temporal judgement tasks (Spence & Driver, 1997).

4.3. The influence of stimulus presentation order in 2AFC tasks

Bausenhart et al. (2015) set out to investigate whether the type B effect (improved performance when the standard is presented prior to the comparison) can be generalised across other standard magnitudes (Bausenhart, Dyjas, & Ulrich, 2015). The authors find that the type B effect is prevalent across a range of standard magnitudes, however it diminishes as the inter-stimulus interval (the gap between the standard and comparison stimuli) is reduced (Bausenhart et al., 2015).

Discrimination sensitivity in 2AFC tasks has shown to be dependent upon the presentation order of standard and comparison stimuli (Hellstrom, 2003; Lapid, Ulrich, & Rammsayer, 2008; Hellstrom & Rammsayer, 2015). When interpreting two-alternative forced-choice designs, it is important to establish whether the standard interval presented was a roving or rather, fixed standard. A fixed standard is a stimulus whose position is fixed on each consecutive trial (for instance, where the standard is always presented before the test stimulus), whereas a roving standard is one whose position is varied (it is presented either first or second) on each trial. A fixed standard may allow subjects to build an internal representation of the interval in their long-term memory, rather than relying on the temporal information provided live during a trial. Furthermore, this fixed standard interval is supported by future trials also employing the same standard duration, thereby allowing a more precise running average of this interval (Pashler, 2001).

To assess temporal sensitivity with a roving standard test interval, Pashler conducted two experiments on the perception and production of short (auditory) temporal intervals presented either in a sequence of 2 or 6 auditory tones. The first experiment employed a two-alternative forced-choice design and the second used similar methodology, however substituted the 2AFC design for interval production of the standard interval. Pashler found the effect of the standard test interval to not be statistically significant suggesting that either a roving, or fixed standard would influence thresholds similarly (Pashler, 2001).

4.4. General stimulus features influencing perception

Marked improvement in performance is found for louder auditory stimuli, compared to quieter stimuli, and for brighter/larger visual stimuli when compared to dimmer/smaller ones (Goldstone & Goldfarb, 1964; Berglund, Berglund, Ekman, & Frankehaeuser, 1969; Zelkind, 1973; Goldstone, Lhamon, & Sechzer, 1978; Xuan, Zhang, He, & Chen, 2007). Higher intensity stimuli are also perceived as lasting longer (Allan, 1979; Fraisse, 1984; Wearden, Edwards, Fakhri, & Percival, 1998).

As are stimuli that are attended to in comparison to their unattended counterparts (Mattes & Ulrich, 1998).

4.5. The influence of response structure on temporal reproduction

In a methodological investigation of interval reproduction, subjects were randomly assigned to one of three groups, where reproducing the interval involved either pressing to start and stop the interval, continuously pressing throughout an interval or pressing at the end of an interval. It was found that the three different methods elicit different results and that the most veridical results were found by pressing a key to start and stop the interval. Contrastingly, continuously pressing a keypress throughout an interval generated the least variability (Mioni, Stablum, McClintock, & Grondin, 2014). Furthermore, correlational analysis has shown that perception and motor production via finger and foot tapping share a common mechanism, suggesting that individuals with low variability with one medium tended to fare similarly in the other medium (Keele, Pokorny, Corcos, & Ivry, 1985). The authors also found that perceptual acuity and the regularity of motor production were also highly correlated (Keele et al., 1985).

4.6. Notable considerations for the investigation of duration perception

Rammsayer and Brandler (2004) have identified two key methodological options to investigate common mechanisms underlying tasks that require precise timing – the first involves the correlational approach and represents the assumption that if two tasks are underlain by the same temporal mechanisms, then a strong correlation should be evidenced in the subjects' performance and judgement variability. Similarly, the second method uses slope analysis (Ivry & Hazeltine, 1995) and compares these across tasks, suggesting that differences in judgement variability as evidenced in slope of the Weber functions, reflect underlying perceptual timing differences. For instance, identical slopes of Weber functions between two tasks suggests the presence of a common mechanism (Rammsayer & Brandler, 2004).

Despite this, it may be possible however, that different mechanisms are employed, each running simultaneously.

4.7. Distinguishing empty and filled intervals

Figure 4.1. Schematic demonstrating the difference between an empty (top half of diagram) and filled (bottom half of diagram) interval of the same temporal length. An empty interval is demarcated by two separate, and brief sensory signals signifying the beginning and end of an interval, whereas a filled interval is signified by an uninterrupted sensory signal presented for the entire length of the interval.

The key distinction here, is that of empty and filled intervals (Figure 4.1). Empty intervals are those defined as having a clear gap between two transient signals demarcating the start and end of that interval. Filled intervals on the other hand, present a continuous and ongoing signal presented throughout the full duration of that interval.

In some of the earliest work in this field, it has been reported that even intervals as short as a few milliseconds are enough to elicit the perception of two separate sounds (instead of one) (Hirsh, 1959). A lengthier interval of between 15-20ms is needed for the listener to report which interval succeeded the other. This interval judgement window occurs irrespective of the length of stimuli, the frequency of stimuli or bandwidth of auditory stimuli (Hirsh, 1959). Additionally, it is reported that listeners are able to discriminate sound intervals of brief durations with remarkable precision and that changes as low as even 5-10% of the size of the interval can be noticed for intervals as short as 100ms (Hirsh, Monahan, Grant, & Singh, 1990). To extend these findings of temporal discrimination, the authors had subjects listen to sequences of 6 tones where one tone was positionally offset in the sequence. Absolute discrimination as part of a sequence was poorer the longer the intervals and thus, the slower the tempo. Measures of relative discrimination however, were better the slower the tempo. They explain their findings by asserting that if tones are played at a fast enough rate, they may be combined and perceived as a single entity thus corrupting discrimination ability (Monahan & Hirsh, 1990).

Cross-modal sensitivity and consistency to tactually and visually designated empty time intervals has also been investigated (Erp & Werkhoven, 2004). Using the two-alternative forced-choice task, subjects were asked to report whether the second interval presented to them was shorter or longer than the first. Two pulses of either tactile or visual signals defined the intervals. A total of 4 standard interval durations were used ranging from 100-800ms. The authors report key differences in perception and that, for tactile intervals to be subjectively perceived as long as visual intervals, they must physically be 8.5% shorter in length. This bias is strongest for shorter intervals and decreases with the lengthening of standard intervals. Furthermore, Weber’s Law holds for the range of intervals (100-800) tested (Erp & Werkhoven, 2004). In a comparison of filled intervals using the auditory and visual pairing, it was found that filled visual intervals in the order of 1 second had to physically be set longer than auditory intervals for the two intervals to be judged as having the same duration (Behar & Bevan, 1961).

Estimates of filled intervals of time can be influenced by a number of different factors, including the sensory modality and the intensity of the stimulus (Indraccolo, Spence, Vatakis, & Harrar, 2016). In examinations of these factors using temporal reproduction and response time judgements, a number of key insights have been found. The first was that visual stimuli evoked longer reproduction times when compared to auditory stimuli of the same durations. Moreover, that longer reproduction times were found for low intensity stimuli when compared to stimuli with higher intensities. The authors used generalised estimating equations to ascertain whether these factors independently influenced participant’s ability to respond. It was found that the sensory modality and stimulus intensity were independent predictors of reproduced durations, as was the stimulus duration. Intriguingly, Indraccolo et al., (2016) also report an additional interaction between stimulus intensity and duration when auditory filled intervals were reproduced, suggesting these factors have a unique influence on the rate of the auditory internal clock. Additionally, that visual and auditory clocks operate at speeds different to one another (Indraccolo et al., 2016). This could be plausible if we consider that the auditory pacemaker-accumulator may run at a faster rate resulting in auditory stimuli accumulating more pulses compared to its visual counterpart for the same veridical duration (see Chapter 4). In Indraccolo et al.'s study however, it is unlikely that different processing speeds were the sole reason for differences between the modalities. For example, despite the fact that auditory stimuli were consistently responded to faster than visual stimuli, auditory stimuli were not consistently underestimated compared to their visual counterparts. These findings suggest that a number of factors influence our perception and production of time, including, but not limited to, the speed of processing, the speed of responding and the rate of activity for the internal clock (Indraccolo et al., 2016).

It has been documented that events during an interval influence the subjective perceived duration of that interval (Hasuo, Nakajima, Tomimatsu, Grondin, & Ueda, 2014) and that the addition of variability in a temporal event can result in a reduction in sensitivity (McAuley & Kidd, 1998). A much documented perceptual distortion is the 'filled duration illusion' which refers to filled intervals being consistently perceived as longer compared to empty intervals even when both interval types present identical physical durations (Hasuo et al., 2014). Strikingly, this illusion remains present and almost identical between auditory, tactile and visual sensory modes (Buffardi, 1971).

Whilst reports of the filled duration illusion are considerably robust (Hasuo et al., 2014), differences have been documented with different methodologies, particularly when comparing the method of adjustment with the method of magnitude estimation. The method of adjustment instructs subjects to directly compare two different time intervals presented in succession. The method of magnitude estimation on the other hand, presents each time interval and requires subjects to make singular temporal judgements after each one. It has been reported that individuals who show the filled duration illusion with one method may not always show the same effect with another method (Hasuo et al., 2014). Generally, the filled duration illusion was more likely to present when using a method of magnitude estimation compared to the method of adjustment. Nevertheless, when the method of adjustment was used the illusion was perceived clearly but only for a minority of subjects. For these same subjects, as the interval duration was lengthened, so was the magnitude of the filled duration illusion (Hasuo et al., 2014).

To investigate differences in duration discrimination of filled and unfilled intervals in relation to their standard durations, Rammsayer (Rammsayer, 2010) presented subjects with two auditory intervals in a two alternative forced-choice procedure in which the task required subjects to respond regarding which interval was longer in duration. Rammsayer compared standard durations between 50-1000ms and reported better discrimination for filled tones than empty tones, however this was only true for shorter durations, i.e. those at a standard interval of 50ms. Rammsayer concludes with the notion of a unitary timing mechanism that overlooks timing of both filled and unfilled intervals irrespective of standard durations for most intervals (Rammsayer, 2010).

Grondin (1993) adapted this method to a cross-modal design and assessed differences in duration discrimination of empty and filled intervals in both, auditory and visual modalities. Conflictingly, Grondin reports that for intervals around 250ms, both modalities show superior performance for unfilled intervals. For shorter intervals around 50ms, superior performance was shown for empty intervals but this was only for visual signals, as there was no difference in auditory discrimination. Differences in methodologies were also highlighted; as auditory discrimination was easier with the forced choice than the single stimulus method. In the last of this series of experiments, Grondin compared four marker-type conditions and their different thresholds by comparing filled and empty auditory and visual stimuli ranging from 125 to 4000 milliseconds of duration. Results report differences in thresholds for marker types particularly for short durations but that with longer durations these differences even out and a standardised Weber's Law holds true (Grondin, 1993).

Further support is provided in a later study by Grondin which suggests that discrimination is markedly improved with empty rather than filled intervals – again, however this result was limited to shorter intervals. This result was replicated in a second experiment employing a single stimulus method documenting the same result – that discrimination was improved with empty rather than filled intervals. This effect was present across auditory and visual modalities, and also to both 400ms and 800ms standard durations. Discrimination was also improved for the auditory compared to the visual mode (Grondin, Meilleur-Wells, Ouellette, & Macar, 1998).

Similarly, in further efforts to understand inter-sensory differences between filled and unfilled intervals, Goldstone and Goldfarb (1963) ran two separate experiments comparing judgements of filled versus unfilled intervals demarcated by lights and sounds. It was found that for both filled and unfilled durations, auditory intervals were judged to last longer, and in a comparison of filled versus unfilled durations, it was found that filled auditory durations were perceived longer than unfilled auditory durations (Goldstone & Goldfarb, 1963; Wearden, Goodson, & Foran, 2007).

Discrepant results have been presented by Rammsayer and Lima (1991) who report that when a standard duration of 50ms was used, filled intervals were discriminated more accurately than empty intervals. They explained this difference by a reliance on perceptual rather than cognitive processes. This was because performance remained unaffected when they used a simultaneous dual-task procedure (which would increase cognitive load) however performance was affected by increasing perceptual load in a backwards-masking task. These results are supported by other studies (Abel, 1972; Craig, 1973; Allan, 1979) who propose the discrimination of short filled intervals is better than that with short unfilled intervals.

The impact of stimulus length was further explored using neuropharmacology and longer intervals (Rammsayer, 1999). Using a range of pharmacological drug treatments to clarify the role of some neurotransmitters in timing, Rammsayer suggested that the temporal processing of longer intervals is regulated by working memory mechanisms, whereas the temporal processing of intervals that fall within the millisecond range (short intervals) depends on levels of dopaminergic activity in the basal ganglia (Rammsayer, 1999). An alternative view is that discriminating short unfilled intervals is inherently more difficult as unfilled intervals elicit lower sensory stimulation compared to filled intervals, consequently result in lower neural firing, higher uncertainty and less efficient responses (Rammsayer & Lima, 1991). A second alternative to why a sustained sound subjectively feels longer than an empty interval of the same duration (the ‘Sustained Sound Illusion’) may be due to pacemaker mechanisms. An internal pacemaker may become increasingly accelerated by a sustained sound resulting in its longer subjective duration (Repp & Marcus, 2010).

To add another dimension of complexity to this debate between filled and unfilled interval discrimination, it has been reported that the marker size can influence the perception of a visual empty interval (Ono & Kitazawa, 2009). As the size of the visual stimulus increases, so does the processing time needed for that stimulus. The authors demonstrate that empty intervals between the presentations of spatially large markers were perceived as lasting longer compared to smaller visual markers. Thus demonstrating that marker size can influence the perceived duration of an interval (Ono & Kitazawa, 2009).

Inconclusive results regarding differences in performance of filled and unfilled intervals can also be found in animal studies (Kraemer, Randall, & Brown, 1997; Santi, Miki, Hornyak, & Eidse, 2006; Macinnis, 2007). The considerable disagreement regarding the structure of temporal intervals is further reinforced by the internal marker hypothesis (Tse & Penney, 2006), that explains heightened performance for empty, compared to filled intervals, by asserting that an empty interval is timed for first marker offset to second marker onset. Alternative views suggest that timing is measured from first marker onset to the second market onset or that timing is measured from the first marker offset to second marker offset. Using EEG however, Tse and Penney found that the unconscious temporal processing occurred from the first marker offset to the second marker onset for empty intervals (Tse & Penney, 2006).

Rammsayer (2014) aimed to explore methodological differences with a set of experiments recording the influence of the type of task (2AFC or reminder task, (when the standard is presented before the comparison)), the type of interval (filled vs. empty), sensory modality (auditory vs. visual) and the base duration (from 100-1000ms) on duration discrimination performance (Rammsayer, 2014). Auditory (compared to visual intervals), and the reminder task (compared to the 2AFC) resulted in improved discrimination. Collective findings from these tasks demonstrate that performance levels are independently influenced by various factors, explicitly, the type of task, sensory modality and base duration.

4.8. Support for centralised timing mechanisms

A theoretical model distinguishes the variety of temporal information we receive into four distinct bands. These are: temporal synchrony, temporal duration, temporal rate and finally, rhythm (Lewkowicz, 2000). Lewkowicz suggests that the development of intersensory temporal perception “emerges in a sequential, hierarchical” order by building upon the previously acquired multisensory temporal processing skills (Lewkowicz, 2000). These findings are also extended to future sensory abilities. A psychophysical study exposing subjects to various temporal sequences has presented evidence that such exposure (to temporal sequences) improves the prediction of future events and also that this finding generalises to untrained stimulus durations (Baker, Dexter, Hardwicke, Goldstone, & Kourtzi, 2014).

Supporting evidence has been provided suggesting that the duration of an auditory event can influence the perception of a co-occurring visual signal (Romei, De Haas, Mok, & Driver, 2011). Romei et al. (2011) demonstrate that auditory and visual congruent streams resulted in higher sensitivity for visual duration discrimination (relative to visual-only conditions), but also that this effect failed to present when auditory and visual streams were presented asynchronously (Romei et al., 2011).

Moreover, psychophysical evidence from cross-modal asynchrony adaptation on perceived audio-visual, audio-tactile and visuo-tactile pairing and temporal order, has shown that after a brief period of repeated exposure to asynchrony in any of the aforementioned sensory pairings, results show marked changes in subsequent temporal order judgements. Simply put, the point of perceived simultaneity shifts towards the level of adaptation asynchrony. These results suggest a single supramodal mechanism may be responsible for the recalibration of multisensory time (Hanson, Heron, & Whitaker, 2008; Heron, Hanson, & Whitaker, 2009).

Additionally, the perception of visual apparent motion has been found to be modulated by both visual and auditory interval adaptation, suggesting a centralised timing mechanism irrespective of sensory modality that mediates visual apparent motion processing (Zhang, Chen, & Zhou, 2012). These findings support previous work establishing auditory dominance in temporal perception but also highlight the presence of centralised mechanisms underlying the complicated range of temporal signals we receive in everyday life.

4.9. Support for distributed timing mechanisms

Contrastingly, the distributed theory of timing posits that human sensory timing relies on mechanisms dedicated to each sensory modality and operating on a largely independent manner (see Chapter 2).

The lack of a cross-modal sensory transfer in perceived durations of visual and auditory stimuli was found as far back as the 1980’s (Walker, Irion, & Gordon, 1981; Walker & Scott, 1981). In recent decades, mounting evidence now supports this hypothesis using a combination of sophisticated methods. To investigate underlying cortical processing, Kaya et al. (2017) adapted subjects to sub-second intervals of visual and auditory signals and later assessed evoked activity of visual apparent motion. Interestingly, they found opposing effects of each modality to apparent visual motion, despite both showing significant changes in the Event-Related potential (ERPs). Furthermore, these changes occurred in different scalp areas (parieto and parieto-central for auditory and occipital, parieto-occipital for visual) and also in different temporal windows (Kaya, Yildirim, & Kafaligonul, 2017).

Evidence for timing mechanisms that are sensory-specific has even been extended to the memory components of these clocks (Gamache & Grondin, 2010). Furthermore, efforts to equate visual perceptual echoes to potential counterparts in the auditory domain have often failed (İlhan & VanRullen, 2012). The authors suggest that the magnitude of cortical oscillations in early visual and auditory sensory processes are not equated and also that alpha band oscillations (8-13Hz) maintain a unique role in visual perception (İlhan & VanRullen, 2012).

Investigations of cross-modal duration perception have presented subjects with various unimodal audio or visual durations, requiring subjects to then report which was longer (Klink, Montijn, & van Wezel, 2011). Notably, the authors presented unimodal target events that were accompanied by distractor stimuli in another modality. The authors note an asymmetrical relationship between the modalities. Specifically, they note that irrelevant auditory temporal information was able to influence the duration estimation of visual stimuli, yet the same visual stimuli was unable to distort auditory duration perception (Klink et al., 2011).

In another study employing a temporal reproduction paradigm and expert percussionists, it was found that temporal rates were reproduced veridically for all auditory stimuli in groups of expert drummers and string musicians. However, notably, only expert drummers were also able to reproduce temporal rate of visual stimulus (presented as brief flashes) (Cicchini, Arrighi, Cecchetti, Giusti, & Burr, 2012). This suggests that it was only this group with specific and extended training that were able to transfer temporal sensitivity from one modality to another.

To further clarify differences in sensory timing, positional sensitivity has been investigated for the visual duration after-effect (Li, Yuan, Chen, Liu, & Huang, 2015). Using an adaptation procedure exclusive to the visual modality, the authors found that the duration after-effect transferred across hemi-fields and was not dependent upon those hemi-fields, suggesting position invariance as a feature of the visual duration after-effect. They further suggest that the absence of spatial specificity implies that duration processing mechanisms in the visual system may lie at a later stage of the processing hierarchy (Li, Yuan, & Huang, 2015). This result has been replicated recently where the transfer of the visual duration after-effect was found to occur across more than 10° of the visual angle. As with Li et al. (2015), Maarseveen et al., (2017) also found that transfer occurs within and across hemifields and similarly concur that the duration encoding occurs at an ensuing stage of the visual processing hierarchy (Li, Chen, Xiao, Liu, & Huang, 2017; Maarseveen, Hogendoorn, Verstraten, & Paffen, 2017)). Evidence of a similar after-effect occurring in the tactile modality with vibro-tactile adaptation has also been documented (Watanabe, Amemiya, Nishida, & Johnston, 2010).

The common timing hypothesis assumes a centralised, amodal timing mechanism overlooking timing information presented to each sensory modality, irrespective of interval duration. Contrastingly, the distinct timing hypothesis suggests two dissociable mechanisms, one that times intervals in the sub-second range and the other that times signals in the supra-second range (Rammsayer et al., 2015). Evidence is now being collated to suggest a “gradual transition” from a modality-specific and sensory automatic timing mechanism to a more cognitively-mediated and amodal timing mechanism. The authors suggest that the window of transition allows for both mechanisms to operate simultaneously, however the influence of sensory-automatic timing mechanisms gradually increase with interval durations and become the dominant processing mechanism (Rammsayer et al., 2015). These findings paired with Stauffer et al., (2012) (see chapter 2) suggest a hierarchical processing architecture that aligns previous work and organises how human brains process temporal input across the scales of physical time and also sensory modalities.

4.10. Neural differences across the senses

It is known that the transduction times for sensory stimuli follow divergent patterns, for example, transduction patterns evoked after visual stimulation are a chemical reaction and much slower than the mechanical transduction engaged in the somatosensory system (Spence & Squire, 2003).

Furthermore, evidence exists to suggest faster reaction times in audition compared to those for vision in the same task (Brebner, 1980). Several theories exist to explain this difference. One such theory is linked to the pacemaker-accumulator theories of timing and specifically, that the pacemaker in auditory cases emits faster impulses for auditory compared to visual stimulus, ultimately resulting in faster processing of auditory compared to visual stimulus. These faster pacemaker pulses are underlain by sensory transduction times. A physiological explanation provided by the authors is that neurally, the central processing stage for auditory information is shorter than that for visual information, resulting in faster reaction times for auditory compared to visual times. This is due to cortical architecture where the primary visual cortex is located in the occipital lobe whereas the temporal lobe is the nucleus for the auditory cortex (Pinel, 2006). Thus, in a comparison of distance between sensory receptors to the primary sensory areas, visual information has a much longer distance to travel. These faster transduction times then result in the pacemaker emitting faster pulses for auditory signals and ultimately, results in faster reaction times for audition when compared to other sensory modalities.

4.11. Timing in clinical populations

The timeframe regarding whether two stimuli presentations will be perceived as simultaneous is around 40ms depending on the modality (Exner, 1875; Herzog, Kammer, & Scharnowski, 2016). It has been noted however, that in certain patient groups, this window of integration is considerably longer, for example in individuals with Schizophrenia (Giersch et al., 2015). This patient group also reports heavily fragmented streams of perception (Giersch et al., 2015; Herzog et al., 2016) indicating serious functional deficits in temporal sensitivity and perception. Whilst isolated deficits in temporal processing (with no other sensory or cognitive deficits) have not been identified in a single patient group to date, inferences regarding temporal processing in other clinical populations may still offer useful insights regarding both, the structural and functional features of time perception.

It is no surprise that our sensory environments naturally possess temporal characteristics (Binetti, Lecce, & Doricchi, 2012), regularly requiring constant real-time motor interaction. These features make it likely that temporal and spatial information may be processed by the same circuits simultaneously (Mauk & Buonomano, 2004). No performance difference was observed between individuals with Parkinson's disease compared to neurotypical controls on a repetitive tapping task, in one of the earliest explorations of this method on clinical populations (Ivry & Keele, 1989). More recent evidence however, has found that responses are more variable in inter-tap intervals for individuals with Huntington's or Parkinson's (compared to neurotypical controls) (Freeman et al., 1996; O'Boyle, Freeman, & Cody, 1996).

A rare case study into the neuropsychology of timing comes from the case of H.M., a patient who underwent a resection to the bilateral medial temporal lobe which, following surgery resulted in considerable memory loss. In tasks where H.M. was to reproduce sub-second durations (between 0.3-1 second), he performed with fairly accurate timing up to durations of 20 seconds. After this point (20s) however, H.M. systematically underestimated durations (Eisler & Eisler, 2001; Meck, 2005), suggesting again the distinction between sub-second and supra-second temporal processing (and their reliance upon different processing mechanisms). Nevertheless, estimating physical time subjectively has been demonstrated as a considerably stable and robust function in non-clinical populations, and deviating from veridicality only in conditions where the individual has suffered cortical trauma, severe psychiatric disorders, brain pathology or toxicological/pharmacological challenges (Meck, 1996, 2005).

4.12. Other considerations: Samples and their size

Typical psychophysical studies consist of long experimental testing sessions. While naïve subjects still replicate the after-effects presented by more experienced subjects, they often fail to respond consistently in long sessions. It is for this reason that the majority of subjects for each experiment were derived from psychophysically-experienced samples. The purpose of naïve subjects was to assess the presence of similar after-effects in less experienced populations.

To address specific challenges experienced by naïve subjects, all experimental testing was conducted in several blocks to prohibit any fatigue effects from impacting performance. Naïve subjects also gathered data on a larger number of trials so as to ensure that any effects were being consistently presented and not just an artefact of that testing block. Furthermore, as results for all experiments were analysed using a within-subjects approach, the impact of experience (or lack of) from naïve subjects was minimal.

Statements about time commonly occur in popular culture, for instance, “time flies when you are having fun”. Regarding, age-related effects on the perception of time, another commonly held view is that time progresses faster as we age (Ferreira et al., 2016, Wittmann & Lehnhoff, 2005). The challenges of truly investigating the effect of age on time perception is that with age, many other deficits also become apparent (such as declining cognitive and memory functions).

Nevertheless, chronological age has also been demonstrated as an inconsequential factor on influencing brief intervals of time (up to a period of seconds) (Hancock & Rausch, 2010). Moreover, age differences in the temporal window of integration and performance on timing tasks in the sub-second to minutes range are typically either subtle or non-existent (Horvath et al., 2007; Rammsayer, Lima & Vogel, 1993). In many cases, any age differences in timing that do exist are a result of differences in other cognitive functions such as working memory and attention (Krampe et al., 2002; Wittmann and Lehnhoff, 2005; Ulbrich et al., 2007; Bartholomew et al., 2015).

General explanations of why time does appear to slow down as we age have been provided by the theory of a slower internal clock and particularly, that this clock takes longer to recover compared to when humans are younger. This “fatigue effect” has been explained through a slow depletion of striatal dopamine as a function of sustained cognitive engagement during skill learning acquisition (Kawashima et al., 2012). This effect is further facilitated by dopamine-related disorders such as typical aging, Parkinson’s and Huntington’s diseases (Malpani et al., 1998; Meck, 2006; Allman & Meck, 2011; Gu et al., 2015). This rapid depletion in dopamine function co-occurs with the sense that the external world is going faster, which may actually be due to our internal clock going slower, and ultimately leading to the sense that a sequence of events are occurring in a shorter temporal window than you would normally expect. Neural explanations underlying age-related differences in time perception arise from fundamental changes in the functioning of the cortico-thalamic-basal ganglia circuits. Specifically, that oscillators within these circuits become increasingly variable and therefore less reliable with age (Allman & Meck, 2012).

For the studies conducted within this thesis, one subject consistently recruited for all experiments (subject DW) was of an older age bracket than of subject AM or other naïve subjects. Subject DW however, consistently reproduced effects that aligned with other participants. The most notable difference was decreased variability in subject DW’s responses which can be expounded by the increased familiarity with psychophysical tasks (rather than age).

In terms of sample sizes, it has been noted that, “it is more useful to study one animal for 1000 hours than to study 1000 animals for one hour” (Skinner, 1938 in Kerlinger, 1999). The justification behind the sample sizes used in the following experiments is two-fold. Firstly, that in studies using small sample sizes, the individual is treated as the replication unit; with each repeated trial effectively acting as a second observation/data point. Resulting models derived from such designs ensure that the functional relationships observed at the individual level can then effectively be employed in such models that can readily be applied to other cohorts of individuals – regardless of the sample size. We therefore, selected the sample size because it is typical of psychophysical studies in which researchers seek to thoroughly characterize the performance of each subject separately, using far more trials than are typical for psychological studies that employ group averages. The second advantage of such designs is that this avoids replicability issues by building in several independent replications (i.e. 1n = 1 replication) (Hickok et al., 2018).

Additionally, from a historical point of view, studies employing small-N designs have produced results that have been consistently replicated and maintained robustness of results (Smith & Little, 2018). Moreover, many of the experiments expanded upon in this thesis have gathered data on subjects who volunteered for substantially long periods of time – especially in experiments utilising psychophysical adaptation, therefore from a practical point of view, and given the time constraints, it would not have been realistically feasible to recruit a much larger sample size.

5. Assessing the modality-specificity of the rhythm after-effect

The work presented within this chapter has been peer-reviewed and published - ***Motala, A.,*** *Heron, J., McGraw, P. V., Roach, N. W., & Whitaker, D. (2018). Rate after-effects fail to transfer cross-modally: Evidence for distributed sensory timing mechanisms. Scientific Reports, 8(1), 924. doi: 10.1038/s41598-018-19218-z*

Human behaviour is fundamentally reliant upon accurate time perception, for example, for speech and planning movement. A current debate in the field concerns the modality-specific nature of temporal processing. Specifically, it remains unresolved whether sensory time perception is mediated by a central timing component regulating all sensory modalities, or by a set of distributed mechanisms, each dedicated to a single sensory modality and operating in a largely independent manner (Figure 5.1).

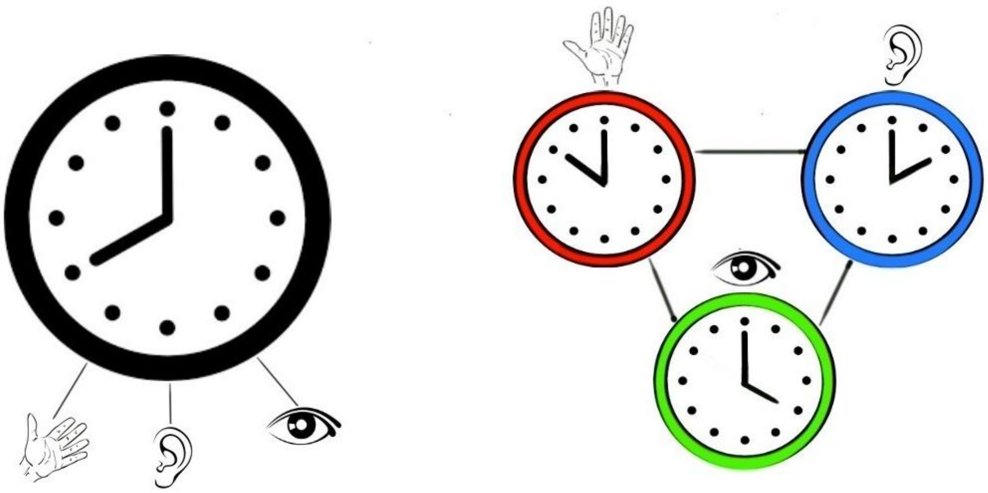


Figure 5.1. Schematic depiction of a centralised timing mechanism (left), compared to a distributed timing network (right). See text for more detail.

5.1 Cross-modal influences on rate and duration perception

Evidence from experiments exploring rate perception demonstrates that when auditory and visual stimuli are presented concurrently, auditory stimuli are able to bias subsequent judgements of visual flicker stimuli (Shipley, 1964; Recanzone, 2003). In studies exploring duration perception however, it has been found that repeated exposure to stimuli of particular durations elicits contingent after-effects present only in unimodal conditions, (and absent in all co-localised cross-modal conditions) (Walker, Irion, & Gordon, 1981). In an investigation of peri- and supra-second durations of both auditory and visual stimuli, it was reported that opposite distortions were perceived in interval timing (Yuasa & Yotsumoto, 2015). Using a combination of two-alternative forced-choice and cross-modal stimuli presentations varying in their temporal features (either flickering/fluttering or continuous), four key findings were reported – firstly, that auditory flutter presentations resulted in underestimated perceived durations (temporal compression), whereas contrastingly, visual flicker presentations resulted in an overestimation of perceived durations (temporal dilation). Furthermore, when both auditory flutters and visual flickers were presented simultaneously, perceived distortions evidenced previously were cancelled out. Lastly, that when the temporal nature of stimuli was altered, that is, auditory flutters were presented with visual stimuli that were constantly presented (and not flickering as before), judgements of visual stimuli were influenced by simultaneously presented auditory flutters – even when participants had been explicitly instructed to ignore the auditory flutters (Yuasa & Yotsumoto, 2015). This suggests that the aforementioned effects were not governed by either attention or presentation order (Yuasa & Yotsumoto, 2015). This finding may perhaps be expected considering previous reports of auditory dominance in time perception (Shipley, 1964; Chen & Yeh, 2009; Bueti & Macaluso, 2010; Grondin & Ulrich, 2011; Li, Yuan, & Huang, 2015).

5.2 Modality-specific constraints on sub and supra-second duration perception

Yuasa and Yotsumoto (2015) further suggest that visual and auditory modalities process peri-and supra-second durations differently, and that these differences may arise from key boundary asymmetries for automatic and cognitive processing Specifically, it is understood that peri-second durations are processed automatically whereas longer durations are processed with a heavier focus on cognitive mechanisms (Lewis & Miall, 2003). Thus, the differences evidenced between the auditory and visual modalities may present as a result of automatic and cognitive processing being defined differently within these modalities (Yuasa & Yotsumoto, 2015). These findings also present further evidence that different modalities employ different processing strategies to time and further support the distributed theory of time. Yuasa and Yotsumoto suggest independent timing mechanisms that govern auditory and visual processing separately but assert that some level of interaction exists between these two systems (Yuasa & Yotsumoto, 2015). Despite not being explicitly related to rate, the study of duration perception has found similar support for distributed timing mechanisms. Using simultaneous sensory adaptation where subjects are presented with two distinct durations defined by two separate auditory and visual stimuli, it was found that the subjectively perceived duration after-effect is modality-dependent, and was also conditional upon the auditory frequency (in audition) but not on orientation (in vision) (Li et al., 2015). These findings suggest one of two things; first, that the after-effects of perceived duration are mediated by modality-specific timers dedicated to each different sensory modality. Secondly, that the timer for the visual modality may be located at a later stage of processing than the auditory timer (Li et al., 2015).

Moreover, modality specific effects for the discrimination of empty time intervals using audition and vision has also been reported (Hocherman & Ben-Dov, 1979). Specifically, subjects were presented with two empty intervals defined by three successive stimuli (leaving combinations of AAA, VVV, AAV, VAA, AVA, VAV, AVV, AAV, and VVA which would define two intervals) on each trial and asked to report which of the two intervals was longer. Modality-independent performance was found when the first two stimuli were of the same sensory modes. Furthermore, strong response biases were elicited as a result of introducing a different modality for either the first or second stimulus (Hocherman & Ben-Dov, 1979).

5.3 The centralised versus distributed debate

The judgement of temporal rate has caused considerable discord in previous literature. According to Levitan and colleagues (Levitan, Ban, Stiles, & Shimojo, 2015), psychophysical adaptation to specific temporal rates elicits repulsive after-effects similar to those evidenced with motion, orientation and other visual features. Specifically, after being exposed to a 5Hz temporal frequency presentation, a 3Hz presentation appears slower (that what it would have pre-adaptation). This effect occurs bi-directionally in the sense that after being exposed to a 1Hz temporal frequency presentation, the same 3Hz presentation now appears much faster. Crucially, Levitan and colleagues used a design employing visual and auditory stimuli and found evidence for cross-modal rate perception which they use to suggest a unified, multisensory theory of timing. This lies at odds with evidence proposed by Becker and Rasmussen (Becker & Rasmussen, 2007) who despite using a similar method, found that after adapting to a specific rate in one sensory modality, effects transfer within the modality (for example, from one ear to the other not adapted to), yet fail to transfer cross-modally (Becker & Rasmussen, 2007) suggesting distinct and independent sensory timing mechanisms. Evidence against a centralised supramodal clock also comes from arguments that even short visual events are encoded via visual neural mechanisms with localised receptive fields, rather than an overarching clock mechanism (Burr, Tozzi, & Morrone, 2007). It is important to note that despite the differences evidenced in the modality-specificity of rhythm after-effects, both Levitan et al. (2015) and Becker and Rasmussen (2007) found evidence that these after-effects were band-limited. This means that the classic adaptation result in terms of rebound after-effects, disappear if the adapting stimuli and test stimuli become too dissimilar.

There are important implications of this debate in understanding the physiological basis of human time perception. A centralised timing mechanism, for instance, purports a single timing mechanism overlooking each sensory modality and implies that sensory processing is disconnected from time estimation (Bruno & Cicchini, 2016). Evidence supporting modality-specific timing mechanisms suggests that to a certain extent, these mechanisms are also tied to other sensory characteristics such as spatial location and visual hemi-fields (Li, Yuan, Chen, Liu, & Huang, 2015; Bruno & Cicchini, 2016). Given this disagreement in the literature, we are still unsure about the centralised or distributed nature of timing. To this end, we used temporal adaptation across three senses and a method of rate reproduction in an attempt to clarify whether this type of after-effect is unimodal or can transfer across the senses.

We employ a paradigm of sensory adaptation (Chen & Zhou, 2014) to investigate sensory time perception using rate reproduction. This involves a rate adaptation paradigm where subjects are presented with a range of rates in an adaptation phase to either the visual, auditory or tactile modalities. The use of the tactile modality is particularly relevant as tactile rhythms have largely been neglected in investigations of human time perception (Jokiniemi, Raisamo, Lylykangas, & Surakka, 2008). Subjects are then presented with a test phase to either the same or a different modality however, critically, this stimulus is always presented at 3Hz for each condition. The final phase of testing requires subjects to reproduce the rate presented during the test phase (3Hz) by tapping on a response device. A total of 9 unimodal and cross-modal pairings are tested, allowing for a comparison of rate perception as a result of rate adaptation for these pairings.

There are several features distinguishing the current experiment from those conducted by Becker and Rasmussen and Levitan and colleagues. Firstly, in comparison to Becker and Rasmussen, a much wider range of frequencies was used, for instance, they used a range of 1.4-3.33Hz, with the control condition consistently presenting a frequency of 2.5Hz. The present experiments use a range of adapting frequencies spanning from 1.05-8.46Hz and our test frequency was slightly faster at 3Hz. In addition to this, we also test a third modality of touch. In comparison to Levitan and colleagues, we use an entirely different method of rate reproduction. Ultimately, we aim to use rate reproduction which hitherto has been relatively underexplored to disentangle the underlying mechanisms of sensory time perception and particularly whether these rely on centralised or distributed mechanisms. There are two potential outcomes to this experiment. The first is that participants will be able to adapt to unimodal rhythms but that resulting after-effects will only present when the adapting and test modalities match. The alternative is that participants will be able to adapt to unimodal rhythms and demonstrate consequent after-effects with congruent and incongruent test modalities demonstrating amodal rate after-effects.

5.4. Methodology

Subjects 3 participants (2 female and 1 male, mean age = 33, standard deviation = 14 years) were used, with self-reported normal hearing and visual abilities. One participant was fully naïve (YL) to the purpose of the experiment. The experiments received ethical approval from the Research Ethics Committee at the School of Optometry and Vision Sciences, University of Cardiff and were conducted in laboratory facilities at the Clinical Research Facility (CRF) located in the School of Optometry and Visual Sciences at Cardiff University.

Stimulus Parameters

General stimulus set-up:

All visual stimuli were temporally constrained by a monitor frame rate of 60Hz. This means that a single frame lasted for 16ms and that was the shortest possible temporal duration we were able to produce. The corresponding auditory stimuli were constrained by a sound card with a sample rate of 44,100Hz however, the duration of a single beep was kept identical to a single flash. The signal for the tactor was also produced using the same sound card to produce a single tap. The following section will elaborate on the stimuli and procedure in more detail. To ensure no millisecond timing errors were resulting from our experimental set up and albeit commonly-used software (Plant & Quinlan, 2013), all program timings were verified using a dual-channel oscilloscope.

Visual stimulus:

Visual stimuli were presented on an Eizo EV2436W monitor driven by an Intel ® Core ™ i5-4460 desktop computer running Microsoft Windows 7. Stimuli were generated using MATLAB 8.6 (Mathworks, USA) and Psychophysics Toolbox 3 (http://www.psychtoolbox.org).

Stimuli consisted of bright white circular flashes of 274 cd/m^2 luminance located centrally at 60 cm viewing distance displayed against a uniform black background at 0.32cd/m^2 mean luminance for a duration of 16ms, presented at varying rates of temporal frequency. Visual stimuli had a diameter of 10.5cm.

During all non-visual sensory presentation, the screen was kept uniformly black at 0.32cd/m^2 mean luminance.

Auditory stimulus:

Auditory stimuli remained constant at a sampling rate of 44,100 KHz. Stimuli constituted of clicks (of durations lasting 16ms) of white noise presented at either 1.06, 1.5, 2.12, 3, 4.24, 6 or 8.46Hz, using Sennheiser HD280 Pro Headphones at an SPL of 70dB. The loudness of auditory stimuli was kept constant throughout the experiment and was set to roughly 65dB.

Tactile stimulus:

Tactile stimuli were square waveforms generated using the ‘audio-out’ voltage and using a Dancer Design Tactor – a miniature electromagnetic solenoid-type stimulator. Tactile stimuli constituted of ‘taps’, each presented for a duration of 16ms and programmed using the same sound card used to present auditory stimuli. As the tactor produced a slight auditory feedback, tactile stimuli were presented alongside white noise (set to roughly 65dB) to mask the sound of the tactor.

Figure 5.2a (left) and 5.2b (right). Figure 5.2a depicts the tactor used to produce tactile taps in all tactile conditions. Figure 5.2b depicts the response disk that subjects were instructed to tap on to reproduce the test temporal frequency.

To mask the sound of tapping during response periods and to reduce the possibility of auditory-feedback, white noise (at the same loudness as other auditory stimuli) was presented during the response period for all trials. To further eliminate auditory confounds, a fabric occluder was used to mask the tactor and subject's hand during all adaptation and test periods using tactile stimulation. Subjects were also explicitly instructed to not watch their finger tap the response device during response periods.

Procedure

Subjects were shown a grey screen and instructed to press the space bar on a keyboard when they were ready to begin.

The experimental trials began with an adaptation period of one randomly assigned temporal frequency to sequences of either auditory, visual or tactile stimuli presented for a duration of 8-10 seconds. Temporal frequencies ranged from 1.06 - 8.46Hz and were spaced in log steps of .15 log units for the adaptation phase. Only one of these was chosen as the adapting frequency. This phase preceded the test phase (after a pause of 400ms), where stimuli were presented at 3Hz for each condition to either the same or a different modality for a period of 2.5-3 seconds. Critically, the test stimuli were presented at 3Hz for each trial irrespective of the modality it was presented to. Finally, the response phase lasting 2 seconds ensued where the participant was instructed to reproduce the test frequency by tapping on a response disk with their index finger (while white noise was played to mask the sound of their response tapping) (see Figure 5.3).

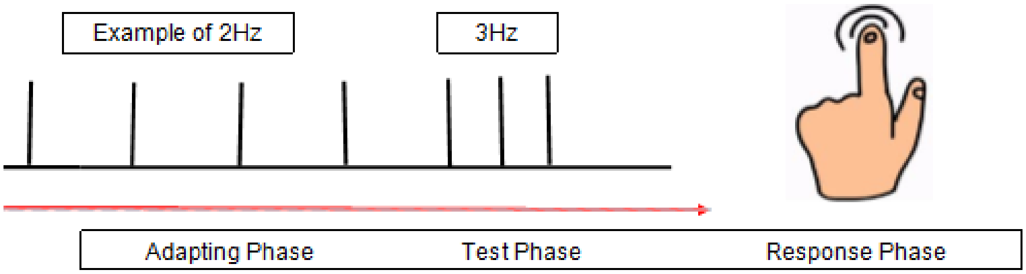


Figure 5.3. Schematic simplifying the experimental set-up described previously. Critically, the test phase remained constant at 3Hz throughout all conditions to allow for a comparison of rate reproduction post-adaptation, and thus measuring the change from baseline. Response phase drawing accessed from www.iconsmind.com.

Different sensory combinations constituted different experimental blocks (such that AV was considered one block and VV another), and the frequencies adapted to within these blocks were randomized. Participants completed each block in random order. This was kept consistent throughout all experiments and subjects in this chapter. In addition, a break of 3 minutes was inserted between each trial period to ensure no adaptation effects crossed-over to subsequent experimental testing. Subject DW repeated each temporal frequency in each sensory pairing 3 times producing data from 189 trials (21 sets of data for each condition) whereas subjects AM and YL repeated each temporal frequency in each sensory pairing 5 times producing data from 315 trials (35 sets of data for each condition).

In timing research, it is fundamental to use responses that are time-limited in their nature. The particular advantage of using a response disk over a spacebar is the increased temporal precision as there is no secondary upwards pressure (as there would be with a spacebar). To further limit any confounding effects, each adapting temporal frequency was tested with each possible modality and tested across either the same, or a different modality for each possible combination of adapting and test modality pairs. Additionally, all testing was randomised at each level to further control for any such effects. Lastly, in extensive efforts to eliminate the possibility of auditory feedback, white noise was played via the headphones throughout the tapping response phase (Wearden, 2003).

5.5 Results

Mean reproduction values and corresponding standard error of the mean were calculated for each adapting temporal frequency for each multisensory combination. These means (and standard errors) were used to plot the data.

A best-fitting curve was fitted to the data to extract relevant parameters such as the magnitude and spread of any adaptation effects. The curve was based on the first derivative of a Gaussian (and fit using the method of least squares and the Levenberg-Marquardt algorithm), namely

$$F_{matching} = F_{mean} + \left[\left(A * log\left(\frac{F_{adapt}}{F_{test}}\right)\right) * e^{-\left(\frac{\left(log\left(\frac{F_{adapt}}{F_{test}}\right)\right)^{2}}{2\sigma^{2}}\right)} \right]$$

where F_{adapt} is the adaption rate, σ is the standard deviation (width) of the Gaussian, A is a constant related to the amplitude of the function, and F_{test} is the rate of the test stimulus (3Hz). Note that when $F_{adapt} = F_{test}$ the equation within the brackets becomes zero, and the matching frequency becomes F_{mean}, the mean vertical height of the function. The maxima and minima of this function occur at adapting rates $\pm\sigma$ log units from the origin, i.e. $\log(F_{adapt}/F_{mean}) = \pm\sigma$. The half-amplitude of this function (μ), which represents the magnitude by which the matching rate deviates from F_{mean} (i.e. the size of any illusion), is therefore given by

$$\mu = A * \sigma * e - \frac{1}{2}$$

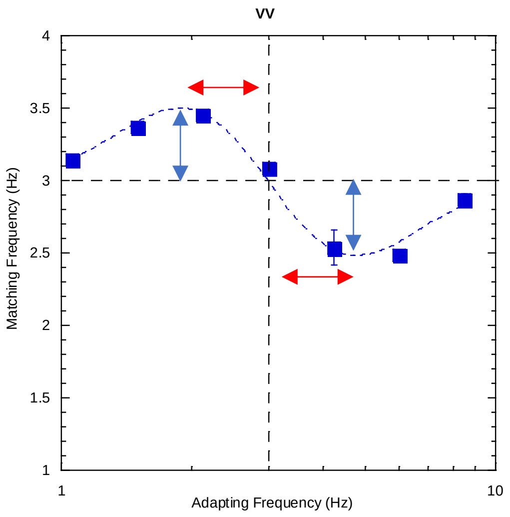


Figure 5.4. Subject DW's mean response values for the unimodal visual condition where the subject was exposed to a range of adapting temporal frequencies in the visual modality and the test stimulus presented was also visual. The x-axis demonstrates the range of temporal frequencies that the subject adapted to and the y-axis demonstrates the rates reproduced by the subject. Vertical arrows indicate the amplitude of the effect and the horizontal arrow indicates the overall spread of effect.

Figure 5.4 shows the results of a sample unimodal condition where subject DW adapted to a visual rhythm and was presented with a visual test rhythm to reproduce. The plot demonstrates marked perceptual responses after adapting to rhythms slower than 3Hz and bi-directionally reflected when adapting to rhythms faster than 3Hz. Specifically, focusing to the left of the vertical midline, it can be observed that after adapting to frequencies slower than 3Hz, this subject reproduces the 3Hz test rhythm as 0.5Hz faster than 3Hz. Similarly, to the right of the same midline the opposite effect can be observed – that adapting to a faster frequency than 3Hz results in the same 3Hz test rhythm now being perceived as much slower (again, around 0.5Hz slower). Whilst this type of effect is seen for every unimodal condition; it remains clearly absent in all cross-modal conditions (Figures 5.5, 5.6, 5.7 and 5.8).

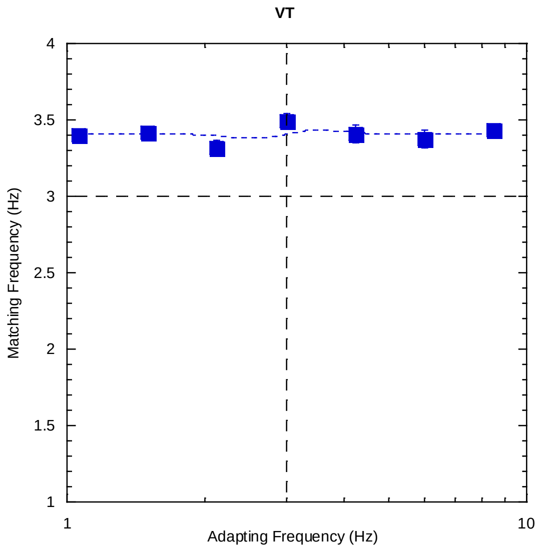


Figure 5.5. Subject DW's mean response values for the visuo-tactile cross-modal condition where the subject was exposed to a range of adapting temporal frequencies in the visual modality and the test stimulus presented at 3Hz in the tactile modality. The x-axis demonstrates the range of temporal frequencies that the subject adapted to and the y-axis demonstrates the rates reproduced by the subject.

Figure 5.5 demonstrates the results from a sample cross-modal condition where subject DW adapted to a range of visual rhythms, but was presented with a tactile test rhythm of 3Hz to reproduce. In the above plot, a clear lack of response bias is observed when subject DW is exposed to a tactile test rhythm after having adapted visually. The lack of change from the baseline is critical. Data presented here indicate that adapting to a given rate in one modality has no effect on the perception of rate of the test modality, if the adapting and test modalities are incongruent. This suggests no cross-modal transfer or rebound after-effects as experienced for the unimodal conditions.

All plots for each subject and each condition are located below, plots from subject AM are presented before plots for subject DW, followed by subject YL. Each plot is labelled with the sensory pairing of the data presented. Specifically, the first letter denotes the adapting modality and the second letter denotes the testing modality. The order of plots presented corresponds to the sensory pairing key below:

TA	TV	TT
VA	VV	VT
AA	AV	AT

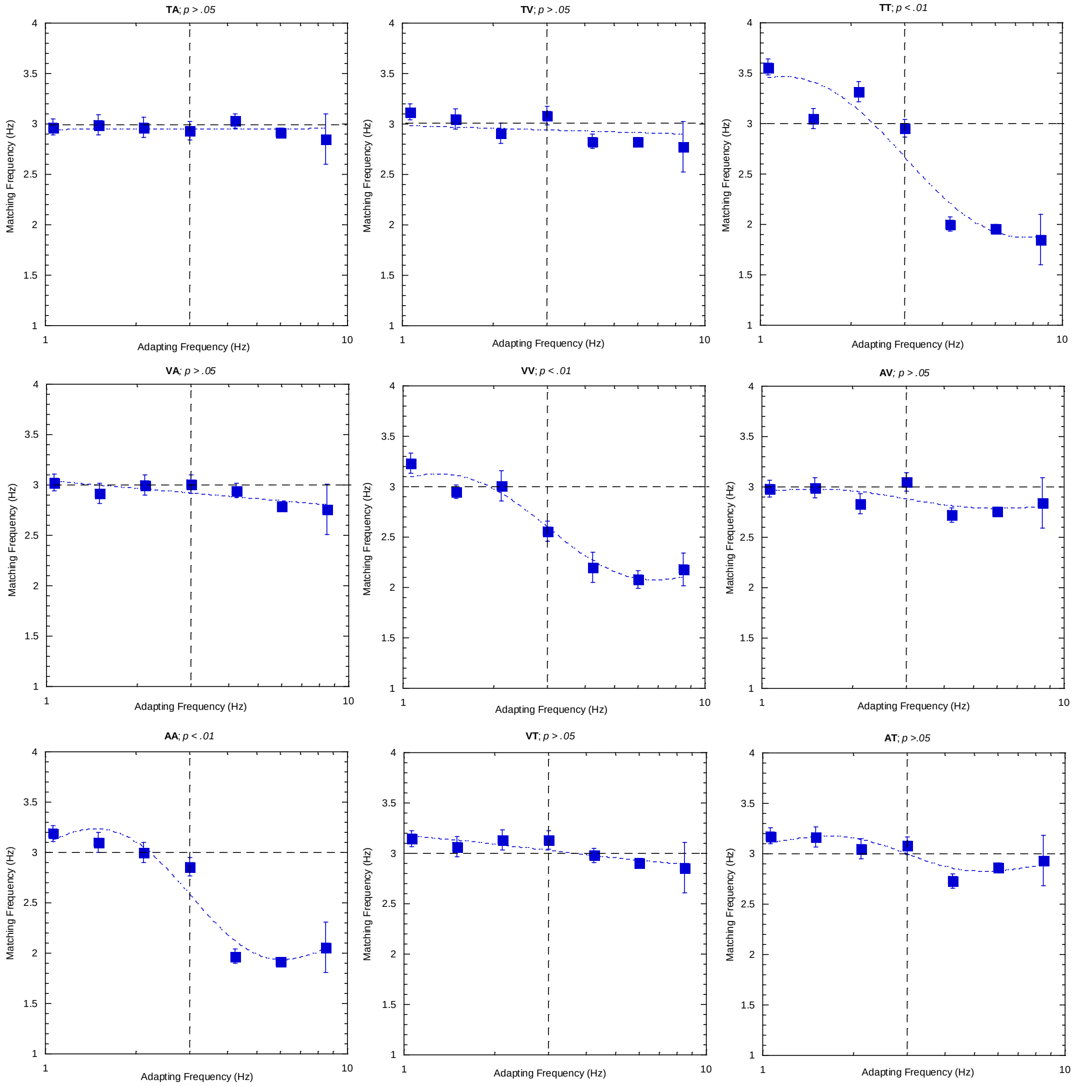


Figure 5.6. All cross-modal and unimodal plots from subject AM. The key above indicates all sensory pairings; specifically the letter ‘A’ refers to the auditory modality, ‘T’ to the tactile modality and ‘V’ to the visual modality.

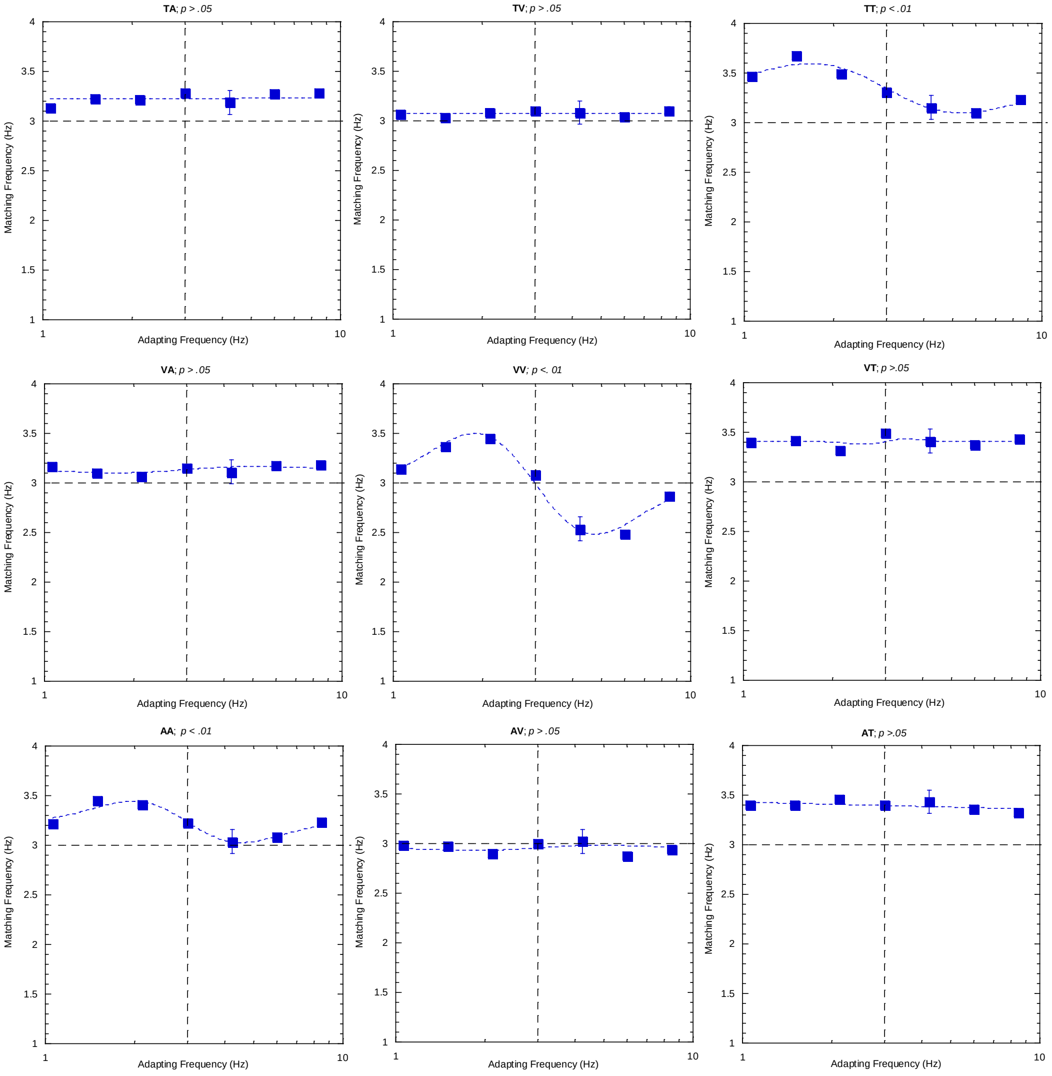


Figure 5.7. All cross-modal and unimodal plots from subject DW. The key above indicates all sensory pairings; specifically the letter ‘A’ refers to the auditory modality, ‘T’ to the tactile modality and ‘V’ to the visual modality.

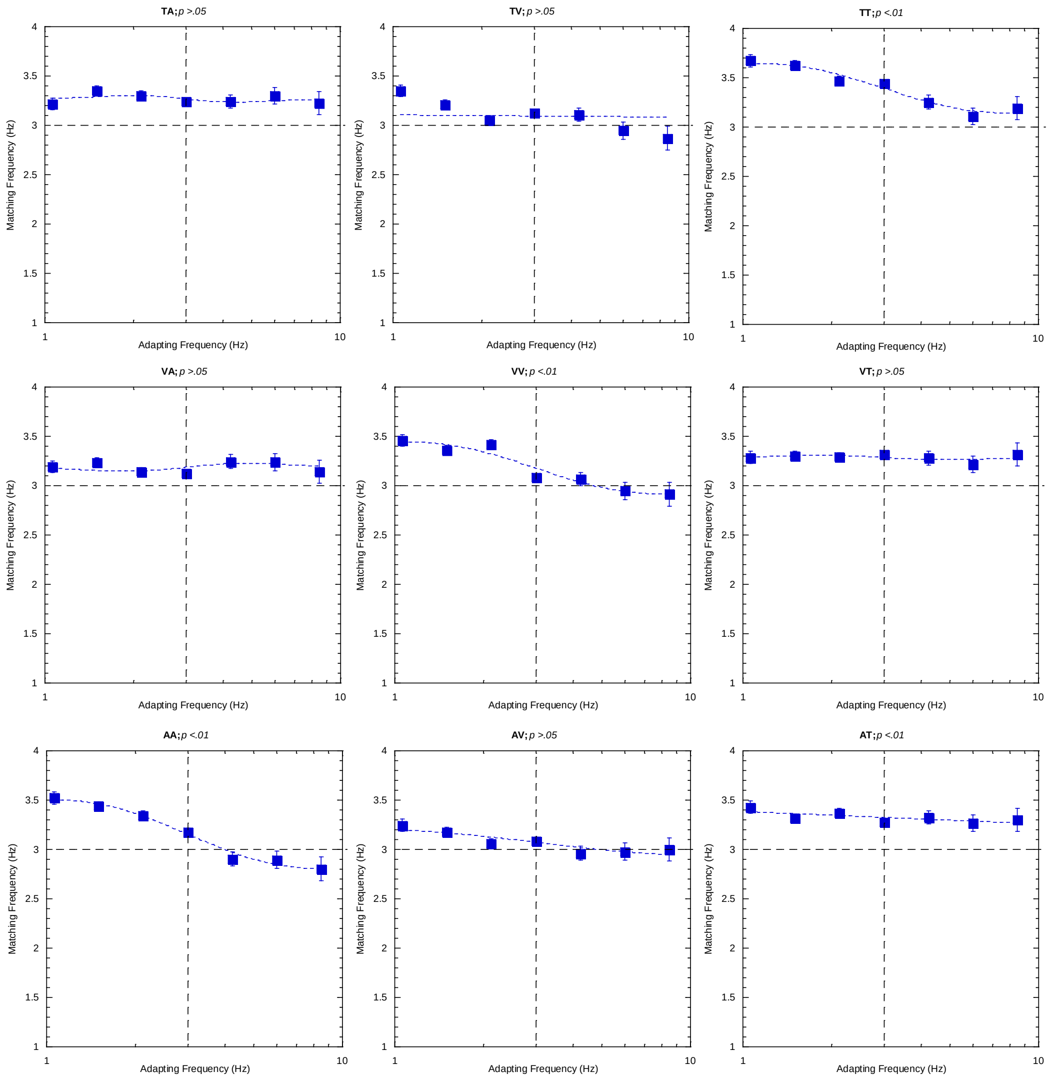


Figure 5.8. All cross-modal and unimodal plots from subject YL. The key above indicates all sensory pairings; specifically the letter ‘A’ refers to the auditory modality, ‘T’ to the tactile modality and ‘V’ to the visual modality.

Effect sizes were then calculated by dividing the µ value by the µ error value for each condition. A two-tailed, one-sample t-test (df= 6) was then conducted for each subject in each condition, Holm-Bonferroni adjusted p-values of which are presented below:

Condition	**Subject AM**	**Subject DW**	**Subject YL**
AA	.002*	<.001*	<.001*
TT	.010*	<.001*	.002*
VV	<.001*	<.001*	.009*
AT	.096	1	1
TA	1	1	1
AV	.740	1	1
VA	1	1	1
TV	1	1	1
VT	1	1	1

(Values marked with an asterisk (*) signify results of statistical significance).

Table 5.1: Adjusted p-values for all subjects across all conditions. The first letter denotes the adapting modality and the second letter denotes the testing modality, ‘A’ refers to the auditory modality, ‘T’ to the tactile modality and ‘V’ to the visual modality. Results for all unimodal conditions are reported first, followed by results for all cross-modal conditions.

	Amplitude of effect – AM (Hz)	**Amplitude of effect – DW (Hz)**	**Amplitude of effect – YL (Hz)**
AA	0.65 *(.09)*	0.21 *(.03)*	0.35 *(.03)*
TT	0.8 *(.14)*	0.25 *(.03)*	0.25 *(.03)*
VV	0.53 *(.06)*	0.51 *(.04)*	0.26 *(.05)*

Table 5.2: Average amplitudes of effect across all unimodal conditions for each subject. Respective standard error values are denoted in brackets.

In a comparison of the average amplitudes of effect, the mean amplitude of effect for all subjects across the three unimodal conditions is 0.423Hz, giving a 14% variance from baseline. In a further comparison of inter-individual difference, subject AM consistently produces effect sizes considerably larger than subjects DW and YL, as evidenced in the table above. Similarly, the average spread of the unimodal curves was 0.334 in log units across all subjects and subject YL produces a larger spread of effect in log units compared to subjects AM and DW (demonstrated below).

	Spread of effect – AM (in log units)	**Spread of effect – DW (in log units)**	**Spread of effect – YL (in log units)**
AA	0.31 *(.06)*	0.18 *(.02)*	0.45 *(.09)*
TT	0.40 *(.15)*	0.25 *(.03)*	0.43 *(.12)*
VV	0.36 *(.08)*	0.21 *(.01)*	0.43 *(.16)*

Table 5.3: Average spread of effect across all unimodal conditions for each subject. Respective standard error values are denoted in brackets.

From these results it is clear that adapting to a slower rate than 3Hz results in a 3Hz temporal presentation appearing faster than it actually is, whereas adapting to a faster rate than 3Hz subsequently makes the same 3Hz presentation feel significantly slower than 3Hz. Whilst this type of effect is seen for every unimodal condition; it remains clearly absent in all cross-modal conditions.

We deduce from these results that adaptation to temporal frequencies happens very flexibly and quickly in sensory systems and that these effects are not particularly difficult to create. However they are limited to each sensory modality independently, and do not cross-over or share effects derived as a result of adaptation between modalities.

Duration perception is modulated by several factors including periodic attention (Shima, Murai, Yuasa, Hashimoto, & Yotsumoto, 2018). It has even been found that attended durations are efficiently and reliably coded whereas their unattended counterparts are either weakly encoded or not encoded at all (Maarseveen, Hogendoorn, Verstraten, & Paffen, 2018). Also, responses to spatially-attended as opposed to unattended targets in all tasks is faster (Jones, 2015). In our current set-up, subjects were always aware of the test modality that the stimulus would be presented to, as this would be the test sequence that they would be required to reproduce. Consequently, there is a possibility that subjects perhaps failed to attend to the adapting stimulus if they were aware that the test stimulus would be presented to a different modality. Becker and Rasmussen (Becker & Rasmussen, 2007) were aware of this concern, however they asserted that it was unlikely to have influenced their findings. Levitan and colleagues (Levitan et al., 2015) introduced a practical gap-counting paradigm during the adaptation phase to control the confounding effects of attention on any adaptation after-effect. To address this concern with our present experiments, we repeated the experimental run in a control experiment where the modality of the test stimulus was unknown to the subject. The experiment was coded such that the subject was aware of the adapting modality however the testing modality on each trial was selected at random. Using the audio-visual pairing, we repeated all experimental trials using this modality pairing where 50% of test trials were randomly auditory, whereas the other 50% were randomly visual. Due to methodological constraints, the tactile conditions could not be tested. This is because any experimental set-up employing the tactor would alert the subject about the upcoming condition to be tested and thereby, negate any attempt to equalise attention during the experimental block and across different conditions. The paradigm is a simple one – any purposeful strategy during the adaptation phase would affect both auditory and visual test stimuli alike, resulting in any after-effect being either present or absent from both conditions. Conversely, should adaptation persist in the unimodal but not cross-modal pairing, then the potentially contaminating role of attention during the adaptation phase can be eliminated. Results from all subjects are presented below:

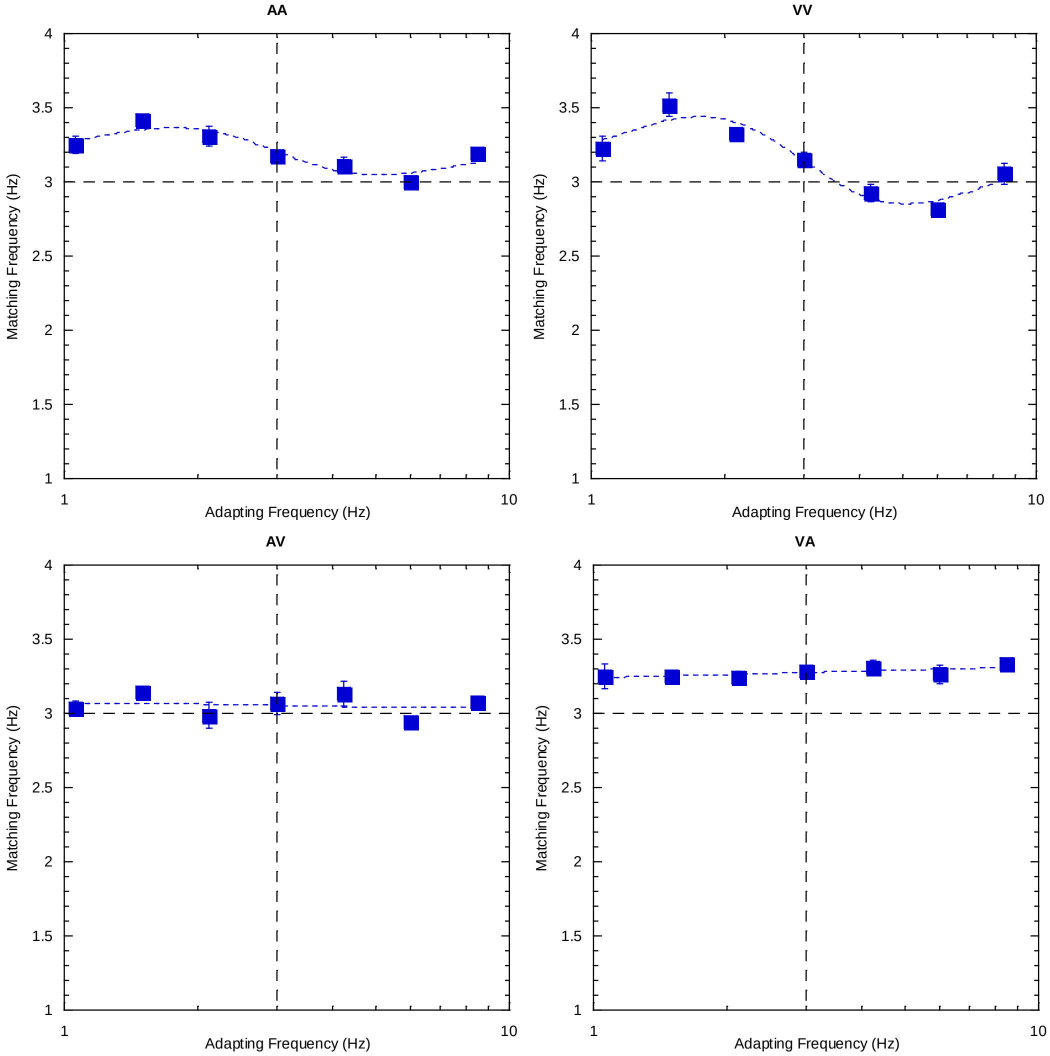


Figure 5.9. Data from the control experiment using the auditory/visual pairing (subject DW). Left-hand plots represent the auditory adaptation condition, right-hand plots visual adaptation. Upper plots represent unimodal conditions (adapt and test same modality), lower plots cross-modal conditions; error bars indicate standard error. See text for a description of the control methodology.

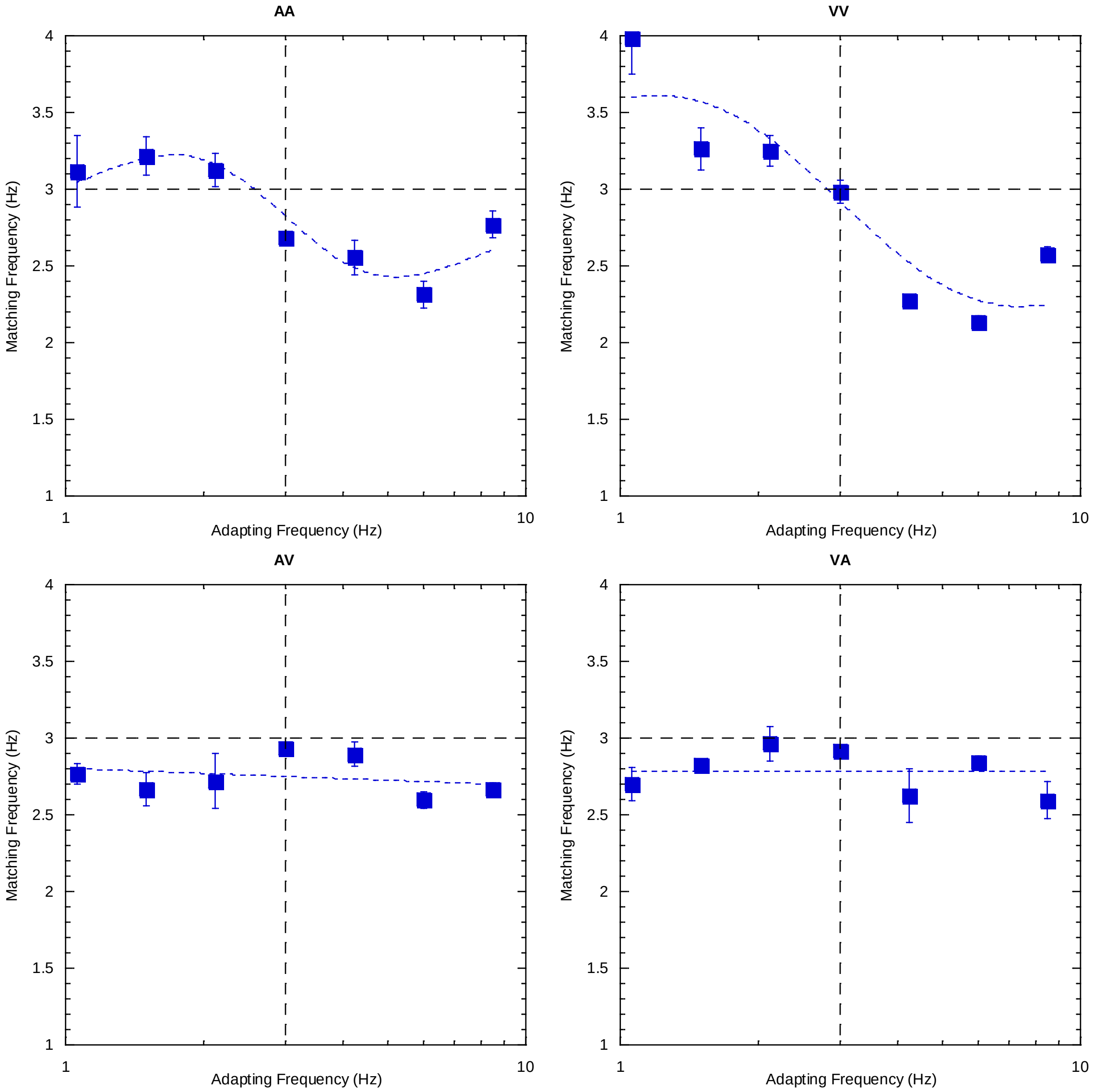


Figure 5.10. Data from the control experiment using the auditory/visual pairing (subject AM). Left-hand plots represent the auditory adaptation condition, right-hand plots visual adaptation. Upper plots represent unimodal conditions (adapt and test same modality), lower plots cross-modal conditions; error bars indicate standard error. See text for a description of the control methodology.

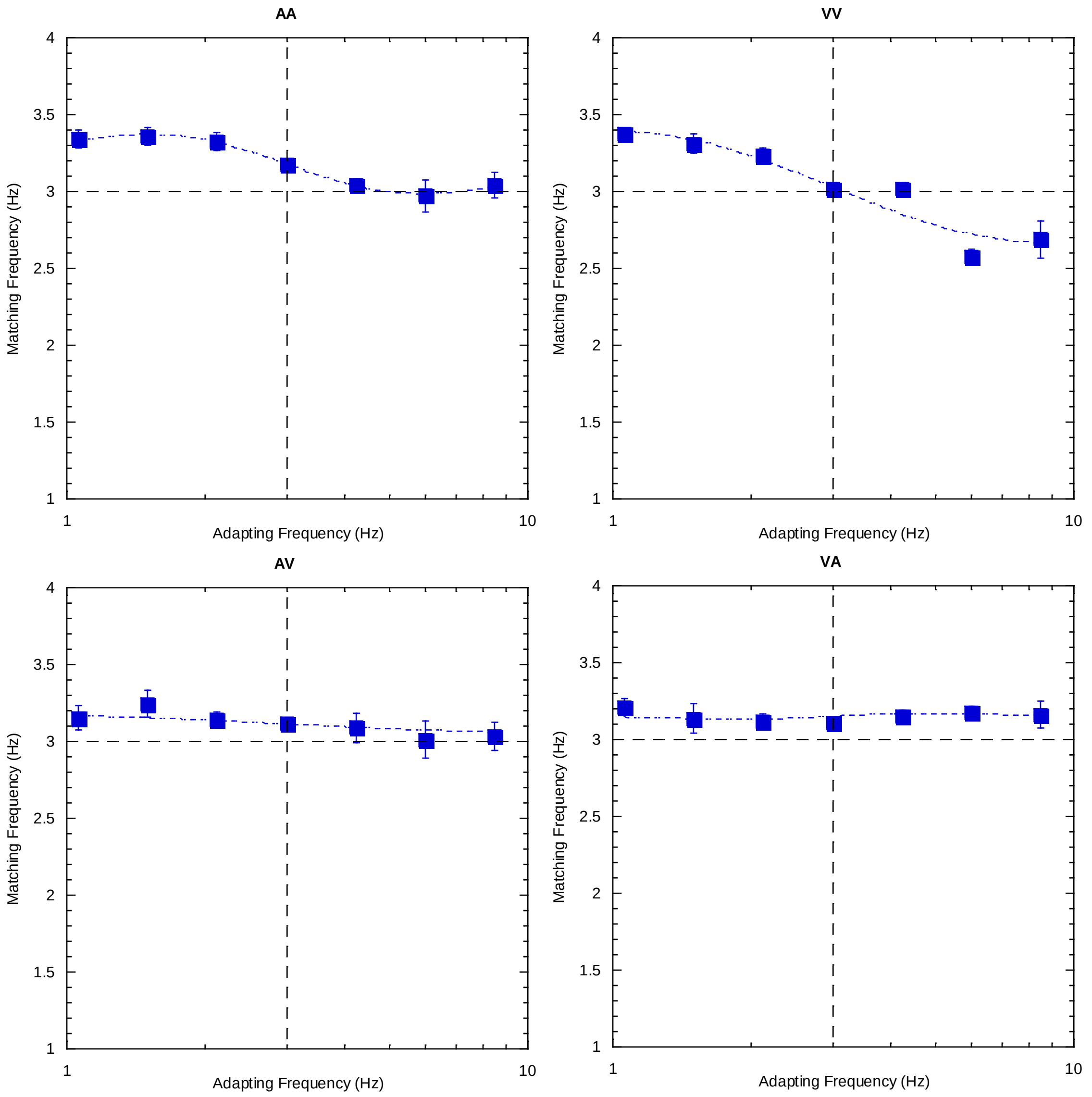


Figure 5.11. Data from the control experiment using the auditory/visual pairing (subject YL). Left-hand plots represent the auditory adaptation condition, right-hand plots visual adaptation. Upper plots represent unimodal conditions (adapt and test same modality), lower plots cross-modal conditions; error bars indicate standard error.

Condition	Subject AM	Subject DW	Subject YL
AA	.007*	.014*	<.001*
VV	.020*	.003*	.036*

AV	1	1	.963
VA	1	1	.963

Table 5.4: Holm-Bonferroni-adjusted p-values for all subjects across all conditions for the control experiment in which subjects were unaware of the test modality. The first letter denotes the adapting modality and the second letter denotes the testing modality, ‘A’ refers to the auditory modality and ‘V’ to the visual modality. Values marked with an asterisk (*) signify results of statistical significance.

	Amplitude of effect – AM (Hz)	**Amplitude of effect – DW (Hz)**	**Amplitude of effect – YL (Hz)**
AA	.40 *(.08)*	.16 *(.04)*	.19 *(.01)*
VV	.69 *(.17)*	.29 *(.05)*	.36 *(.10)*

Table 5.5: Average amplitudes of effect across both unimodal conditions for each subject. Respective standard error values are denoted in brackets.

	Spread of effect – AM (in log units)	**Spread of effect – DW (in log units)**	**Spread of effect – YL (in log units)**
AA	.24 (.04)	.23 (.05)	.29 (.01)
VV	.39 (.20)	.23 (.03)	.49 (.27)

Table 5.6: Average spread of effect across both unimodal conditions for each subject. Respective standard error values are denoted in brackets.

Conclusive findings are observed – after-effects prevail for both auditory and visual adapting conditions but only when the test stimulus is presented to the same modality, and not when they are presented to a different modality. Thus, in line with results from the main adaptation experiments, adapting to a specific temporal frequency does influence the perception of rate in typical rebound-type effects however, this result is exclusive to unimodal conditions where the adapting and testing modalities are the same as these effects fail to transfer cross-modally.

5.6. Discussion

Since the publication of Blakemore and Campbell's 1969 paper on selectivity in the human visual system (Blakemore & Campbell, 1969), much more evidence has been garnered towards the idea of the sensitivity of sensory systems towards specific features of perception. The present data delivers further support for selectivity in sensory systems, in this instance, within the domain of time. Sensory systems are able to flexibly and rapidly adapt to changes in temporal rate, and demonstrate strong band-limited repulsive after-effects in all three sensory systems of audition, touch and vision. Since the repulsive after-effects failed to transfer across modalities, we suggest that in line with Becker and Rasmussen (2007), sensory timing abilities operate with distributed timing mechanisms, each dedicated and largely independent to each sense. Adaptation is understood to be a consequence of sensory history within neural populations. Specifically, when adapting and test stimuli address overlapping neural populations, the resulting perceptual artefacts are observed as repulsive after-effects. Had such effects transferred cross-modally, it would have suggested that multiple senses operate using the same temporal principles and would have allowed a suggestion of a central, supramodal timing mechanism shared between the senses - as suggested by Levitan et al. (2015). However, our data refute this possibility. We suggest that recent sensory history does indeed influence the perception of rate, but crucially, that the resulting after-effects are modality-specific.

Certain evidence suggests spatially-specific processing of sensory time (Burr et al., 2007; Johnston, Arnold, & Nishida, 2006). It cannot be ruled out, therefore, that cross-modal after-effects only present when all sensory streams are spatially-overlapped. Perhaps the lack of cross-modal after-effects evidenced here is due to the fact that auditory stimuli were presented over headphones whereas visual stimuli were presented on a display. To address this concern, a second control experiment was conducted in which auditory and visual stimuli were spatially co-localised. Specifically, to achieve this aim, visual stimuli were projected on a thin fabric sheet enabling acoustic signal transparency. Within this set-up, auditory stimuli were simultaneously projected using a loudspeaker placed directly behind the acoustically-transparent screen. All other features were kept consistent with the experimental set-up of the main experiment and first control experiment. Data were gathered for all possible pairings encompassing the auditory and visual modalities (AA, VV, AV, VA) over a minimum of 105 trials for each subject. Data for all subjects is presented below:

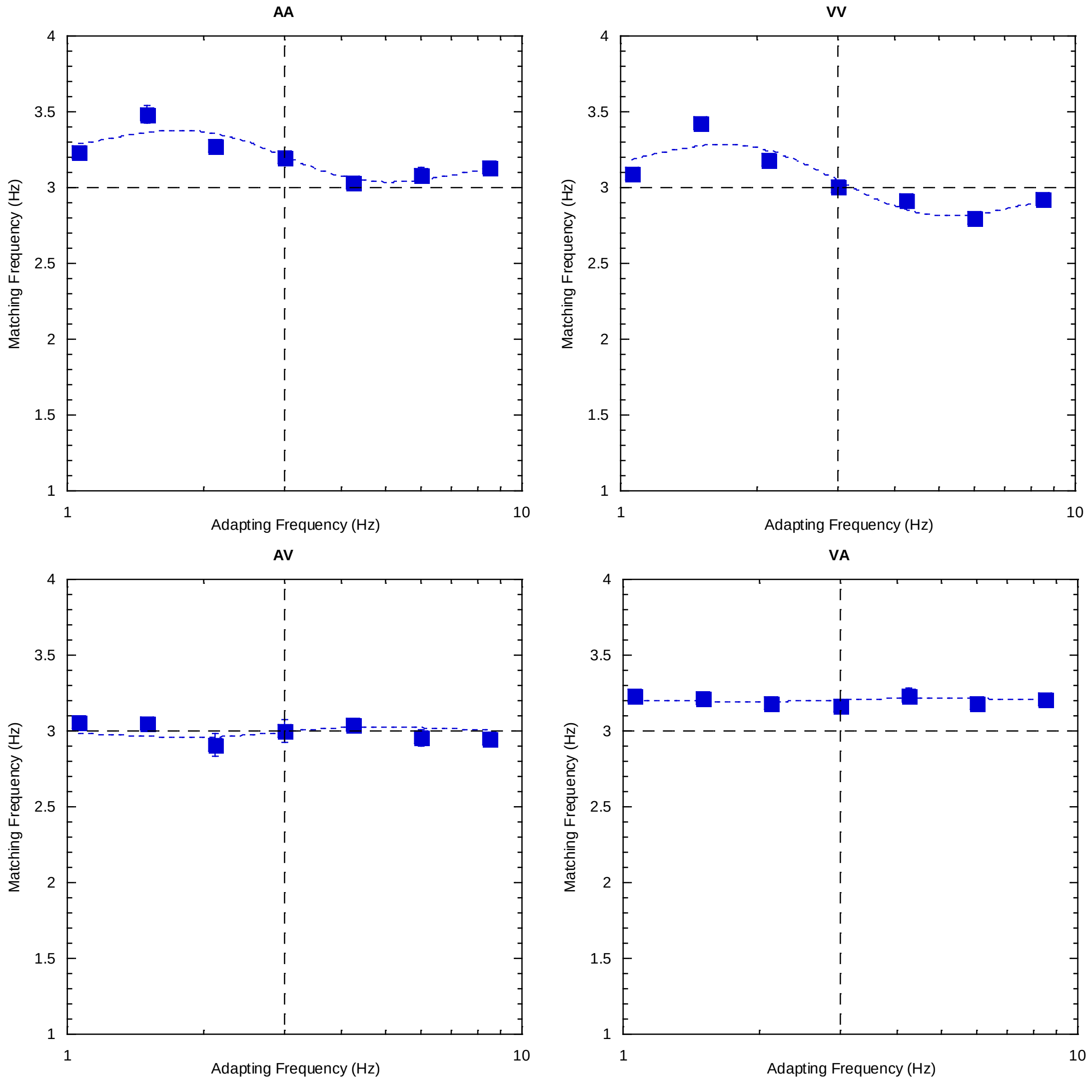


Figure 5.12. Data for all four adapt/test stimulus pairings for subject DW where stimuli were spatially and temporally overlapped. The sensory combination is shown at the top of each plot. The two unimodal conditions are shown in the top panel (left; AA, right; VV). Error bars indicate standard error. See text for further description.

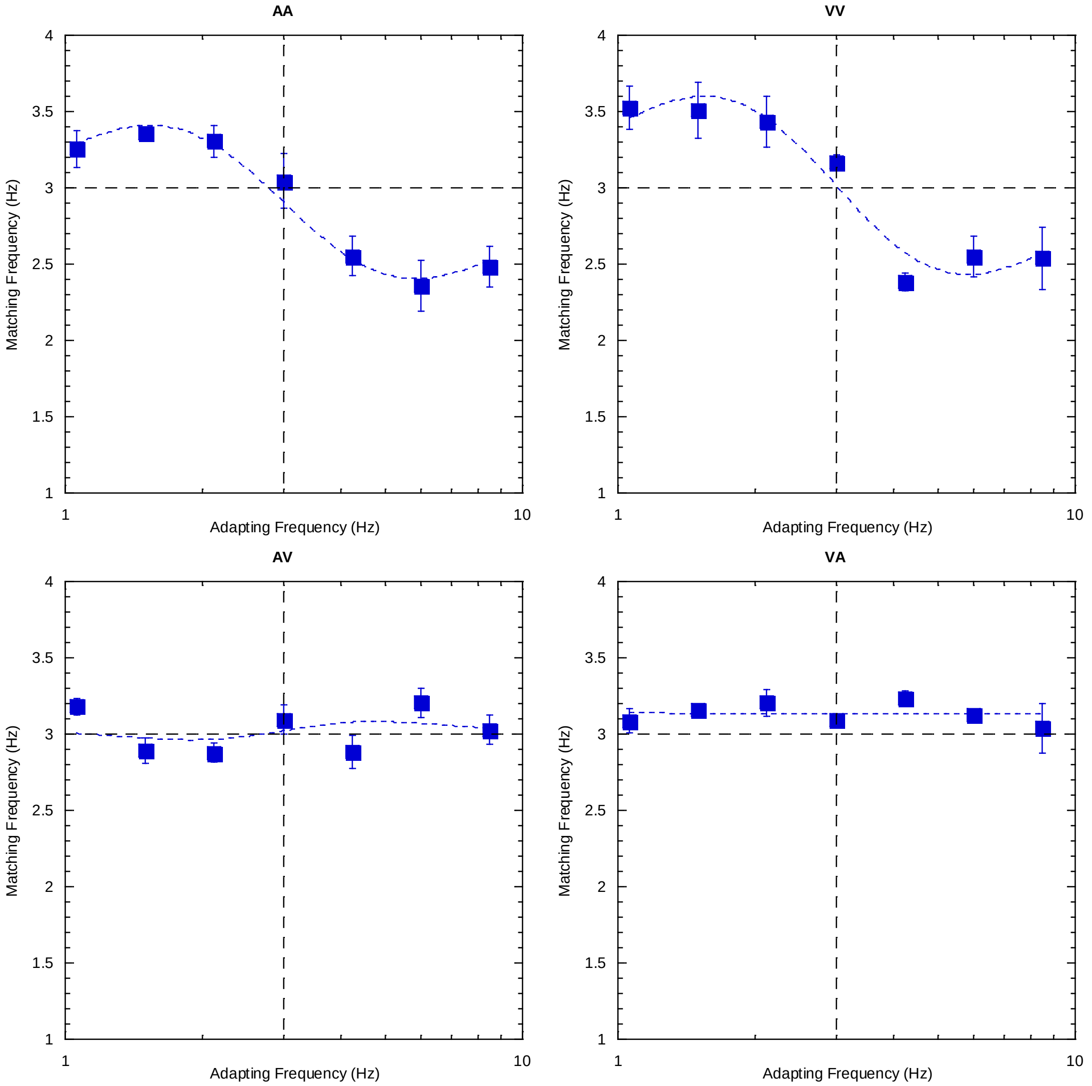


Figure 5.13. Data for all four adapt/test stimulus pairings for subject AM where stimuli were spatially and temporally overlapped. The sensory combination is shown at the top of each plot. The two unimodal conditions are shown in the top panel (left; AA, right; VV) whereas the cross-modal conditions are presented in the lower panel

(left; AV, right; VA). Error bars indicate standard error. See text for further description.

	Subject AM			**Subject DW**		
	Amplitude (μ)	**p-value**	**Spread (σ)**	**Amplitude (μ)**	**p-value**	**Spread (σ)**
AA	0.50±.04	<0.001	0.29±.03	0.17±.05	0.030	0.23±.05
VV	0.58±.08	<0.001	0.29±.04	0.24±.06	0.022	0.25±.06

Table 5.7: Amplitudes of adaptation effect (μ), spread (σ in log units) of adaptation effect and Holm-Bonferroni-adjusted p-values across all unimodal conditions for each subject for the control experiment in which visual and auditory stimuli were spatially and temporally overlapped. The first letter denotes the adapting modality and the second letter denotes the testing modality, 'A' refers to the auditory modality and 'V' to the visual modality. All cross-modal conditions were found to be not statistically significant (p>0.05).

Results are consistent with findings from the main and first control experiments – adapting to a given rate in the adaptation phase only affects the perception of a test rate when the adapting and test sensory modality match. Importantly, our findings also suggest that while some temporal after-effects present only when the adapting and test stimuli are co-localised spatially, this experimental feature is not sufficient in eliciting cross-modal transfer within rhythm adaptation.

There are several plausible explanations as to why the results presented here are in such contrast to those of Levitan and colleagues. Firstly, Levitan et al., used a "missing pulse" task, where several pulses were missing in an otherwise regularly-paced rhythm. This required subjects to self-report how many pulses were missing, and possibly inducing a reporting bias. Such a design would also encourage the subject to count the pulses (and in some instances, even tap along) in order to correctly report how many were missing, this consequently shifts the focus toward counting the signals (thereby becoming a measure of numerosity), irrespective of modality and also make any resulting after-effect more likely to be a supramodal one. Additionally, as there was no explicit reference standard, subject responses were a less direct measure of rate (“was this faster, or slower?”). This meant that the internal mean was exposed to a criterion bias and also that there was nothing to stop this internal mean from being modified on trial to trial.

Recent work using fMRI suggests evidence of ‘neural tuning’ in temporal representations within humans (Hayashi et al., 2015). Repeatedly presenting a participant with stimuli of the same duration resulted in a significantly decreased level of activity within the right inferior parietal lobule (IPL). Further experimentation on an extensive range of subsecond durations produced the same result, indicating preferential tuning of IPL neurons to specific subsecond durations (Hayashi et al., 2015). Furthermore, similar tuning effects have also been found within the auditory modality in the transfer of auditory temporal learning (Wright, Buonomano, Mahncke, & Merzenich, 1997; Karmarkar & Buonomano, 2003), and in the somatosensory modes for interval discrimination (Nagarajan, Blake, Wright, Byl, & Merzenich, 1998; Johnston et al., 2006). Collectively, these findings indicate that the processing of specific timeframes is underlain by the presence of dedicated circuitry; reinforcing the support for band-limited tuning (also known as temporal channels) within the cortex.

It is demonstrated that adapting to lower temporal frequencies produces higher frequency tapping responses but this eventually returns to veridical upon reaching the extreme ends of frequency range tested, (and vice-versa for adaptation to higher temporal frequencies). It is plausible that because these responses are for frequencies at the more polar ends of our adapting spectrum, perceptual effects begin to dissipate as perceptually, the difference in adapting and testing temporal frequencies are too dissimilar. In other words, classic band-limited ‘rebound’ adaptation effects are evidenced here, as elaborated upon earlier (see Becker and Rasmussen, 2007). Importantly, experimental observations of duration reproduction and sensory adaptation have similarly failed to find evidence of cross-modal effects (Nemes, Whitaker, Heron, & McKeefry, 2011). However, what remains consistent across the sample assessed by Heron et al., (2012) and Levitan et al. (2015), along with the current data is a gradual diminishing of effects as the difference between the adapting and test stimuli becomes greater - a ‘tailing’ off at the end of the observed curves. As the channels framework has been supported in duration perception already (Heron et al., 2012), we use these current findings to assert support for the channels framework in rate research too, however the true nature of these channels remains yet to be elucidated as Levitan and colleagues propose these are multisensory timing channels whereas our data support more modality-specific segregation of timing information. In any case, emerging evidence continues to support a channels-based model of timing perception.

It may be possible that “more adaptation” (i.e. cycles present within the same window) could have occurred for stimuli presented at higher temporal frequencies. For example, adapting to visual stimulus at 8.46Hz for a duration of 10 seconds may result in stronger adaptation effects compared to adapting at 1.06Hz of the same stimulus for an identical duration. However, if this were true, we would have found greater magnitudes for reproduction responses at higher frequencies, yet this was not the case, as no such effects were found. Importantly, the duration of the adaptation period was set to 8-10 seconds, but the exact duration for that particular trial was programmed at random therefore negating the possibility that any subject may have been able to count individual signals during either the adapting or test periods. Furthermore, the profile of each curve indicates equal levels of adaptation effects across conditions and if adapting to higher temporal frequencies produced stronger adaptation effects, then we believe these effects were still negligible as they were not evidenced in responses.

Modality-specific benefits are observed in beat perception for the auditory versus visual modality (Patel, Iversen, Chen, & Repp, 2005). Claims have been made that despite the impeccable human ability to perceive and act on a beat, evidence of robust visual beat perception is rare (Grahn, 2012) and that generally, beat perception is an ability that has stronger ties to the auditory modality (Glenberg & Jona, 1991; Repp & Penel, 2002). Grahn notes that this is particularly surprising for a number of reasons, such as the fact that one of the most fundamental features of time is that it is ‘amodal’ – being perceived and processed by a range of different modalities. Substantial overlap exists between neural substrates of auditory and visual timing (Schubotz, Friederici, & Von Cramon, 2000; Wiener, Turkeltaub, & Coslett, 2010) and so any advantages in temporal accuracy afforded by auditory timing should be of comparable benefit to the visual modality also (Grahn, 2012). Thus, there is an urgent need to clarify suggestions of rhythm perception, or any other temporal phenomena being more exclusive to one modality over another. The results here distinctly demonstrate that all three sensory modalities presented comparable rhythm after-effects despite the rhythms being presented unimodally. And perhaps importantly, that each modality experimented with here is distinctly able to process rhythm and additionally, is able to adapt to rhythms presented within this modality for a range of adapting temporal frequencies.

The use of structural equation modelling has been made to demonstrate a model inclusive of both modality-specific and modality-independent levels of temporal processing (Stauffer, Haldemann, Troche, & Rammsayer, 2012). Specifically, the model promotes modality-specific processing at the earlier stages of processing, and suggests modality-independent processing of temporal information occurs during later stages of the processing hierarchy (Stauffer et al., 2012). In other words, it may be possible that cognitive systems employ both modality-independent and modality-specific mechanisms to understand and interpret temporal signals. Indeed, in a manipulation of duration ranges and sensory modalities, a common, supramodal cognitive mechanism overlooking timing in audition and vision was suggested (Noulhiane, Pouthas, & Samson, 2009). However, it has also been suggested is that audition has an additional dynamic process “superimposed” to this overarching mechanism. They attribute this feature to a longer lasting auditory sensory memory, allowing for a learning mechanism to become part of the process. Notably, they mark the 3 second point as a key indicator of the indifference interval regardless of the range of durations presented, in both modalities. One assertion that can be made here is that these results provide support for centralised timing mechanisms from their suggestions of supramodal cognitive mechanisms underlying timing in both audition and vision. However, the ranges of durations used were considerably longer (1.0 - 5.5s in one group and 1-10s in the second group) as traditionally methodology in interval timing literature uses durations shorter than 1 second to control the confounding effects of cognitive processes (Hellstrom & Rammsayer, 2004). A viable avenue to explore would be to replicate this design using much shorter durations and rates, (as we have here) to assess the automatic processing of timing in vision and audition. It would also be interesting to present auditory and visual stimuli closely but not simultaneously to investigate the unintentional cross-over of attentional effects.

Using data derived from sensory adaptation and rate reproduction experiments we show that sensory systems are rapidly able to adapt to a range of temporal frequencies. It is clear that this modulation is relatively simple to elicit. However, whilst these effects were able to transfer **within** modalities, crucially, they did not transfer across modalities. In other words, when modalities of the adapting and test phases were not identical, no significant effects were found. This supports the existence of distributed timing mechanisms, each independent and specific to a particular sensory modality.

6. Sensory rate perception – Simply the sum of its parts?

6.1. The evolutionary basis of rhythm

Human appreciation of rhythm represents one of the earliest forms of communication assisting our ancestors, and accordingly, has a universal and ancient history. Intricately carved specimens of bone flutes have been predated to exist thousands of years before the earliest known cave paintings (Conard, Malina, & Münzel, 2009), and beats exist in almost every known culture (Nettl, 2000). Interactions between movement and sound in the form of sensori-motor synchronronisation have been evidenced at birth, indicating the importance of temporal processing from the very onset of life (Provasi, Anderson, & Barbu-Roth, 2014). This continues early on in life, for instance, many caregivers hold and bounce infants rhythmically whilst also singing to them simultaneously. These social interactions demonstrating interpersonal synchrony can not only shape their early music and rhythm perception but also encourage social cognition and prosocial behaviour (Cirelli, Trehub, & Trainor, 2018). Evidence suggests that special social preferences and "selective prosociality" is shown towards individuals singing socially-learned and familiar melodies indicating how intertwined rhythms are in our everyday lives (Cirelli et al., 2018). One can ask whether our intuitive ability to pick up on rhythm is what encourages us to incorporate rhythms in every aspect of life (from the very onset of birth), or rather whether this ability has been developed consciously as one of the most efficient tools of communication and survival. Moreover, this affinity for rhythms extends throughout a human being's lifespan. The ability to tap along to a beat almost unconsciously and automatically is, one of the most remarkable abilities that humans possess. This is specifically due to the level of competence needed to extract a beat from a larger and more complicated musical ensemble and the magnitude of abstraction required regarding the temporal structure of the stimulus (Tal et al., 2017). Additionally, once rhythms are consciously perceived as rhythms, these become harder to distinguish into the component intervals that constitute them (Garner & Gottwald, 1968).

The precise mechanisms by which the brain processes rhythms across the range of timescales and sensory inputs it receives is currently not fully understood. This forms the central question underlying the work conducted here. In particular, we aim to establish the perceptual link between rhythms and the component intervals that form those rhythms. From a Fourier perspective, any signal derived from a 333ms empty interval will present substantial commonality with the signal derived from a 3Hz rhythm. Additionally, information about one is all that is needed to construct the other. An example of this is knowing that a 1.5Hz rhythm can be constructed into its component 667ms intervals. Despite this direct link, the phenomenological differences between rhythms and their component intervals could not be starker. Subjectively, rhythms feel much more intuitive and ‘natural’ when compared to their component intervals, and accordingly, have a much more automatic response (for instance, tapping along to a beat).

6.2. Differences and similarities between duration and rate perception

In order to comprehensively understand time, it is important to distinguish between temporal duration and temporal rate. Duration explicitly refers to the temporal extent of a sensory event whereas rate refers to explicit frequency of sensory presentation. It has been suggested that a single model may be able to account for both, duration and rate perception as these two concepts are heavily intertwined and a mechanism capable of processing one, would, in theory, also be able to fully make sense of the other (Hartcher-O’Brien, Brighouse, & Levitan, 2016).

Discrimination of intervals, both as part of a sequence but also individually allows us to distinguish between beat-based and interval-based timing systems. An interval-based timer specifically refers to a clock-like mechanism that times and stores the representation of a temporal event (or duration) (Pashler, 2001a). Beat-based timers, on the other hand, refer to temporal mechanisms that contain information regarding the presentation of consecutive intervals, each presented in quick succession of the last one (akin to a rhythm). Thus, because of their very nature, interval timers are universal in their abilities as even sequences of intervals (or beats) can be broken down into their building blocks of single intervals. Beat-based timers on the other hand, do not possess the same utility due to their specialised nature. Support for a beat-based model is presented by McAuley and Kidd (McAuley & Kidd, 1998) who presented subjects with a pattern of four tones, each segregated by a 400ms interval and then four comparison tones, each separated by an interval either longer, or shorter than 400ms. The design culminated in a two-alternative forced choice where the subject was requested to report whether the comparison interval was shorter or longer than the standard. Performance was found to be poorer when the comparison tone was presented earlier than expected compared to the beat created by the first set of standard tones, suggesting activity of beat-based timing mechanisms.

In efforts to deduce the differences in beat-based (relative) and duration-based (absolute) auditory timing, Teki and colleagues, (2011) assessed a functional dissociation of the cortical networks mediating these two processes. Using sequences of regular and irregular intervals, and by controlling for the interval between the reference and test stimuli, it was hypothesized that in irregular sequences, each separate interval would need to be timed and calculated individually. Therefore, to process such a sequence the brain would employ duration-based mechanisms. Contrastingly, regular sequences would recruit beat-based timing mechanisms to calculate the regularly repeated intervals. The contribution of neural circuitry was found to be clearly dissociated for the two sequences. Specifically, Teki et al. found that duration-based timing was mediated by the olivocerebellar network employing the cerebellum and inferior olive whereas beat-based timing employed a striatio-thalamo-cortical network involving several areas including the putamen, caudate, thalamus, pre-SMA/SMA, premotor and dorsolateral prefrontal cortex (Teki et al., 2011) (see Chapter 2). Later evidence from Teki et al. (Teki, Grube, & Griffiths, 2012) suggests a high-level of co-dependence between these networks, further implicating the interconnected nature of these networks through the cerebral cortex and numerous synaptic pathways (Teki et al., 2012).

Duration and rate perception arising from closely-related mechanisms have been suggested by work conducted by Johnston et al. (2006). Specifically, prolonged viewing of visual stimuli with a fixed temporal frequency induce compression in the perceived duration of a subsequently presented test stimulus. It has been suggested that this duration compression occurs even when the perceived temporal frequency of the test stimulus is veridical (Johnston, Arnold, & Nishida, 2006). A number of other studies have also shown an interaction between duration and temporal frequency and specifically, that the perceived duration of a stimulus is mediated by a closely presented temporal frequency in the visual and auditory modalities (Treisman, Faulkner, Naish, & Brogan, 1990; Kanai, Paffen, Hogendoorn, & Verstraten, 2006; Ono & Kitazawa, 2011). Specifically, they suggest that the clock for perceived duration is overseen by the temporal frequency of a stimulus (Kanai et al., 2006).

Nevertheless, the relationship between temporal rate and duration is still under scrutiny (Maarseveen, Paffen, Verstraten, & Hogendoorn, 2019) . Conflicting evidence has been suggested by Bruno et al. (2015), who present evidence of duration compression despite controlling for temporal frequency and that this duration compression occurs independent of the temporal frequency of the stimulus, suggesting that these are clearly dissociated processes (Bruno, Ayhan, & Johnston, 2015). To further this debate, it is known that rhythms sets up psychological expectancy (Grahn, 2012b) but despite this, many temporal events in nature have no supporting rhythm, any system dedicated to the processing of rhythm would have to occur alongside an interval-based timing system (Grahn, 2012a). The idea that sequences of intervals (beats) may be processed similarly to single intervals (duration) in a repeated-loop fashion has been briefly suggested previously (Keele, Nicoletti, Ivry, & Pokorny, 1989). More recent work has attempted to directly examine the bi-directionality between rate and duration (Hartcher-O’Brien et al., 2016) asserting that information from one is all that is needed to deduce characteristics for the other. For instance, knowing that an interval lasts for 500ms every time it is presented is enough for us to assert that the same interval presented in a sequence of signals (rate) would correspond to 2Hz (as similarly explained previously).

Despite human fondness for rhythm, the building blocks of time are single intervals and it is due to this that models of time perception have typically focused on duration estimates (Hartcher-O'Brien et al., 2016). It has been found that after adapting to fixed, filled intervals of either auditory or visual stimuli, distortions in the subjective duration of a reference interval become evident (a duration after-effect) (Heron et al., 2012). Similar to rate after-effects, the authors further extend their conclusions by suggesting evidence for duration channels that are also tightly tuned in their bandwidths and modality specific (see discussion - Chapter 4). The duration after-effect already established in audition and vision has recently also been demonstrated in the tactile modality (Li & Chen, 2018 – in personal communication at IMRF 2018 (Low-level duration after-effect occurs in tactile modality, IMRF Toronto 2018)).

Investigating empty intervals is an important opportunity for the current debate as it provides one method to disentangle whether single intervals are temporally unique in their own sense and how trains of intervals i.e. rhythms are processed. For example, are 'filler' characteristics (features that would fill an otherwise empty temporal interval) necessary to evoke distortions in time, or would an empty interval suffice? One method to enable this is sensory adaptation to empty intervals.

Present work thus aims to deconstruct the concept of rhythm (also known as rate) and clarify how exactly the brain processes a sequence of sensory signals. Explicitly, we examine the existence of a rhythm after-effect and, using single interval comparisons, question whether this effect is simply an extension of the previously documented duration after-effect (Heron et al., 2012) via a culmination of repeatedly presented single intervals. There are two potential outcomes to this experiment. The first is that subjects will be able to adapt to various rates and demonstrate after-effects when presented with single, empty durations (suggesting dependent processing mechanisms for rate and duration). The alternative is that adapting to a given rate will have no influence on subsequently presented empty intervals (suggesting independent processing mechanisms for rate and duration).

6.3. Methods

Subjects: Four participants (1 female and 3 male, mean age = 30, standard deviation = 14 years) participated, with self-reported normal hearing and visual abilities. Following initial practice sessions, a lengthy process (20-25 hours) of data collection began, in a series of sessions spread over several days. Two participants had previous experience of psychophysical data collection and two participants (ND and SA) were completely naïve to psychophysical research and the purpose of these experiments. The experiments received ethical approval from the Research Ethics Committee at the School of Optometry and Vision Sciences, University of Cardiff (U.K.) and all experiments were performed in accordance with relevant guidelines and regulations. Informed consent was obtained for study participation.

General Stimuli: Brief (16msec duration) sensory stimuli were presented – either in the auditory or visual modalities and all stimuli were grossly suprathreshold. Stimulus generation and presentation was controlled by an Intel ® Core ™ i5-4460 desktop computer running Microsoft Windows 7. The programming environment involved MATLAB 8.6 (Mathworks, USA) in combination with Psychophysics Toolbox 3 (http://www.psychtoolbox.org). Stimulus timing was verified using a dual-channel oscilloscope. Stimuli features such as volume, brightness and size were kept consistent with those used for experiments conducted in Chapter 5.

Visual:

Visual stimuli were presented on an Eizo EV2436W monitor. These were bright (274 cd/m^2) white circular flashes presented centrally against a uniform dark background (0.32cd/m^2). Stimulus duration was a single frame (approximately 16ms at the monitor frame rate of 60Hz). The viewing distance was kept constant at 60cm and visual stimuli had a diameter of 10.5cm.

Auditory:

Auditory stimuli consisted of brief (16ms duration) bursts of white noise generated by a Xonar Essence STX (ASUS) soundcard (https://www.asus.com/us/Sound-Cards/Xonar_Essence_STX/) with a sampling rate of 44,100Hz. Stimuli were delivered using Sennheiser HD280 Pro Headphones at an SPL of 70dB. Auditory stimuli were specifically chosen to be lacking in any possible pitch, timbre or dynamic variations to avoid confounding influences on rhythm (Rammsayer & Brandler, 2004).

Procedure

Pilot Experiments:

The first pilot experiment examined the ability for subjects to adapt to empty intervals of time – i.e. intervals with no temporal (or otherwise) filler characteristics. We are aware of such experiments using filled intervals (Heron et al., 2012) however, to our knowledge, we are not aware of a replication using empty temporal intervals. To summarise, subjects were adapted to 10 pairs of empty intervals demarcated by two sensory signals of either the auditory or visual modalities, separated by a gap of 160ms. The remainder of the task depended on whether the participant was completing the interval reproduction or two-alternative forced choice procedure.

Interval Reproduction (Experiment 1.1):

This experiment involved adapting the subject to 10 pairs of empty intervals demarcated by two sensory signals of either the auditory or visual modalities, separated by a gap of 160ms. In the test phase they were then presented with a reference interval of 333ms presented to the same modality that they adapted to, and were asked to reproduce this interval by tapping on a response device (a piezoelectric transducer) used to record interval reproduction. This response specifically involved tapping twice to indicate the start and stop of the interval (Mioni, Stablum, McClintock, & Grondin, 2014b). The resulting voltage output was fed to the ‘audio in’ of the soundcard as a recording which was analysed within MATLAB to extract the duration of the reproduced interval. The transducer was enclosed in a sound-dampening environment and shielded from sight of the subject. To further eliminate the possibility of auditory feedback, white noise was played via the headphones throughout the tapping response phase (Wearden, 2003).

Two-Alternative Forced-Choice (2AFC) (Experiment 1.2):

Subjects were adapted to 10 pairs of empty intervals demarcated by two sensory signals of either the auditory or visual modalities, separated by a gap of 160ms. The subject was then presented with a 2AFC testing period where on each trial, a reference interval of 333ms was presented in the modality not adapted to, followed by a test interval that could be one of seven intervals separated by linear steps ranging from 282-383ms and centred on 333ms (presented in the modality adapted to). The response period required subjects to respond to whether the test stimulus was longer than the reference stimulus by pressing a key on a keyboard. The time taken to respond was not recorded.

A pause of 200ms occurred between trials and performance feedback was not provided on any trial. Baseline data with no adaptation period were gathered for both subjects for all conditions. A minimum of 100 trials were conducted for the interval reproduction experiment and a minimum of 45 trials were conducted for the 2AFC method for subjects across both modalities.

Main Experiments:

The specificity of rate adaptation and resulting after-effects were investigated by adapting subjects to temporal rates (either 1.5Hz or 6Hz, fixed within a block) and testing with single, empty intervals using both interval reproduction and two-alternative forced choice methods. The interval reproduction method requires subjects to recreate their internally perceived durations after adaptation and therefore provides a very explicit response. Despite this, the interval reproduction method has been criticised for exposing a larger criterion-dependent bias (García-Pérez, 2014). The interval reproduction method was therefore used alongside a more conservative two-alternative forced-choice (2AFC) duration discrimination task. In this task, an unfilled reference interval was presented to the non-adapted modality (e.g. vision), followed by a variable (282-383ms in seven linear steps) unfilled test interval presented to the other (adapted) modality (e.g. audition). The task required the subject to report (via keypress) whether the test stimulus was shorter or longer than the reference stimulus. Following initial practice sessions, a process of data collection (approximately lasting 6 hours) began in a series of sessions spread over several days. Interval reproduction response periods can have a tendency to produce noisier results (Shi et al., 2013), and therefore more data was gathered for each condition involving an interval reproduction response period. Furthermore, in order to prohibit response delays that may be incurred from shifting the attentional focus between different modalities (Spence, Nicholls, & Driver, 2001; Spence, Shore, & Klein, 2001), separate blocks were used to presenting trials in different modalities.

Interval Reproduction (Experiment 2.1):

The interval reproduction experiment began with a 10 second adaptation phase where a train of stimuli with a fixed rate was presented to subjects. The sensory modality (either auditory clicks or visual flashes) and presentation rate (either 1.5Hz or 6Hz rate) of the adapting stimuli was held constant within an experimental session. The adaptation phase was followed by a test period composed of an empty reference interval of 333ms presented within the adapted modality. The test interval was identical on each trial. As before, subjects then reproduced this empty interval by tapping twice on the response disk (a piezoelectric transducer) used to record interval reproduction (Mioni, Stablum, McClintock, & Grondin, 2014a). The resulting voltage output was fed to the 'audio in' of the soundcard as a recording which was analysed within MATLAB to extract the duration of the reproduced interval. The transducer was enclosed in a sound-dampening environment and shielded from sight of the subject. To further eliminate the possibility of auditory feedback, white noise was played via the headphones throughout the tapping response phase (Wearden, 2003). A minimum of 150 trials were conducted for the interval reproduction experiment.

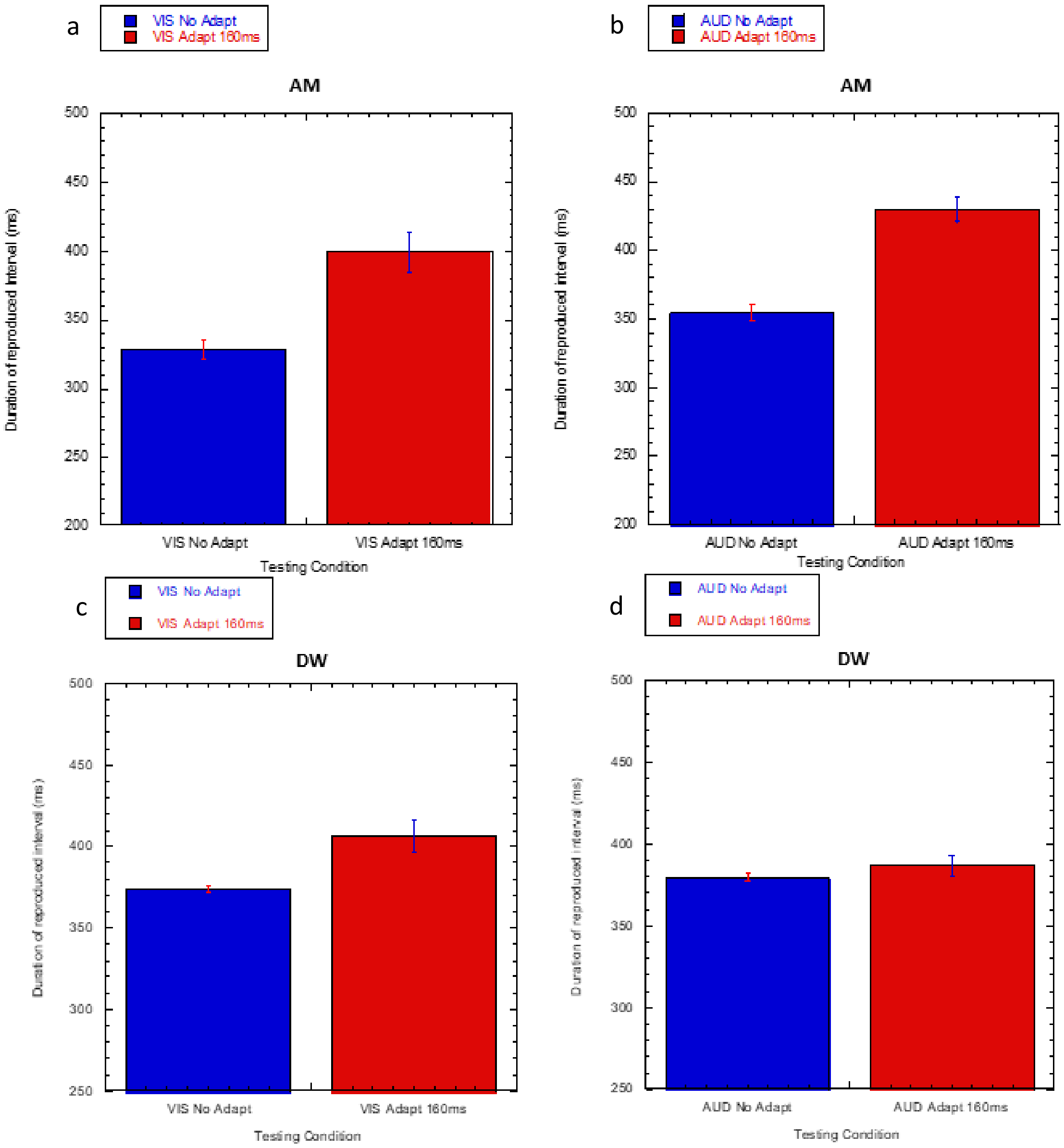
Two-Alternative Forced-Choice (2AFC) (Experiment 2.2):

The 2AFC experiments were identical to their interval reproduction counterpart for the adapting sequence. After this phase, the participant was presented with an empty reference interval of 333ms to the modality not adapted to, followed by a test interval presented in the same modality adapted to (ranging 282-383ms in seven linear steps, centred around 333ms). The response period required subjects to respond to whether the test stimulus was shorter or longer than the reference stimulus via a keypress. A minimum of 45 trials were conducted for the 2AFC method for subjects across both modalities. Performance feedback was not provided during either task. Baseline data collected without prior adaptation, were gathered for all conditions. The experiments were blocked according to sensory modality and adapting frequency. The order of blocks was randomised. This was kept consistent throughout all experiments and subjects in this chapter.

6.4. Results

Experiment 1.1:

Mean values for test intervals reproduced after adapting to empty intervals and rate were then averaged to provide a value for each condition. Data for both subjects are plotted below.



Figures 6.1a-d**.** Left plots indicates results from interval reproduction methods after having adapted to empty visual intervals, whereas right plots show results from interval reproduction methods after adapting to empty auditory intervals; values plotted are indicative of mean intervals reproduced for one subject and error bars indicate standard error. Upper panel of plots represent results for participant AM, lower plots for participant DW.

Results from Figures 6.1a-d demonstrate the key finding that empty intervals are sufficient to elicit distortions in interval perception and that ‘filler’ characteristics are not always necessary. The similar after-effects evidenced through both methods of interval reproduction and two-alternative forced-choice further reinforce the strength of this effect and suggest that the timing system is incredibly dynamic in adapting to temporal information. Very simple input, such as two single sensory signals can produce rapid distortions of time.

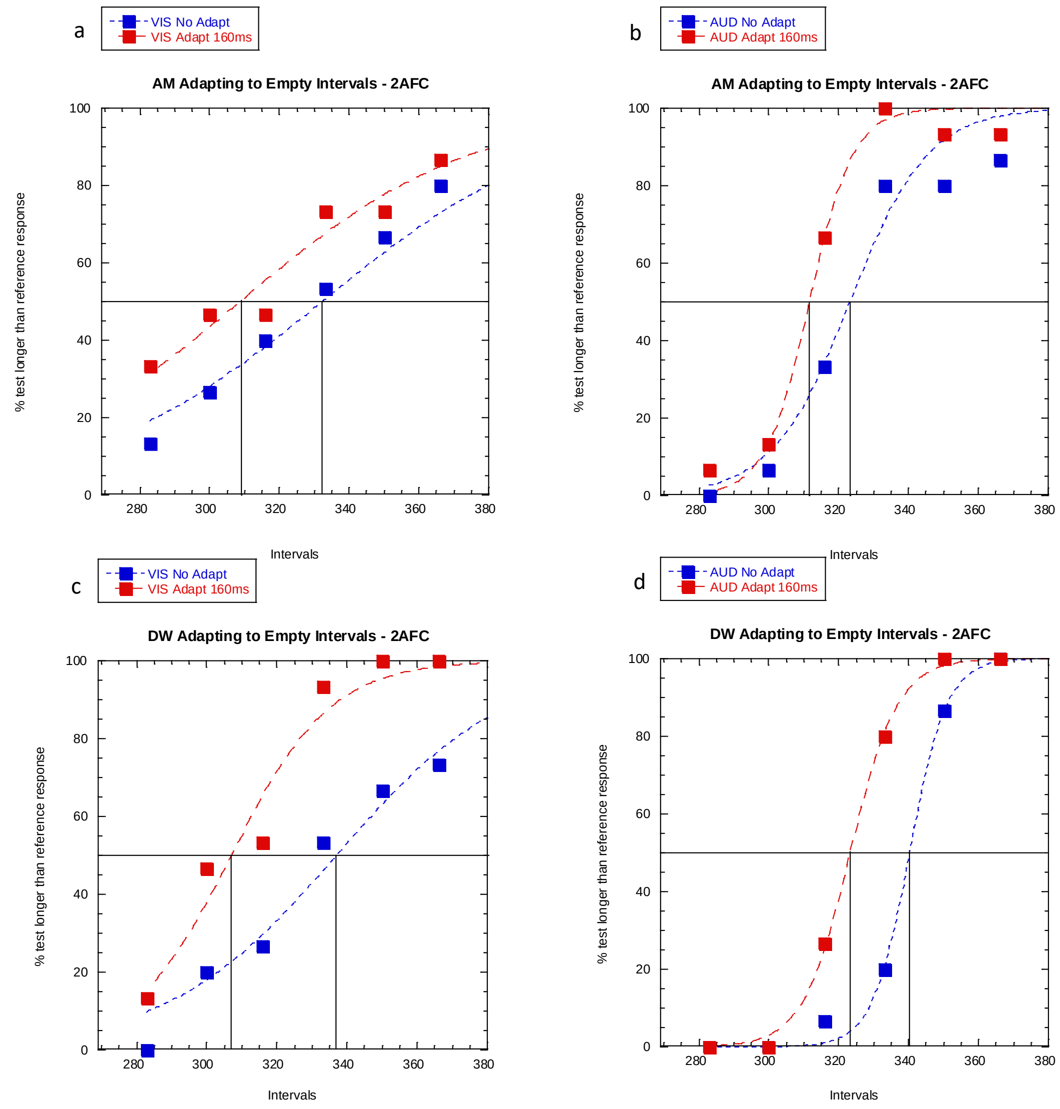
Mean values for test intervals reproduced after adapting to empty intervals were averaged to provide a value for each condition for each subject. Using SPSS (IBM Corp. Released 2015. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.), paired-samples t-tests were conducted on interval reproduction data from experiments on auditory and visual conditions to test the hypothesis that adapting to empty intervals of 160ms (in the auditory and visual modalities) would result in negative, rebound after-effects when responding using interval reproduction. Results were found to be statistically significant for both, visual and auditory conditions ($p<.05$) for subject AM and for the visual condition for subject DW ($p<.05$). Results for subject DW for the auditory condition were not statistically significant ($p>.05$).

Experiment 1.2:

As in Heron et al.’s paper (Heron et al., 2012) data from the 2AFC experiments were plotted with a psychometric function using the subject’s interval discrimination judgement of the proportion of responses of ‘test longer than reference’. The functions were then fitted with a logistic of the form

$$\boldsymbol{y} = \frac{100}{1 + e^{\left(x - \left(\frac{\alpha}{\theta}\right)\right)}}$$

where ‘$\boldsymbol{\alpha}$’ denotes the point of subjective equality (PSE – the 50% response level on a psychometric function) and ‘$\boldsymbol{\theta}'$denotes an estimate of the duration discrimination threshold.



Figures 6.2a-d**.** Right plots indicate results from two-alternative forced-choice methods after adapting to empty auditory intervals. Left plots indicate results from two-alternative forced-choice methods after adapting to empty visual intervals. Upper panel of plots demonstrate results from participant AM, whereas lower panel indicate results from participant DW.

Results from 2AFC experiments clearly reaffirm the pattern of results presented with interval reproduction methods for both subjects. Subject DW showed clear rebound after-effects after adapting to empty intervals of 160ms for both, auditory and visual conditions using 2AFC methods demonstrating the criterion-dependent bias that can sometimes influence results derived from interval reproduction methods.

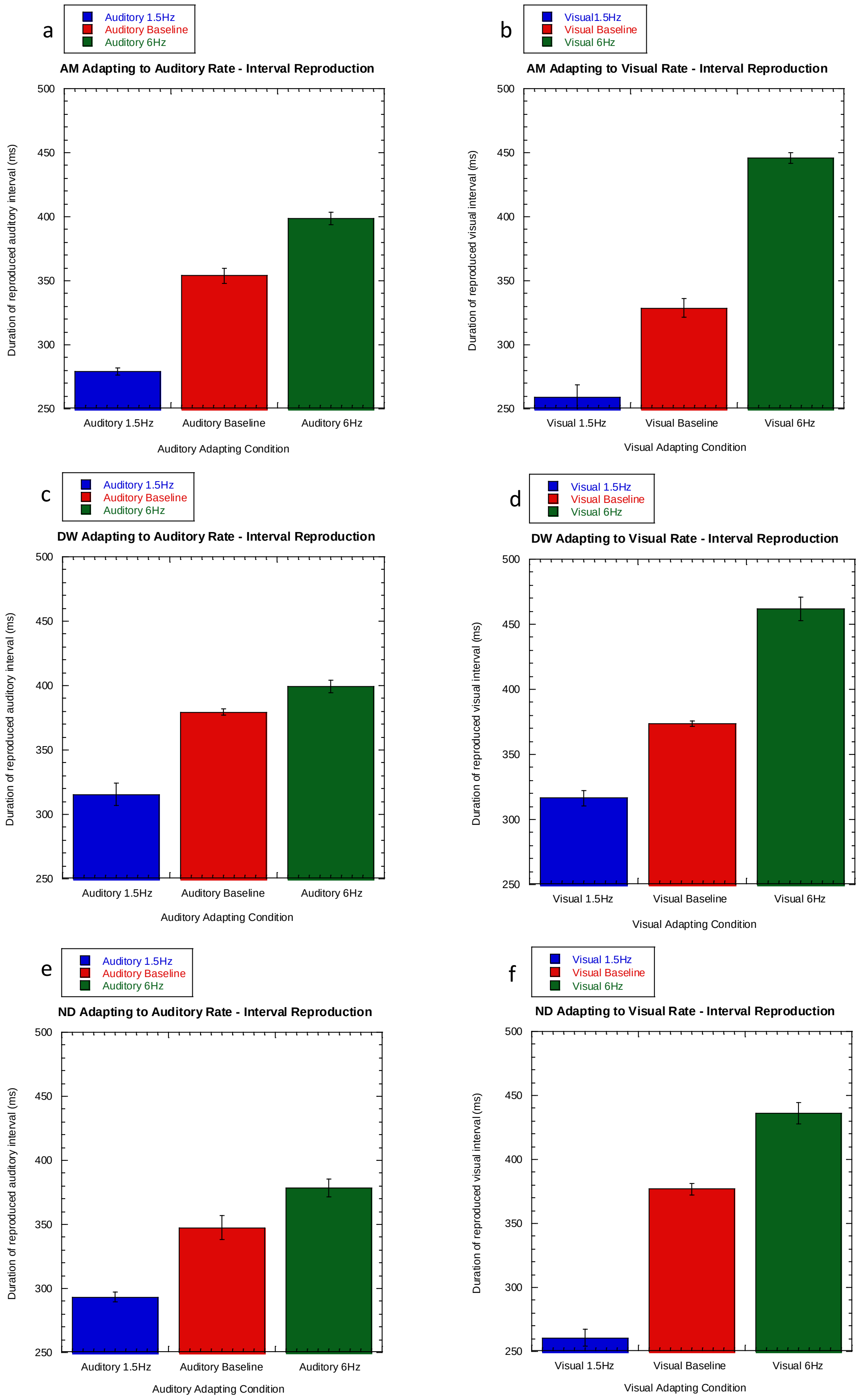
Confidence intervals at 95% were calculated for data to test the hypothesis adapting to empty intervals of 160ms for auditory and visual modalities would result in a shift in thresholds for the percentage of 'test longer than reference' responses. Results for both subjects are presented below in Table 6.1:

Experimental Condition	$\bar{x}_1 - \bar{x}_2$	95% CI for $\mu_1 - \mu_2$
(AM) Baseline-Adapt 160ms (A)	11.6	(5.2, 18)
(AM) Baseline-Adapt 160ms (V)	23.54	(10.7, 36.4)
(DW) Baseline-Adapt 160ms (A)	16.96	(15.3, 18.6)
(DW) Baseline-Adapt 160ms (V)	30.02	(20.9, 39.2)

Table 6.1: Results from Experiment 1 for subjects AM and DW with 95% confidence intervals; $\bar{x}_1 - \bar{x}_2$ represents the mean difference between data plotted using logistic functions and μ_1 -μ_2 represents the difference between mean squared error; (A) denotes the auditory modality and (V) denotes the visual modality.

Experiment 2.1:

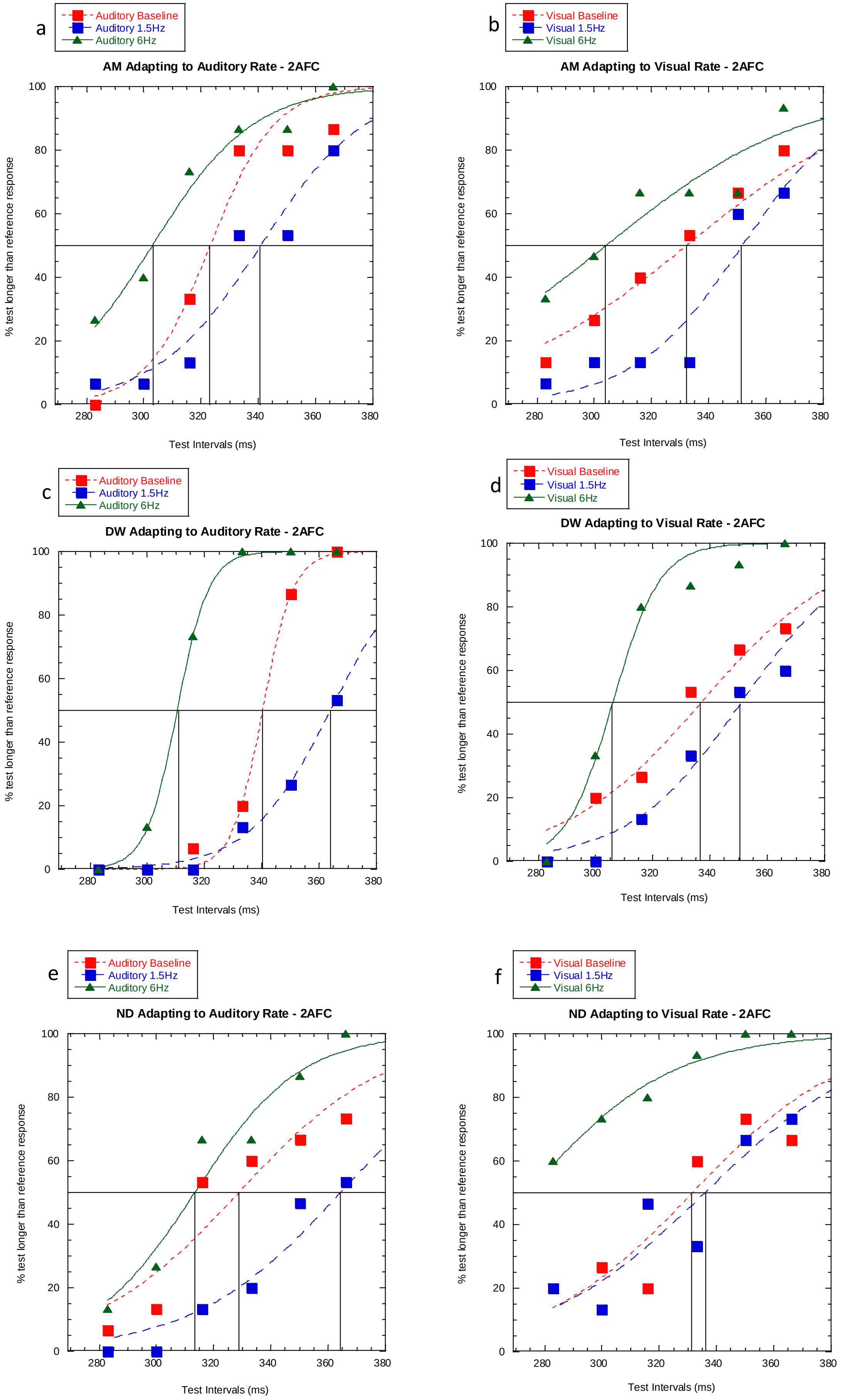
An auditory rebound after-effect pattern of results is evidenced through the first method of interval reproduction (Figure 6.3a, participant AM). Without adaptation, this subject reproduces the reference stimulus of 333ms at around 354ms. After adapting to a slow rate of 1.5Hz, the same reference is reproduced closer to 280ms and conversely, after adapting to a much faster rate of 6Hz, the same interval is reproduced closer to 398ms. A similar pattern of results is seen in the visual condition (Figure 6.3b) and across the other two subjects.


a
Auditory 1.5Hz
Auditory Baseline
Auditory 6Hz
AM Adapting to Auditory Rate - Interval Reproduction
Duration of reproduced auditory interval (ms)
500
450
400
350
300
250
Auditory 1.5Hz
Auditory Baseline
Auditory 6Hz
Auditory Adapting Condition
b
Visual1.5Hz
Visual Baseline
Visual 6Hz
AM Adapting to Visual Rate - Interval Reproduction
Duration of reproduced visual interval (ms)
500
450
400
350
300
250
Visual 1.5Hz
Visual Baseline
Visual 6Hz
Visual Adapting Condition
c
Auditory 1.5Hz
Auditory Baseline
Auditory 6Hz
DW Adapting to Auditory Rate - Interval Reproduction
Duration of reproduced auditory interval (ms)
500
450
400
350
300
250
Auditory 1.5Hz
Auditory Baseline
Auditory 6Hz
Auditory Adapting Condition
d
Visual 1.5Hz
Visual Baseline
Visual 6Hz
DW Adapting to Visual Rate - Interval Reproduction
Duration of reproduced visual interval (ms)
500
450
400
350
300
250
Visual 1.5Hz
Visual Baseline
Visual 6Hz
Visual Adapting Condition
e
Auditory 1.5Hz
Auditory Baseline
Auditory 6Hz
ND Adapting to Auditory Rate - Interval Reproduction
Duration of reproduced auditory interval (ms)
500
450
400
350
300
250
Auditory 1.5Hz
Auditory Baseline
Auditory 6Hz
Auditory Adapting Condition
f
Visual 1.5Hz
Visual Baseline
Visual 6Hz
ND Adapting to Visual Rate - Interval Reproduction
Duration of reproduced visual interval (ms)
500
450
400
350
300
250
Visual 1.5Hz
Visual Baseline
Visual 6Hz
Visual Adapting Condition

Figures 6.3a-f. The after-effect of adapting to different temporal rates of 1.5Hz (blue bar) and 6Hz (green bar) demonstrated through interval reproduction for auditory (left column) and visual (right column) conditions. Three different subjects are plotted at the top, middle and bottom of the panel. Values plotted are mean reproduction values of the test interval and error bars indicate standard error.

Mean values for test intervals reproduced after adapting to different rates were averaged to provide a value for each condition for each subject. Paired-samples t-tests were then conducted on interval reproduction data from auditory and visual conditions to test the hypothesis that adapting to different rates of 1.5Hz and 6Hz would result in rebound after-effects when responding using interval reproduction. Results were found to be statistically significant for both visual and auditory conditions ($p<.05$) for each subject.

Experiment 2.2:


a
Auditory Baseline
Auditory 1.5Hz
Auditory 6Hz
AM Adapting to Auditory Rate - 2AFC
% test longer than reference response
Test Intervals (ms)
b
Visual Baseline
Visual 1.5Hz
Visual 6Hz
AM Adapting to Visual Rate - 2AFC
% test longer than reference response
Test Intervals (ms)
c
Auditory Baseline
Auditory 1.5Hz
Auditory 6Hz
DW Adapting to Auditory Rate - 2AFC
% test longer than reference response
Test Intervals (ms)
d
Visual Baseline
Visual 1.5Hz
Visual 6Hz
DW Adapting to Visual Rate - 2AFC
% test longer than reference response
Test Intervals (ms)
e
Auditory Baseline
Auditory 1.5Hz
Auditory 6Hz
ND Adapting to Auditory Rate - 2AFC
% test longer than reference response
Test Intervals (ms)
f
Visual Baseline
Visual 1.5Hz
Visual 6Hz
ND Adapting to Visual Rate - 2AFC
% test longer than reference response
Test Intervals (ms)
0
20
40
60
80
100
280
300
320
340
360
380

Figures 6.4a-f. The after-effect of adapting to different temporal rates of 1.5Hz (blue curve) and 6Hz (green curve) demonstrated through two-alternative forced-choice is shown for auditory (left column) and visual (right column) conditions. Three different subjects are plotted at the top, middle and bottom of the panel (as before).

Results from the auditory condition (Figure 6.4a) of the 2AFC method show that the Point of Subjective Equality (PSE - the physical test duration producing a perceptual match with the 333ms reference duration) shifts from 323ms (no adaptation baseline) to 303ms or 341ms after adapting to a relatively slow (1.5Hz) rate or fast rate (6Hz), respectively. These distortions represent a rate adaptation-induced distortion of perceived duration. For example, adapting to a fast rate expanded the perceived test interval duration, therefore requiring correspondingly shorter unfilled test durations (and thus a smaller PSE value) to maintain perceptual equivalence with the (non-adapted) test duration. A similar pattern of results is evidenced with the visual condition (Figure 6.4b) and across the other two subjects. From eyeballing the data it is clear that the green function moves one way (with respect to the baseline), whereas the blue function is shifted in the opposite direction. Calculating 95% confidence intervals for the PSE of each function indicated that this difference was significant.

Results suggest that adapting to a given rate does indeed influence the temporal perception of a single empty interval. This effect is present bi-directionally, across 1.5Hz and 6Hz of adapting frequencies and is also evidenced through both, interval reproduction and two-alternative forced choice methods. Consistency in results between two distinctly different experimental methods strongly verifies the relationship between single intervals and rate.

Ultimately, our results show that after-effects of adapting to rate influence the perception of single intervals of time (Figure 6.4a-f). If such effects were truly the result of being exposed exclusively to rate, no such after-effects should be evidenced when the test period presents any temporal input other than rate. Evidence of adaptation to rate influencing single intervals therefore suggests that actually, the adaptation is not to rate specifically, but rather to the durations of repeatedly presented intervals.

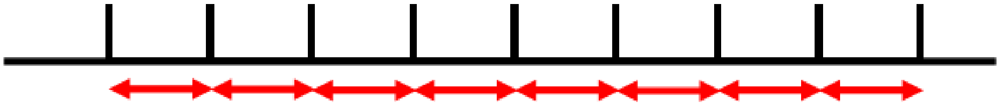
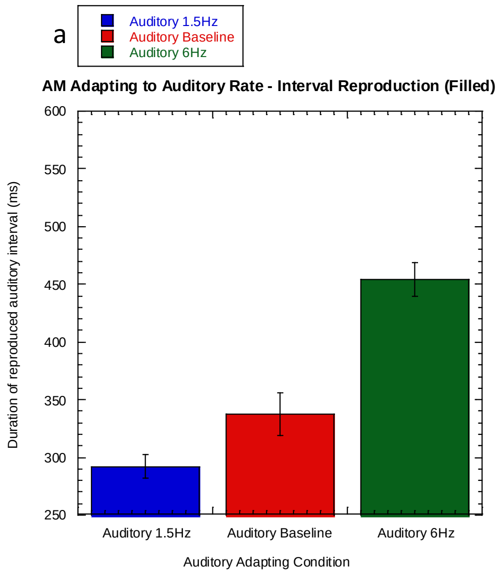
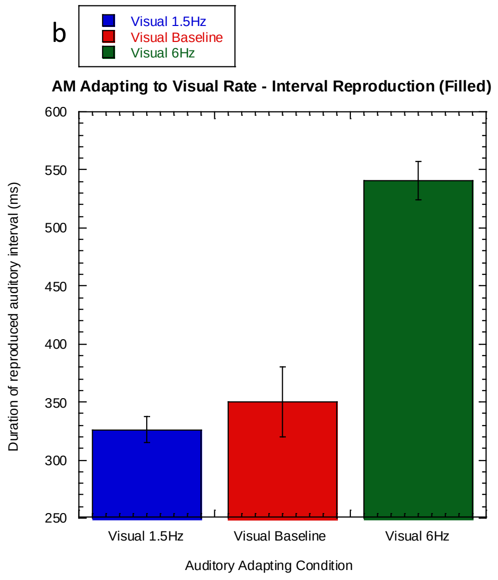
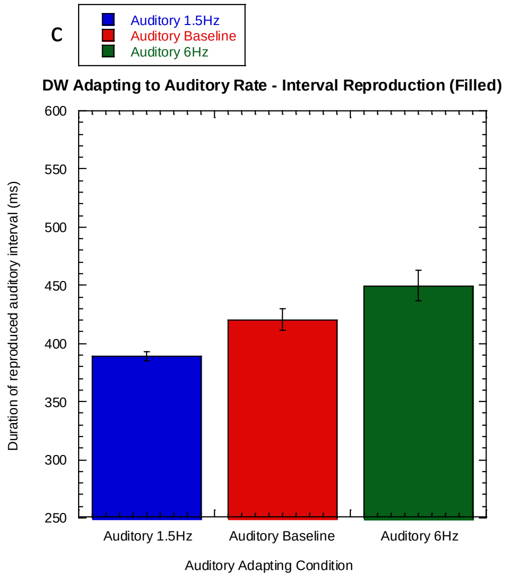


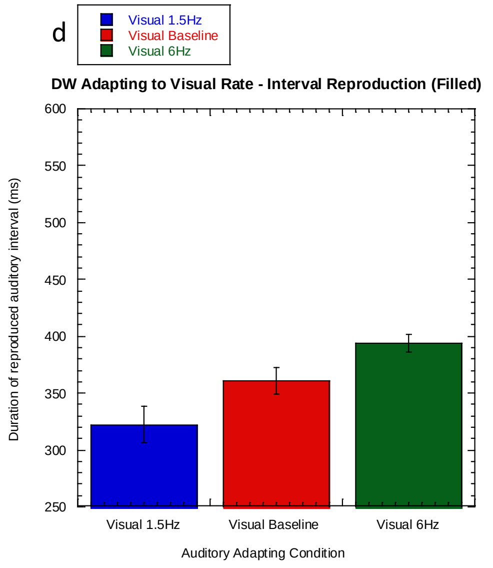
Figure 6.6. A comparison of rate and single intervals: the sequence in black indicates a typically rhythmic pattern whereas the red arrows indicate single intervals being repeated regularly (see text for further detail).

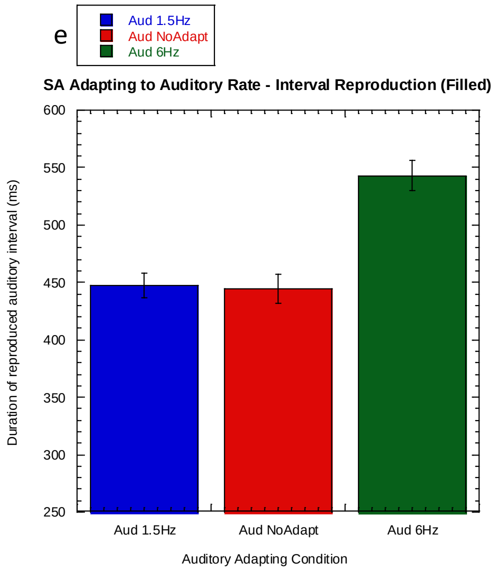
Despite gross phenomenological differences between a single empty interval and rate, the two share an undeniable physical commonality – empty intervals delineated by brief sensory stimuli. We therefore ran a further control experiment testing whether adaptation to rate is able to manipulate the perceived duration of a stimulus that is both physically and phenomenologically different to the adaptor. Participants were asked to reproduce the duration of a *filled* interval (an interval demarcated by continuous signaling). To recap, this meant that the subject was exposed to 10 seconds of adaptation to a particular rate (as before), and then presented with a continuous reference interval lasting 333ms. As before, they were then instructed to reproduce this interval using a continuous keypress to indicate the beginning and end of the test interval. Results from three subjects are presented below:

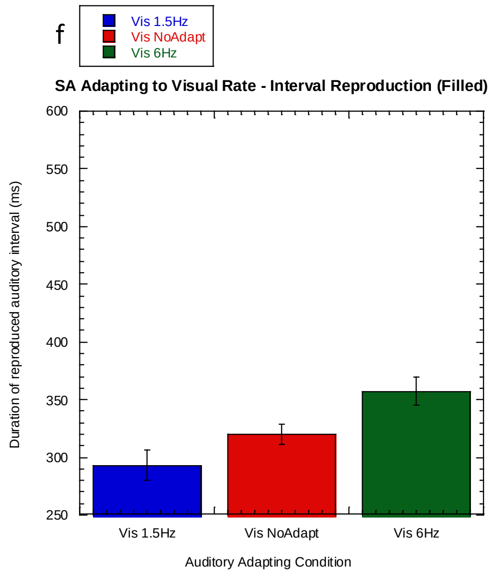

a
Auditory 1.5Hz
Auditory Baseline
Auditory 6Hz
AM Adapting to Auditory Rate - Interval Reproduction (Filled)
600
550
500
450
400
350
300
250
Duration of reproduced auditory interval (ms)
Auditory 1.5Hz
Auditory Baseline
Auditory 6Hz
Auditory Adapting Condition


b
Visual 1.5Hz
Visual Baseline
Visual 6Hz
AM Adapting to Visual Rate - Interval Reproduction (Filled)
600
550
500
450
400
350
300
250
Duration of reproduced auditory interval (ms)
Visual 1.5Hz
Visual Baseline
Visual 6Hz
Auditory Adapting Condition


c
Auditory 1.5Hz
Auditory Baseline
Auditory 6Hz
DW Adapting to Auditory Rate - Interval Reproduction (Filled)
600
550
500
450
400
350
300
250
Duration of reproduced auditory interval (ms)
Auditory 1.5Hz
Auditory Baseline
Auditory 6Hz
Auditory Adapting Condition


d
Visual 1.5Hz
Visual Baseline
Visual 6Hz
DW Adapting to Visual Rate - Interval Reproduction (Filled)
600
550
500
450
400
350
300
250
Duration of reproduced auditory interval (ms)
Visual 1.5Hz
Visual Baseline
Visual 6Hz
Auditory Adapting Condition


e
Aud 1.5Hz
Aud NoAdapt
Aud 6Hz
SA Adapting to Auditory Rate - Interval Reproduction (Filled)
600
550
500
450
400
350
300
250
Duration of reproduced auditory interval (ms)
Aud 1.5Hz
Aud NoAdapt
Aud 6Hz
Auditory Adapting Condition


f
Vis 1.5Hz
Vis NoAdapt
Vis 6Hz
SA Adapting to Visual Rate - Interval Reproduction (Filled)
600
550
500
450
400
350
300
250
Duration of reproduced auditory interval (ms)
Vis 1.5Hz
Vis NoAdapt
Vis 6Hz
Auditory Adapting Condition

Figures 6.7a-f. The after-effect of adapting to different temporal rates of 1.5Hz (blue bar) and 6Hz (green bar) demonstrated through interval reproduction for auditory (left) and visual (right) *filled* interval conditions. Upper panel represented subject AM's results whereas the middle panel represents subject DW's results and the lowest panel represents results from subject HS. Values plotted are indicative of mean intervals reproduced and error bars indicate standard error.

As before, paired-samples t-tests were used to compare filled duration reproduction after adapting to 1.5Hz and 6Hz unimodal rhythms, results of which were found to be statistically significant for both, visual and auditory conditions ($p<.05$) for each subject. These results suggest that typical after-effects as a result of exposure to temporal frequencies can be evidenced when subjects are presented with filled intervals, as have also been demonstrated earlier for empty intervals.

6.5 Discussion

Using a combination of psychophysical experiments, namely temporal reproduction and two-alternative forced choice, we investigated the ability of human subjects to adapt to unfilled intervals of time across the auditory and visual sensory modes. A second set of experiments examined whether the effects of adapting to varying temporal frequencies (1.5Hz, 3Hz and 6Hz) also presented to one of the aforementioned sensory modalities consequently influenced the perception of empty and filled intervals of time. Results suggested two main findings, the first of which indicated that subjects are able to rapidly adapt to empty intervals of time, just as they are with filled intervals (Heron et al., 2012) and secondly, that adapting to varying temporal rates rapidly distorts the perception of single empty and filled intervals.

The present findings that subjects are rapidly able to adapt to different visual rates and are also able to communicate this effect through judgements of subsequently presented single empty intervals suggests that what was previously understood as after-effects of adapting to rate (or rhythm), may actually be an extension of duration after-effects of adapting to repeatedly presented single (filled) intervals of time, which are already known to occur (Heron et al., 2012). The picture of results here encapsulates the remarkably dynamic nature of time perception. The present findings support existing ideas that the processing and perception of rate are closely interlinked as information about one is all that is necessary to predict the other (Hartcher-O’Brien et al., 2016). We also find strong consistency between the two different methods used of 2AFC and interval reproduction. Where bi-directional distortions are evidenced when subjects are instructed to reproduce empty intervals after a period of temporal frequency adaptation, similar bi-directional shifts in PSE are evidenced in 2AFC responses. The consistency between comparisons further strengthens the case that subjects are able to rapidly adapt and show the same distortions in perception despite the variety of response tasks presented.

The sensory-specificity of after-effects from exposure to empty intervals of time are also further reinforced here. In our 2AFC experiment, responses were made after comparing the standard interval (presented to a different sense) to a test interval. As noted by Heron and colleagues (Heron et al., 2012) in their exploration of filled intervals, if empty interval adaptation effects were to occur cross-modally, then the perception of both the reference stimulus and the test should become equally distorted; the result of which would be diminished (if any) duration after-effects. In actuality, our results suggest that after-effects prevail and duplicate those observed in the absence of reference stimuli (as gathered from the collection of interval reproduction experiments).

Fundamentally, something as simple as equal durations of sounds being judged as longer than the same duration of a visual signal indicates that there is a difference between these timing systems. Additionally, it has been found that visual intervals must exceed tactile intervals by at least 8.5% to be subjectively judged as equally long (van Erp & Werkhoven, 2004). Whether that is due to differences in rudimental physiological and neural transduction times or due to differences in their respective perceptual systems remains a topic of discussion. Nevertheless, the critical difference in temporal perception across sensory systems cannot be ignored. Previous studies have documented the difference in modality specificity between rate and duration. Specifically, it has been suggested that the perception of duration is sensory-specific (Heron et al., 2012) whereas rate has been shown to be processed cross-modally (Levitan, Ban, Stiles, & Shimojo, 2015). These findings pose a serious concern for the field of human time perception. If the processing of rate (which has been postulated as cross-modal) is simply an extension of the processing of single intervals (which has been reported to act unimodally), at what point does cross-modal temporal integration come into play? One approach to reconcile these findings is to invite other findings of Becker and Rasmussen, who using adaptation to auditory and visual rhythms, suggested that rate is in fact processed unimodally (Becker & Rasmussen, 2007) as rhythm after-effects transferred between ears however failed to transfer cross-modally. The present findings thus present a considerable addition to the growing body of literature and solve this contention - rate is an extension of duration perception and that these features are both modality-specific. Future work may incorporate intermodal stimuli (stimuli demarcating a single interval with signals from two different sensory modalities, for example, a flash and an auditory beep) to further elucidate these processes. Adapting to such stimuli would not only allow us to clarify the extent of cross-modality in rate but also gauge potential cross-modal correspondences across duration and rate. The world as we know is inherently cross-modal and whilst strictly controlled unimodal stimuli in a lab allow us to pinpoint the specifics of time perception, the use of intermodal stimuli would allow our results to extend to more ecologically ‘real’ environments – a generalization that is sometimes unintentionally neglected.

Distinctions between beat-based and interval-based timers in the production of brief durations have been investigated previously (Pashler, 2001b). Explicitly, Pashler wanted to test the existence of beat-based timers by comparing precision between interval processing compared to precision when processing beats. In the first experiment, Pashler exposed participants to a string of regular standard tones and then presented the same participants with two test tones. The task required subjects to compare the interval between the standard tones to the interval between test tones. Pashler’s justification was that if a beat-based timing mechanism had been elicited in the prior presentation of standard tones, optimum efficiency should be elicited in conditions where the intervals between the standard and test tones matched. This effect however, was not found. In other words, the unique benefit afforded (in theory), by beat-based timers did not become evident when processing beat-based stimuli. In the second experiment, participants reproduced the inter-tone interval between two test tones that they had just heard. Entrainment to the beat was more evidenced in this condition. Specifically, the first keypress response clustered around the standard interval. Pashler concludes that this suggests evidence for an interval-based timer that is responsible for the timing of brief intervals but that the same timing mechanism can, in certain instances, operate in a cyclical, loop manner to generate rhythmic responses. As beat-based timers cannot in theory process intervals, Pashler suggests that one timing mechanism is used to process both intervals and also beats (Pashler, 2001b). If a beat-based mechanism truly exists, there is no reason why it should not be activated and mediate the processing of rhythms in some form. The rhythms used in many of the studies conducted hitherto have employed simple rhythms that can be deconstructed into their component intervals considerably easily. This may explain why evidence suggests recruitment of interval-based timers to process these stimuli. The use of more complex rhythms that cannot easily be disintegrated into their component intervals may be a more comprehensive method to test the existence of beat-based and interval-based timing mechanisms.

Literature concerned with examining the perception of isochronous (regular) against anisochronous (irregular) temporal patterns has, on numerous occasions, suggested that discrimination of anisochronous sequences is significantly worse compared to performance for isochronous sequences (Drake & Botte, 1993; Madison & Merker, 2004; Miller & McAuley, 2005; Horr & Di Luca, 2015). Evidence gathered from the present set of experiments suggests that rate may not be a distinctly independent temporal feature and instead is processed as repeatedly presented single intervals of time. These findings are neatly able to add to the body of literature being gathered on anisochrony perception and allow for another explanation of poorer performance for irregular sequences. Since rate as a feature is processed on an interval-by-interval basis, an anisochronous sequence thus fails to allow a consistent representation of an internal mean and instead is processed as a collection of successive, yet unrelated intervals, thereby providing one explanation for poorer performance with irregular compared to regular sequences.

In conclusion, consistently organised temporal information in the form of rhythm, aids the brain’s efficiency in processing temporal information but this does not mean that the processing of simpler info (for instance, single intervals) is compromised. Our results suggest that the same mechanism is shared across duration and rate and that this is modality-specific. This collectively adds to the idea that local temporal processing mechanisms are modality-specific. An important question is to deduce whether temporal channels are sensitive to the temporal frequency or to temporal intervals. In other words, what language do these channels code input in – frequency or temporal distance? Encoding durations as temporal frequencies would be akin to encoding visual input as spatial frequencies in vision. Future work may aim to explore this further.

7. Dissociating rhythm and interval discrimination through unimodal temporal pattern adaptation

A regular (also referred to as periodic) rhythm refers to a rhythm with equally spaced sensory and temporal signals demonstrating perfect regularity in presentation. Deviations from such a rhythmic presentation may be referred to as aperiodic/irregular or anisochronous rhythms. An isochronous pattern is one where all sequential intervals are equal, like a metronome. Thus, all isochronous sequences are rhythmic however not all rhythmic sequences are isochronous (Ravignani, Honing, & Kotz, 2017; Kotz, Ravignani, & Fitch, 2018).

Following on from the debate presented in Chapter 6 regarding interval-based and rhythm-based temporal processing, one approach to clarify the exact nature of processing for intervals and regular rhythms is to study temporal patterns and anisochrony. This is because temporal patterns allow the experimenter to magnify the features of time that the brain selects and prioritises for temporal processing. These features then allow us to further investigate the processing strategies employed for temporal perception.

7.1. Interval discrimination and the Multiple Look Model

In order to assess temporal processing, it is first important to establish what constitutes a single percept of time. In some of the earliest work in this field, it was reported that for visual percepts, the scale within which one stimulus can be distinguished from another lies in the range of 120-240ms, whereas for an auditory durations, the range lies between 120-170ms (Efron, 1970).

Drake & Botte (1993) assessed auditory interval discrimination as part of a sequence with inter-onset intervals ranging from 100-1500ms (Drake & Botte, 1993). In their first experiment, they compared JND discrimination for a single interval compared to that of a sequence. Explicitly, the task presented subjects with two sequences that differed in tempo consecutively, the subject was then required to respond to which they perceived as faster. Using a sample of 4 subjects, it was found that the JND was optimal for medium tempi sequences of 2 and 6 intervals. Generally, it was found that mean JNDs decreased as the number of intervals in the sequence increased, suggesting that as the number of tones in a sequence increases, so does the listener's ability to discriminate the faster of two sequences (Drake & Botte, 1993).

Drake and Botte (1993) also assessed discrimination of anisochronous (irregular) sequences composed of four intervals (lasting a duration of 2 seconds). It was found that performance improved as the sequence grew closer to isochrony (regularity). Also that JNDs for anisochronous sequences were at an average level in between those recorded for single intervals and those recorded for perfectly regular temporal sequences (Drake & Botte, 1993). The authors also note that for both single intervals and regular sequences, musicians displayed higher sensitivity to changes in tempo than non-musicians, this finding held even for irregular sequences. These results highlight the influence that training has on temporal performance.

Drake and Botte use the 'Multiple-Look' Model to explain their findings and suggest that a sequence of successive interval durations allow a better 'capture' of a representative mean compared to being exposed only to a single interval alone. Thus, 'multiple looks' of the same interval in a sequence result in higher temporal sensitivity compared to only one interval. Furthermore, a regular sequence employing a standardised inter-onset interval encourages a decrease in the average sampling error (compared to a single interval) thereby leading to improved discrimination thresholds (Miller & McAuley, 2005).

7.2. Different types of temporal sequences

Anisochrony perception can be assessed by eliciting either 'local' or 'global' anisochrony – or potentially even both (Figure 7.1). 'Local' anisochrony can be generated by creating a 'jitter' or simply by shifting one sound location in an otherwise regular sequence. On the other hand, 'global' anisochrony is created by shifting each individual sound in a sequence from its original isochronic position (Ehrlé & Samson, 2005). The authors used a two-alternative forced-choice design and assessed the role of Weber's Law in anisochrony discrimination. The simplest interpretation of Weber's Law asserts that the value of shifts necessary to discriminate a regular to irregular sequence should be proportional to the IOI of the continuous sequence (expressed either through a percentage of the IOI or a JND) (Halpern & Darwin, 1982; Grondin, 2001). The authors conclude that a good fit was predicted between anisochrony discrimination and Weber's Law but only for IOIs between 250-1000ms and no shorter (Ehrlé & Samson, 2005).

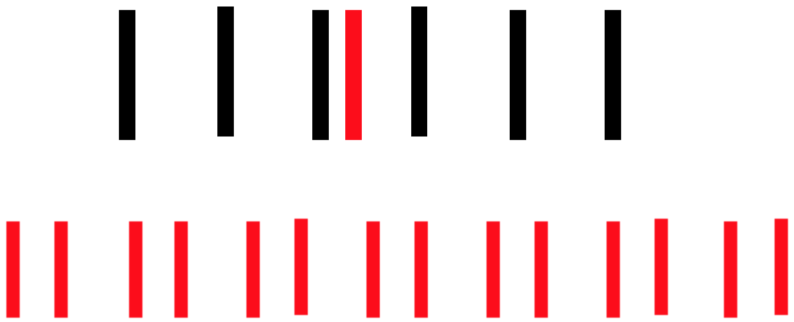


Figure 7.1. An example of local and global anisochrony. Top half of panel indicates an example of local anisochrony (red bar indicates anisochronous signal) whereas lower half of panel indicates an example of global anisochrony (see text for more detail).

7.3. Neural basis behind anisochrony

Interestingly, it has been found that beat perception has even been evidenced in newborn infants (Winkler, Háden, Ladinig, Sziller, & Honing, 2009). Furthermore, that even passive exposure to rhythms is enough to activate and employ motor regions in the brain (Chen, Penhune, & Zatorre, 2008). It is no surprise that temporally predictable (rather than irregularly timed) stimuli allow for more efficient neural and perceptual processing. It has been shown with fMRI and EEG (van

Atteveldt et al., 2015), that participants responded faster and more accurately when stimuli were embedded in rhythmic rather than unpredictable stimuli. Moreover, perception is most accurate when stimulus presentation occurs in line with beat prediction and progressively decreases as the temporal discrepancy between beat prediction and the actual stimulus onset increases (Jones, 1976; Large & Jones, 1999). These findings are explained through the 'dynamic attending theory' which suggests tones presented within a regular sequence entrain attentional oscillations and that these oscillations allow for more efficient processing of subsequent tones presented within a regular (and expected) rhythm (Jones, 1976).

How exactly temporal expectation modulates perception neurally was investigated using a paradigm combining psychophysics with electrophysiological recordings (Cravo, Rohenkohl, Wyart, & Nobre, 2013). Psychophysical data suggested that temporal expectation elicited through regular temporal cues enhanced a visual target's contrast sensitivity. Electrophysiological data gathered simultaneously showed that the phase of delta oscillations (1-4Hz) could be used to predict target processing quality but only in events that were presented within regular streams. Additionally, in anticipation of the predicted events, it was found that the optimum phase for these oscillations also coincided, suggesting that the entrainment phase of low-frequency oscillations enhances sensory processing (Cravo et al., 2013). Irregular temporal sequences prohibit the build-up of expectation (Zeni & Holmes, 2018). The authors thus link the methods by which temporal expectation modulates perception by proposing that contrast sensitivity is enhanced by temporal expectation and that this is accompanied by the phase entrainment of low-frequency oscillations (Cravo et al., 2013).

In another study using duration estimates and neural entrainment it was suggested that timing judgements are mediated by the structural composition of temporal stimuli (Horr, Wimber, & Di Luca, 2016). Using a 2AFC design paired with EEG, the authors presented subjects with one isochronous and another anisochronous sequence and asked subjects to judge which was longer, results demonstrated that isochronous sequences were regularly overestimated whereas anisochronous sequences regularly led to underestimation (Horr & Di Luca, 2015; Horr et al., 2016).

The authors report that this bias appears to be temporally-specific as no distortions are elicited when varying the regularity of non-temporal stimuli features such as pitch or sound intensity. These findings are explained by suggesting that presentation of isochronous intervals leads to an increase in neural responsiveness, which in turn leads to a heightened representation of the specific features of the interval being filled, and ultimately resulting in an increase in perceived subjective duration (Horr et al., 2016). They also support earlier findings of isochronous and predictable sequences leading to longer perceived duration compared to anisochronous and unpredictable sequences.

Isochronous tapping to a visual metronome allowed Merchant et al. (2015) to classify response properties of neurons in the primate medial premotor cortex as either sensory or motor (Merchant et al., 2015). Sensory cells presented in two distinguishable clusters; one cell population group exhibited short response-onset latencies to previously presented stimuli (sensory-driven neurons), whereas the other cell population was actively predicting the incidence of the next stimulus (stimulus-predicting neurons). Clear differences were also observed in the behaviours of these cells, for example, sensory-driven neurons displayed a bias towards visual stimuli whereas stimulus-predicting neurons were bimodal. The authors further expand and suggest that as the task progressed, sensory-driven cells diminished in their functional impact whereas motor cells increasingly gained in importance throughout the task and were likely responsible for the progression of rhythmic taps in the task (Merchant et al., 2015).

Detection of isochrony in more natural environments comes most commonly in the form of beat perception. Beat perception through tapping along to a beat captures a range of cognitive and underlying neural mechanisms, some of which include mental timekeeping and establishing a relationship between perception and action (Patel, Iversen, Chen, & Repp, 2005). Recently, evidence that humans are also able to extract meaningful information about a tactile beat and reproduce the beat in synchrony within the tactile modality has also been presented (Brochard, Touzalin, Despres, & Dufour, 2008). Whilst beat perception is neurally explained through basal ganglia activity, more recent evidence implicates its role specifically in beat prediction (Grahn & Rowe, 2012) whereas absolute timing is supported by the cerebellum (Grube, Cooper, Chinnery, & Griffiths, 2010).

Furthermore, evidence presented also suggests non-metric rhythm is localised to the right hemisphere (Horvath et al., 2011), and also that cross-modal effects are present in the perception of meter (Celma-Miralles, de Menezes, & Toro, 2016). This result has recently been extended to cross-modal stimuli (Su, 2016). Using point-light figure dances and auditory metrical rhythms, Su found an influence of visual rhythm on auditory detection suggesting a multisensory integration of metric perception (Su, 2016).

Temporally periodic stimuli commonly occur in the natural world, for instance in footsteps. It has been found that periodic stimuli contribute to temporal predictability and the simultaneous neural entrainment that occurs. These have widely been thought to contribute a perceptual advantage via perceptual enhancement of those periodic signals (Lawrance, Harper, Cooke, & Schnupp, 2014). More recently, it has even been found that temporal predictability and expectation can be marked by the pupillary response (Akdoğan, Balcı, & van Rijn, 2016).

7.4. Temporal irregularity and the senses

Temporal detection in anticipation and prediction have been found to be influenced by the rhythmicity of the stimulus and sensory cues. It was hypothesized that temporally predictive information combined within and across sensory modalities should facilitate the detection of sound, exceeding the advantage provided by unisensory cues alone (ten Oever, Schroeder, Poeppel, van Atteveldt, & Zion-Golumbic, 2014). Indeed, two experiments requiring participants to detect tones within noise that was either random or rhythmic found that detection was improved on rhythmic versus random trials. Interestingly, on half the trials, the experimenters presented a predictive visual signal before the sequence of tones that was to be detected within the noise. Detection was improved for audio-visual, compared to audio only presentation. They conclude that to optimally process predictable sounds and enhance detection, information from both the available sensory modalities and rhythmicity is used (ten Oever et al., 2014). They suggest that combining multiple cues results in improved temporal estimation and suggest our sensory systems dynamically adopt all useful information within the surrounds to process the world efficiently.

It has been established that isochronous temporal patterns are both better discriminated and more accurately reproduced when presented to the auditory modality compared to the visual modality (Handel & Buffardi, 1968; Glenberg & Jona, 1991). Patel and colleagues used a cross-modal tapping design to auditory and visual isochronous and anisochronous beats and compared how metrical structure affects synchronisation to a beat period. It was reported that for auditory patterns, synchronisation to a beat remained relatively stable when presented alongside strong isochronous rhythms. The presence of non-isochronous rhythms however, resulted in an impairment in tapping. For visual patterns, it was reported that subjects were unable to synchronise their motor responses to either non-isochronous rhythms or very fast isochronous rhythms. The authors propose that beat perception and synchronisation hold unique affinity with the auditory system (Patel et al., 2005), a result supported with ERP data (Pasinski, McAuley, & Snyder, 2016) and other work (Grahn, 2012).

Furthermore, in a design using patterned sequences of tones and an assessment of temporal jitter, it was concluded that a separate channels model best explained the data (Sorkin, Boggs, & Brady, 1982). They assert the nature of the multichannel system in which incoming marker signals are simultaneously segregated by different spectral bands. Nevertheless, classical musicians regularly follow the visual beat of a conductor and so high level beat perception and synchronisation are functional in the visual modality, but perhaps only evidenced after periods of extensive training (Patel et al., 2005).

7.5. Factors affecting sequence perception

In conflicting evidence to assertions made by the Multiple Look Model, it has been reported that discrimination of jitter in an isochronous pattern is not dependent upon the number of sounds in the sequence – as such, performance remains stable despite an increase in the number of sounds in an isochronous sequence (Ehrle & Samson, 2005). The authors also implicate left temporal lobe structures in perceiving inter-onset-interval increments in isochronous sequences but also in familiar musical tunes (Samson, Ehrle, & Baulac, 2001) and assert the specialised role of left temporal lobe structures in the processing of fast auditory patterns (Ehrle, Samson, & Baulac, 2001).

The influence of ascending and descending temporal structures on durations has also been documented (Matthews, 2013). Nevertheless, a large proportion of literature on anisochrony revolves around the detection of tempo change and difficulties in interpreting the range of findings have been documented (Madison, 2004), mainly due to the differences in stimuli and methods. For instance, stimuli have included drumbeats (Miller & Eargle, 1990), sequences of metronome sounds (Kuhn, 1974), distorted samples of real music (Geringer & Madsen, 1984) and music-like stimuli (Ellis, 1991; Madison, 2004). Many of these studies have focused on subjects crudely detecting the tempo and whether this is increasing or decreasing rather than assessing sensitivity in anisochrony and temporal sensitivity with comprehensive methods. For instance, some methods involve tempo changes that occur discretely and rhythmically, for example, with every nth signal being displaced rather than the whole sequence being dynamically displaced and creating a global sense of anisochrony.

Upon closer inspection of this field it becomes apparent that no clear consensus has been reached. This may be in part due to the difficulty in interpreting the range of results reported due to contrasting methodologies for instance, constant stimuli, adaptive stimuli and stimuli requiring adjustment all present varying results (Ehrlé & Samson, 2005). Moreover, varying detail in methodology such as IOI's, pitch, luminance, number of intervals in a sequence and so forth, all contribute to the breadth of research conducted in this field. It will be the aim of future research in the field to consolidate these findings in more comprehensive research assessing when exactly the number of intervals in a sequence act influentially.

Sensory timing forms an integral part of human experience. Hitherto, findings have largely added to the understanding that statistical regularities in natural environments aid perceptual constancy and maintain efficient processing (Rhodes & Di Luca, 2016). Moreover, that adapting to a fast tempo makes a medium tempo feel slower (Chapter 4 and 5). Despite the progress made in the field, key questions still remain to be elucidated. Notably, these include the influence of repeating intervals in the perception of rhythms and to what extent these dictate the human experience of rhythm. Expanding on this, are humans able to adapt to anisochronous (irregular) rhythms? And to what extent? Evidence from previous chapters (5 and 6) indicate that perceptual systems are rapidly able to adapt to temporal signals, importantly, however, it remains to be investigated whether the same systems are able to adapt to anisochronous signals. Will prolonged exposure to such rhythms result in anisochronous rhythms being perceived as more regular? Or will exposure to such rhythms lead to isochronous (regular) rhythms being perceived as more irregular also?

There is now mounting evidence that adaptation to specific temporal features elicits after-effects akin to other features of sensory processing such as visual motion and orientation. The present study aims to temporally adapt subjects to patterned temporal sequences (long interval – short interval – long interval and so on) and investigate whether consecutive regularity in time is a prerequisite for such temporal after-effects. The following experiments specifically aim to explore how patterned rhythms shape the perception of more complex temporal sequences, and how adaptive exposure to rhythms varying in anisochrony influence the perception of regularity across the audio, visual and tactile modalities using the 2AFC and single stimulus methods. There are two potential outcomes to this experiment, the first is that subjects will be able to adapt to patterned rhythms as demonstrated in shifted thresholds for what constitutes isochrony. The second possible outcome is that subjects will not adapt to patterned rhythms resulting in anisochronous stimuli being perceived as more irregular.

7.6. Methods

To assess the perception of temporal regularity in unimodal conditions, we ran a combination of adaptive two alternative forced choice and criterion-dependent response measures across the auditory, tactile and visual modalities. These experiments involved being presented with two sequential trains of sequences, one that was temporally regular and another that was temporally irregular via patterned sensory presentation. Subjects responded either by responding to whether the sequence was regular (for the single stimulus experiment) or they responded to which of the two sequences was regular (in the two-alternative forced-choice experiment). The subject was instructed to respond by pressing a keyboard key indicating which they perceived as irregular. A range of anisochronies was adapted to and presented on each trial.

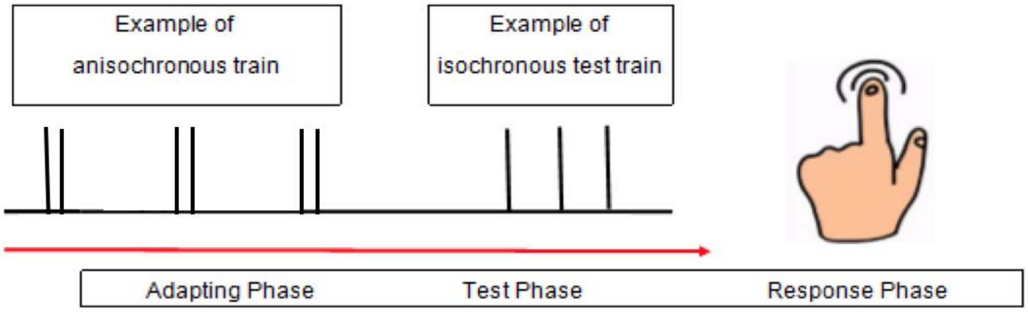


Figure 7.2. Schematic demonstrating an example trial where the subject is exposed to an anisochronous adapting sequence followed by an isochronous test sequence before responding to whether the test sequence was regular or not via keypress. This trial demonstrates a typical single stimulus method trial that was employed. Response phase drawing accessed from www.iconsmind.com.

Subjects: 3 participants (1 female and 2 males, mean age = 33, standard deviation = 15 years) participated, with self-reported normal hearing and visual abilities. Following initial practice sessions, a lengthy process (20-25 hours) of data collection began, in a series of sessions spread over several weeks. Two of the participants had previous experience of psychophysical data collection. The third participant (KH) had no such experience and was naïve to the purpose of the experiments. The experiments received ethical approval from the Research Ethics Committee at the School of Optometry and Vision Sciences, University of Cardiff and all experiments were performed in accordance with relevant guidelines and regulations. Informed consent was obtained for study participation.

Stimuli: Brief (16msec duration) sensory stimuli were presented – either in the auditory, visual or somatosensory modality and all stimuli were grossly suprathreshold. Stimulus generation and presentation was controlled by an Intel ® Core ™ i5-4460 desktop computer running Microsoft Windows 7. The programming environment involved MATLAB 8.6 (Mathworks, USA) in combination with Psychophysics Toolbox 3 (http://www.psychtoolbox.org). Stimulus timing was verified using a dual-channel oscilloscope.

Visual:

Visual stimuli were presented via a 20mm LED bulb presenting a highly luminous green/yellow flash at 568nm. These were bright flashes (274 cd/m^2) presented centrally against a uniform dark background (0.32cd/m^2), and lasting 16ms. At the viewing distance of 60cm the circular flash subtended a diameter of approximately 2° of visual angle.

Auditory:

Auditory stimuli consisted of brief (16ms duration) bursts of white noise generated by a Xonar Essence STX (ASUS) soundcard (https://www.asus.com/us/Sound-Cards/Xonar_Essence_STX/) with a sampling rate of 44,100Hz. Stimuli were delivered using Sennheiser HD280 Pro Headphones at an SPL of 70dB.

Tactile:

Tactile stimuli were produced using the amplified (LP-2020A+ Lepai Tripath Class-T Hi-Fi Audio Mini Amplifier) ‘audio-out’ voltage of the sound card which controlled a miniature electromagnetic solenoid-type stimulator (Dancer Design Tactor www.dancerdesign.co.uk/products/tactor.html). Using brief (16ms) audio bursts of white noise the Tactor produced taps to the index finger of the left hand. The Tactor was enclosed within a fabric occluder in order to eliminate the possibility of auditory cues.

Procedure

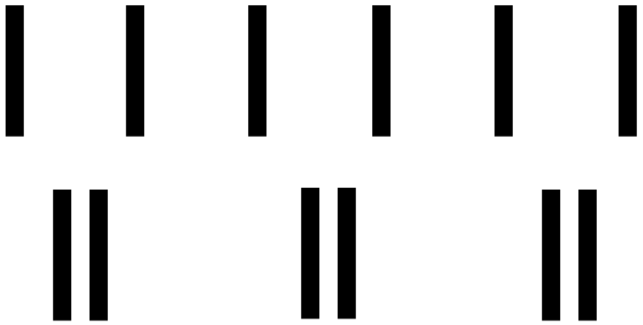


Figure 7.3. Schematic depiction of a temporally regular control sequence of 3Hz (upper panel) versus a temporally offset anisochronous sequence (lower panel). The anisochronous conditions included 3 possible shifts in a 3Hz temporal sequence of either 25, 50 or 100ms. After adapting to an irregular sequence, subjects were exposed to a second sequence and instructed to respond to whether this sequence was regular or not (via keypress). See text for further descriptions of task.

Experiment 1a

In Experiment 1, subjects adapted to a temporal sequence of either 0, 25, 50 or 100ms offset from 3Hz synchrony. This meant that each alternate signal was presented either 0, 25, 50 or 100ms early and the next signal would then be presented with a delay of the same time (Figure 7.3). This manner of irregularity resulted in a ‘temporal pattern sequence’ being presented where one interval was longer than average and the next interval would be shorter than average, in other words, an irregular rhythm. Which interval began the sequence (either long or short) was randomised on each trial to prevent the subject from attempting to familiarise themselves with the sequence. Subjects were then presented with a test sequence averaging 3Hz but offset by 0-100ms in the same way. Subjects were instructed to indicate whether the test sequence was ‘regular’ or not via a keypress (‘<’ for not regular, ‘>’ for regular). A two-alternative forced-choice (2AFC) experiment was also run to gather baseline measurements of subject’s thresholds for regularity. Subjects were presented at random with a reference isochronous sequence of 3Hz presented to one modality, followed by a test sequence presented to the same modality and also averaging 3Hz yet offset in the arrival of each sensory signal (anisochronous) by 0-100ms for each signal. Subjects indicated which sequence was regular via keypress (‘<’ for the first sequence, ‘>’ for the second sequence). Both experiments were run on all three subjects and with all three sensory modalities (audition, vision, touch). A minimum of 30 trials were gathered for each subject across each condition and sensory modality combination.

A single trial began with an adaptation period (8-10 seconds) of a pre-chosen temporal nature (ranging between 0-100ms anisochrony). This was followed by a test phase (lasting between 2 and 2.5 seconds) of either a regular or irregular sequence. A method of single stimuli allowed response data to be generated by instructing subjects to respond to whether the sequence was ‘regular’ by pressing a key on a keyboard. This data was then used to plot psychometric thresholds for each subject.

Baseline data was gathered (data without any adaptation) using both, a method of single stimuli but also a two-alternative forced-choice design (2AFC). Justification behind this was to allow a more conservative and precise measure of anisochrony sensitivity.

Experiment 1b

In two-alternative forced choice experiments, the first sequence presented was randomly either regular or irregular thereby, allowing each trial to present a different order of regular or irregular sequences and inhibiting any attempts by the subject to 'guess' which sequence was regular. After presentation of both sequences, subjects were instructed to respond to "Which sequence was regular, the first or second?" again, by pressing a key on a keyboard. Whether the first interval of the sequence presented was long or short was determined randomly for all sequences presented across baseline single stimulus and two-alternative forced-choice experiments. Justification behind this was also to expose the subject to as 'pure' a form of temporal anisochrony as we could and prohibit the subject from adapting to a predictable rhythm.

Experiment 2

Due to the variability that can sometimes result when using criterion-dependent measures (such as the first single stimulus experiment), the second experiment attempted to expose subjects to identical periods of anisochrony adaptation as the first experiment however, within this experiment subjects would be tested with a 2AFC method. To elaborate, subjects were presented with a baseline 2AFC method identical to the first experiment. In the experimental conditions, the subject adapted to either a perfectly regular sequence presented at 3Hz (and therefore with a 0ms offset) or an irregular sequence averaging 3Hz but offset by 100ms. Subjects were then presented with a 2AFC judgement where at random a reference sequence was presented at 3Hz, followed by a period of top-up adaptation (identical to the first adapting sequence) and then the test sequence also averaging 3Hz however offset by either 0-50ms. Subjects then indicated which sequence they perceived as regular via a keypress ('<' for the first sequence, '>' for the second sequence).

Each experimental block incorporated a total of 10 repeats of each test sequence, and this was repeated at least 3 times resulting in a minimum of 30 trials for each condition. A break of at least 3 minutes was then taken before adapting to a different temporal irregularity - this was to ensure no adaptation effects crossed-over from one run to the next. The order of testing conditions was randomised for each condition (adapting level) and each sensory modality.

Experiment 2 was run on all three subjects and with all three sensory modalities (audition, vision, touch) and all conditions, resulting in a total of a minimum of 540 trials per subject across auditory, tactile and visual modalities.

Both Experiments (1 and 2), were blocked according to sensory modality. The order of blocks and level of adaptation in each was randomised. This was kept consistent throughout all experiments and subjects in this chapter.

7.7. Results

As in the study of Heron et al. (Heron et al., 2012), data from the single stimulus and 2AFC experiments were fitted with a psychometric function using the subject's discrimination judgement of the proportion of responses that were correctly identified as anisochronous. The functions were then fitted with a logistic of the form

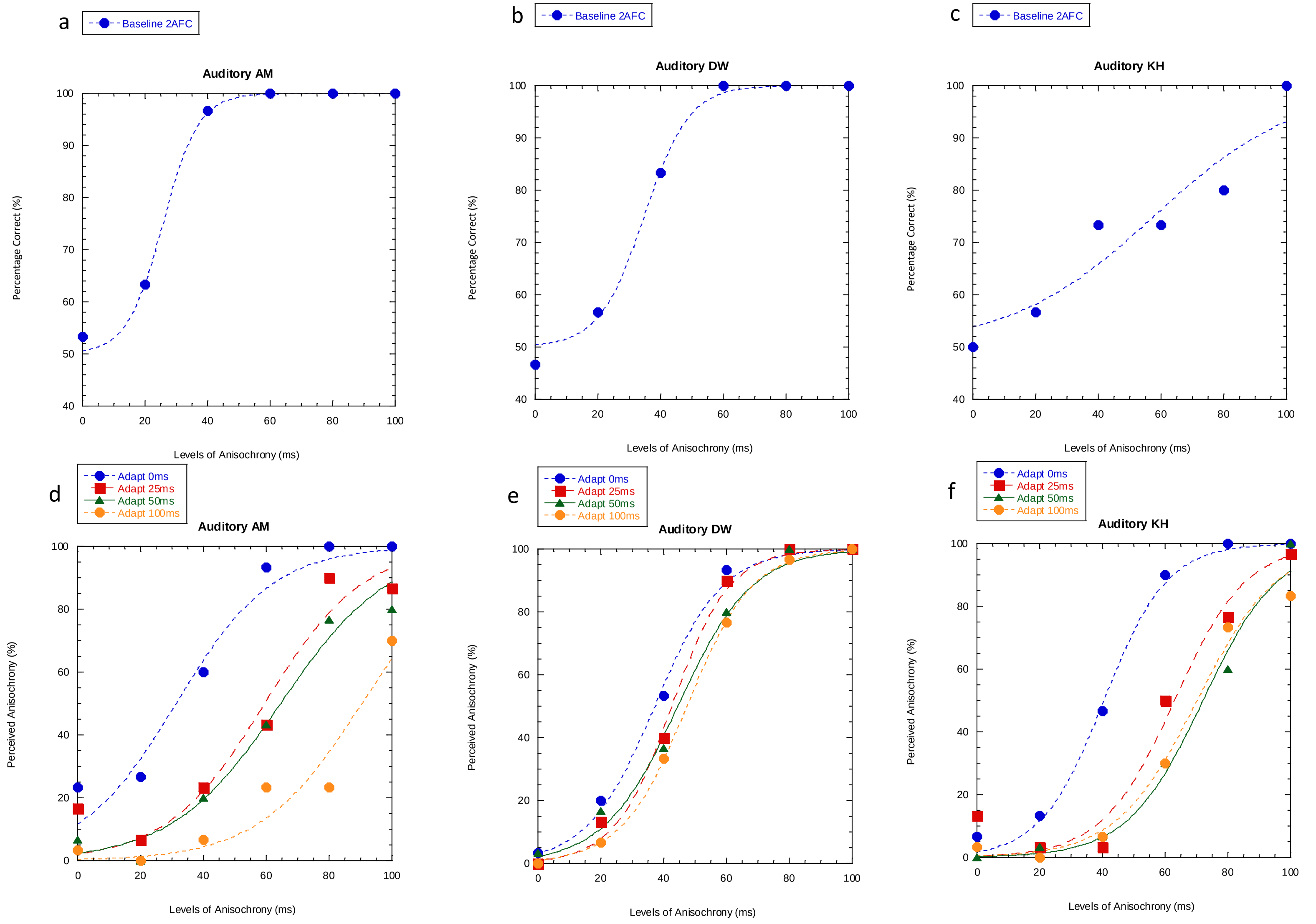
$$y = 50 + (50 + e^{-\frac{\alpha}{\theta}})$$

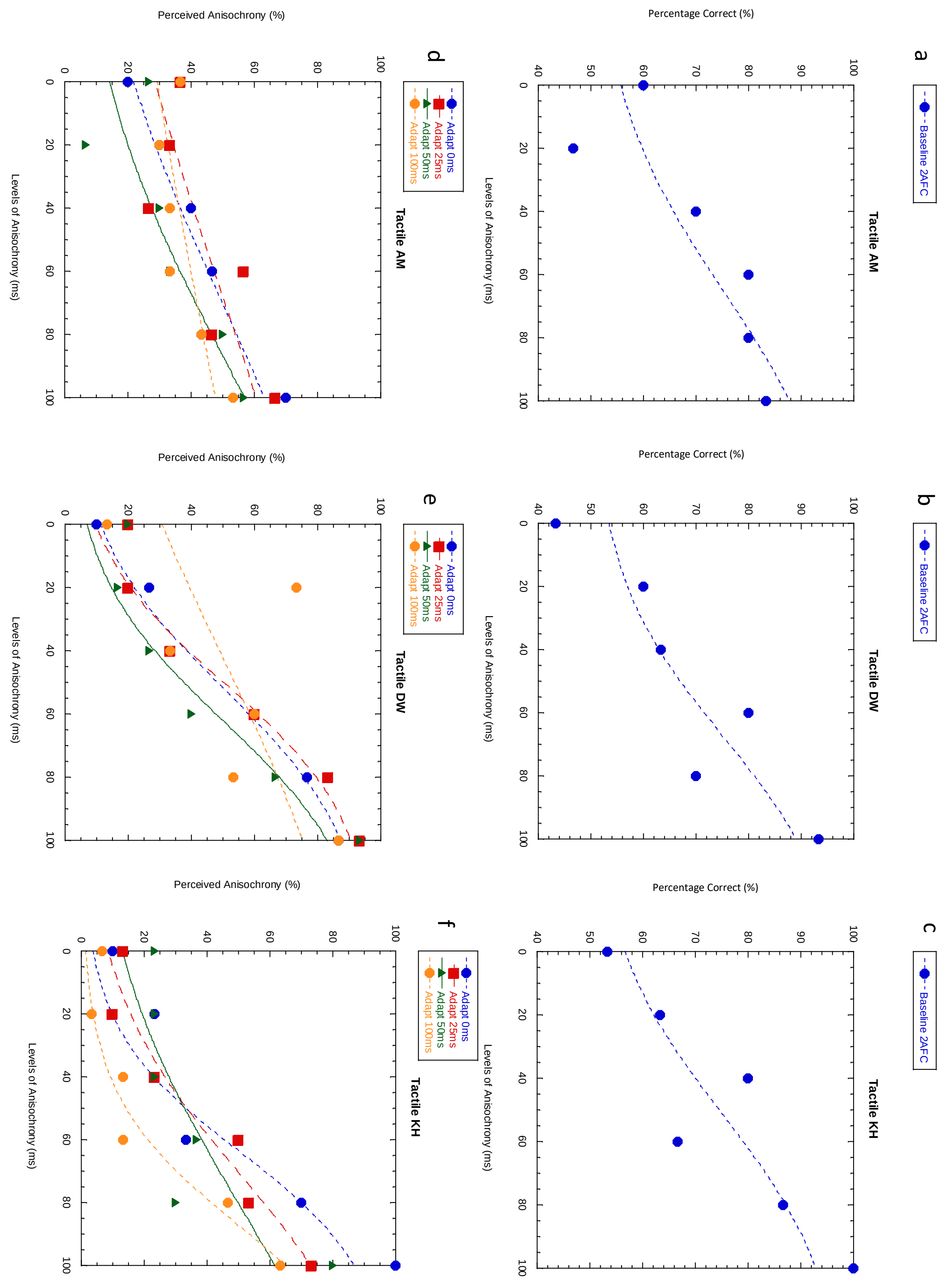
where '$\boldsymbol{\alpha}$' denotes the point of subjective equality (PSE – the 75% response level on a psychometric function) and '$\boldsymbol{\theta}'$ denotes an estimate of the anisochrony discrimination threshold.

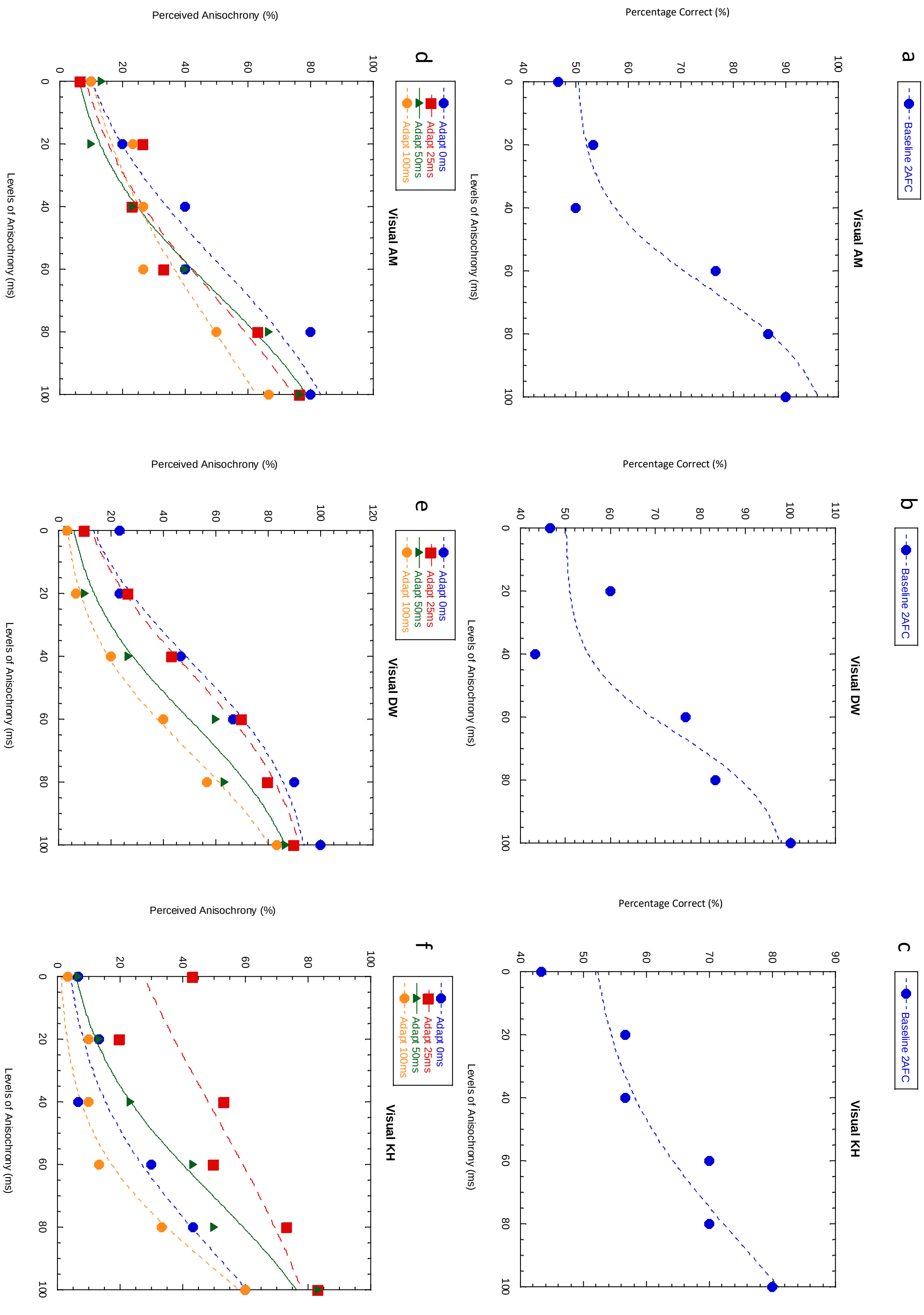
Subjects were first tested with a baseline two-alternative forced-choice method randomly comparing an entirely uniform sequence centred around 3Hz, with a 'patterned' sequence also centred around 3Hz, with signals offset by one of six different possibilities ranging from 0-100ms. Subjects were then adapted to various levels of temporal patterns (averaging 3Hz but offset by either 0ms, 25ms, 50ms or 100ms) across three modalities of audition, touch and vision, and tested with a single stimulus method. Subjects responded to whether the test sequence presented was regular or not via a keypress.

Experiment 1a and 1b

Data plotted below are results for all three subjects from Experiment 1a and 1b.


a
Baseline 2AFC
Auditory AM
Percentage Correct (%)
40
50
60
70
80
90
100
0
20
40
60
80
100
Levels of Anisochrony (ms)
b
Baseline 2AFC
Auditory DW
Percentage Correct (%)
40
50
60
70
80
90
100
0
20
40
60
80
100
Levels of Anisochrony (ms)
c
Baseline 2AFC
Auditory KH
Percentage Correct (%)
40
50
60
70
80
90
100
0
20
40
60
80
100
Levels of Anisochrony (ms)
d
Adapt 0ms
Adapt 25ms
Adapt 50ms
Adapt 100ms
Auditory AM
Perceived Anisochrony (%)
0
20
40
60
80
100
0
20
40
60
80
100
Levels of Anisochrony (ms)
e
Adapt 0ms
Adapt 25ms
Adapt 50ms
Adapt 100ms
Auditory DW
Perceived Anisochrony (%)
0
20
40
60
80
100
0
20
40
60
80
100
Levels of Anisochrony (ms)
f
Adapt 0ms
Adapt 25ms
Adapt 50ms
Adapt 100ms
Auditory KH
Perceived Anisochrony (%)
0
20
40
60
80
100
0
20
40
60
80
100
Levels of Anisochrony (ms)


a
Baseline 2AFC
Tactile AM
Percentage Correct (%)
Levels of Anisochrony (ms)
b
Baseline 2AFC
Tactile DW
Percentage Correct (%)
Levels of Anisochrony (ms)
c
Baseline 2AFC
Tactile KH
Percentage Correct (%)
Levels of Anisochrony (ms)
d
Adapt 0ms
Adapt 25ms
Adapt 50ms
Adapt 100ms
Tactile AM
Perceived Anisochrony (%)
Levels of Anisochrony (ms)
e
Adapt 0ms
Adapt 25ms
Adapt 50ms
Adapt 100ms
Tactile DW
Perceived Anisochrony (%)
Levels of Anisochrony (ms)
f
Adapt 0ms
Adapt 25ms
Adapt 50ms
Adapt 100ms
Tactile KH
Perceived Anisochrony (%)
Levels of Anisochrony (ms)
40
50
60
70
80
90
100
0
20
40
60
80
100



a
Visual AM
Baseline 2AFC
Percentage Correct (%)
Levels of Anisochrony (ms)
b
Visual DW
Baseline 2AFC
Percentage Correct (%)
Levels of Anisochrony (ms)
c
Visual KH
Baseline 2AFC
Percentage Correct (%)
Levels of Anisochrony (ms)
d
Visual AM
Adapt 0ms
Adapt 25ms
Adapt 50ms
Adapt 100ms
Perceived Anisochrony (%)
Levels of Anisochrony (ms)
e
Visual DW
Adapt 0ms
Adapt 25ms
Adapt 50ms
Adapt 100ms
Perceived Anisochrony (%)
Levels of Anisochrony (ms)
f
Visual KH
Adapt 0ms
Adapt 25ms
Adapt 50ms
Adapt 100ms
Perceived Anisochrony (%)
Levels of Anisochrony (ms)

Figures 7.4a-c (top panel) demonstrate auditory baseline data from Experiment 1 for all three subjects (AM – left, DW – centre and KH – right). Figures 7.4d-f (lower panel) demonstrate auditory adapting data for the same subjects in Experiment 1 after having adapted to auditory sequences of 0, 25, 50 and 100ms anisochronies. Figures 7.5a-c (top panel) demonstrate tactile baseline data from Experiment 1 for all three subjects (AM – left, DW – centre and KH – right). Figures 7.5d-f (lower panel) demonstrate tactile adapting data for the same subjects after having adapted to tactile sequences of 0, 25, 50 and 100ms anisochronies. Figures 7.6a-c (top panel) demonstrate visual baseline data from the same experiment for the same subjects (AM – left, DW – centre and KH – right). Figures 7.6d-f demonstrate visual adapting data for the same subjects after having adapted to visual sequences of 0, 25, 50 and 100ms anisochronies.

For all baseline conditions, the mid-point of the curve (demonstrating the probability at which the subject will accurately detect an irregular sequence on 75% of trials) is reported below. Also reported are the standard errors associated with this value. For all other conditions, 95% confidence intervals were calculated on the PSE values (point of subjective equality at which the test sequence is equally likely to be considered perfectly regular or 'irregular') and their respective standard errors for threshold plots from Experiment 1b. The results of these are below:

Auditory			
	AM	**DW**	**KH**
Baseline	25.6 ± 0.9	34.7 ± 1.3	57.7 ± 6.9
Adapt 0ms	31.3 ± 8.8	37.1 ± 2.9	40.4 ± 2.9
Adapt 25ms	59.5 ± 11.6	42.6 ± 3.0	62.9 ± 8.3
Adapt 50ms	64.6 ± 7.8	44.5 ± 4.1	72.0 ± 5.6
Adapt 100ms	90.4 ± 10.5	47.3 ± 0.7	70.1 ± 5.1
Tactile			
	AM	**DW**	**KH**
Baseline	64.7 ± 11.0	67.2 ± 10.5	51.1 ± 10.5

Adapt 0ms	70.4 ± 19.3	51.7 ± 5.3	63.5 ± 14.8
Adapt 25ms	68.1 ± 38.2	49.9 ± 8.6	70.3 ± 10.7
Adapt 50ms	86.8 ± 29.5	62.1 ± 13.2	80.4 ± 37.0
Adapt 100ms	111.9 ± 68.8	42.4 ± 55.0	87.2 ± 8.6
Visual			
	AM	**DW**	**KH**
Baseline	65.0 ± 5.0	65.4 ± 6.7	86.3 ± 8.3
Adapt 0ms	57.6 ± 12.9	41.2 ± 9.4	88.1 ± 10.6
Adapt 25ms	61.7 ± 12.7	44.7 ± 4.0	43.4 ± 28.6
Adapt 50ms	67.7 ± 7.0	60.1 ± 9.2	70.3 ± 9.3
Adapt 100ms	80.8 ± 14.3	70.5 ± 4.1	93.2 ± 7.5

Table 7.1. 95% Confidence interval values. Specifically, the values in the table describe the PSE for each condition for each subject, ± the confidence interval for that condition.

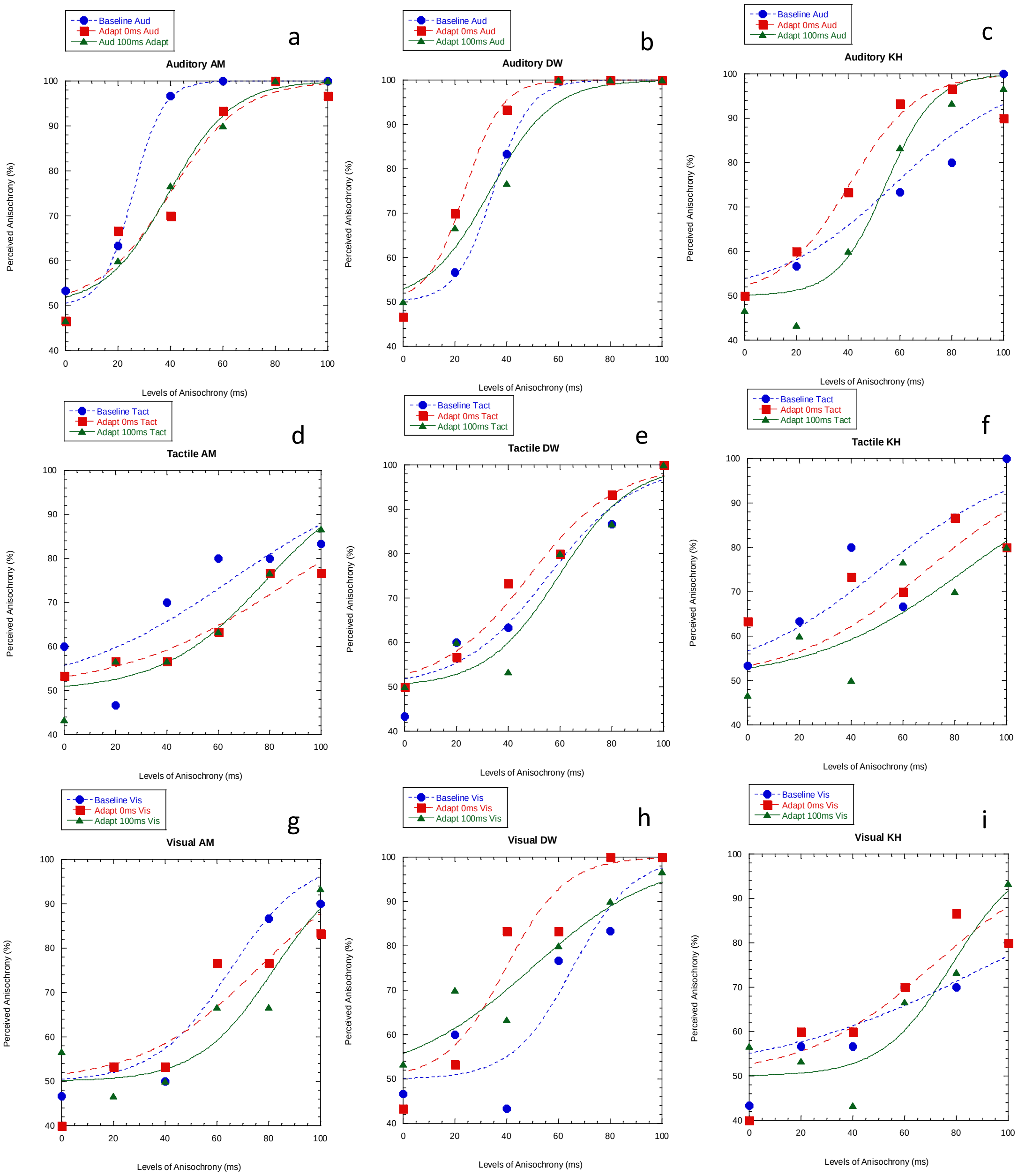
If we find that humans are able to adapt to more complex temporal sequences, such as patterns, we may expect to find that baseline and adapt 0ms PSE values should be relatively close to one another. We may also expect to find that after adapting to various temporal patterns the same PSE values should increase if we are able to adapt to patterns as a larger offset in the patterns would be needed to elicit the same magnitude of irregularity. On the other hand, if adapting to patterns in fact makes the same subject more sensitive to temporal offsets, we would expect to see the same JND values to decrease post-adaptation to patterns.

In Figures 7.4a-f, the data loosely show that after adapting to an irregular rhythm, a much greater test anisochrony is needed before the stimulus is perceived as anisochronous. Adapting to an irregular auditory rhythm does not appear to influence the slope of psychometric functions, but does influence their horizontal positions. In terms of differences between subjects, it is clear that DW shows a much smaller, but consistent effect. Subject KH also shows a strong effect whereas subject AM shows the weakest effect of all subjects. Figures 7.5a-f demonstrate the same data for the tactile condition. The decreased slope of the psychometric functions is immediately noticeable suggesting that subject performance in terms of sensitivity to anisochrony is much reduced after adaptation here. The results are largely noisy across subjects and across conditions indicating the lack of a consistent effect within this sensory modality. Figures 7.6a-f present data from the same experiment replicated with the visual sensory modality. A similar effect to that seen within the auditory results (Figures 7.4a-f) is again evidenced here.

Data across all subjects show that in almost every modality, PSE values increase for conditions where subjects have adapted to a 100ms offset in rhythm. Additionally, it is clear that large levels of noise and variability presents itself across other offset patterns and particularly within the tactile modality.

Experiment 2

As single stimulus response measures rely heavily on an internal criterion that the subject sets themselves within the experiment, these measures are open to large amount of variability. To conduct a more conservative and robust method, the design was adapted to be run as a 2AFC task. In attempts to measure the underpinnings of pattern perception in the most robust way, the previous was adapted to include an additional (top-up) period of adaptation thereby presenting the subject with identical adapting sequences before each reference and test sequence during the 2AFC judgement. Results from this experiment (Experiment 2) are below.



Figures 7.7a-c (upper panel): Results for all three subjects (AM – left, DW – centre, KH – right) after adapting to auditory temporal patterns offset by 100ms in a 2AFC procedure with top-up adaptation. Figures 7.7d-f (middle panel): results for all three subjects (AM – left, DW – centre, KH – right) after adapting to tactile temporal

patterns offset by 100ms in a 2AFC procedure with top-up adaptation. Figures 7.7g-i (lower panel): results for all three subjects (AM – left, DW – centre, KH – right) after adapting to visual temporal patterns offset by 100ms in a 2AFC procedure with top-up adaptation.

As before, for all baseline conditions, the mid-point of the curve (demonstrating the probability at which the subject will accurately detect an irregular sequence on 75% of trials) is reported below. Also reported are the standard errors associated with this value. For all other conditions, 95% confidence intervals were calculated on the PSE values (point of subjective equality at which the test sequence is equally likely to be considered perfectly regular or ‘irregular’) and their respective standard errors for threshold plots from Experiment 2. The results of these are below:

Auditory

	AM	**DW**	**KH**
Baseline	25.6 ± 0.8	34.7 ± 1.3	57.7 ± 6.9
Adapt 0ms	39.8 ± 12.0	23.8 ± 4.3	40.4 ± 10.5
Adapt 100ms	39.3 ± 6.1	33.6 ± 8.6	54.5 ± 8.6

Tactile

	AM	**DW**	**KH**
Baseline	64.7 ± 11.0	55.5 ± 5.0	51.1 ± 10.5
Adapt 0ms	88.5 ± 14.3	48.1 ± 7.7	69.3 ± 45.0
Adapt 100ms	78.6 ± 11.7	59.5 ± 12.5	84.3 ± 32.6

Visual

	AM	**DW**	**KH**
Baseline	65.0 ± 5.0	65.4 ± 6.7	92.8 ± 16.5
Adapt 0ms	75.1 ± 23.5	39.6 ± 14.7	71.1 ± 25.6
Adapt 100ms	81.7 ± 15.1	49.6 ± 16.2	78.2 ± 14.5

Table 7.2. 95% Confidence interval values. Specifically, the values in the table describe the PSE for each condition for each subject, ± the confidence interval for that condition.

The data presented here (Figure 7.7) generally demonstrate the same line of results as the preceding experiment. A difference is evidenced between adapting to a perfectly regular rhythm and an anisochronous rhythm offset by 100ms (Figure 7.7a, d, f, g and h). In other words, adapting to an irregular rhythm does lead to a subjective shift in the perception of subsequent rhythms, as evidenced in PSE values. This effect however, is only partially present for each sensory modality, moreover, it is not present for all subjects. In addition, adapting to a regular rhythm (adapt 0ms condition), should, in theory, make an irregular rhythm 'feel' more irregular (as evidenced by a leftward shift of the right curve (indicating a higher PSE), however, this effect too is only found occasionally across subjects and equally occasionally across sensory modalities.

7.8 Discussion

If a temporal pattern is consciously perceived by humans as separate yet consecutive intervals, an anisochronous or patterned sequence should be processed as an average of the intervals composing the pattern. If, however, humans treat temporal sequences as a whole (rather than the sum of its parts), we should find that subjects are systematically able to adapt to anisochronous sequences (perceiving them as more regular) thereby distorting the perception of isochrony (as this should appear more irregular). We ask whether we are able to adapt to temporal patterns demarcated by alternating long and short intervals. Evidence for such processes would suggest multiple duration adaptation mechanisms working in synchrony with one another simultaneously. The experiments discussed in this chapter focus on this question and aimed to explore whether such effects can exist to modulate the perception of time.

We adapted a group of subjects to anisochronous, patterned sequences using a single stimulus (yes/no) response and also a two-alternative forced-choice design employing a period of top-up adaptation. Using the single stimulus method, we found that subjects demonstrated an after-effect of adapting to anisochronous sequences, i.e. that post-adaptation sequences that were perceived to be irregular pre-adaptation, were now thought to be regular. This method, however, is heavily dependent upon the subject's internal criterion of what constitutes regularity (and irregularity) and is therefore open to a criterion-dependent bias. Thus, whilst this method allows us to infer that exposing an subject to a period of irregular sequences shifts their internal criterion, a more robust method was needed to assess this effect objectively.

We then modified the experiment adapted to a 2AFC design using a period of top-up adaptation. Despite largely noisy results across several conditions and modalities, one pattern present in several conditions was across adapting to a perfectly regular rhythm (0ms) versus adapting to an irregular rhythm offset by 100ms. This demonstrates a subjective shift in the temporal perception and specifically that after exposure to temporal irregularity, a subject's perception of regularity becomes distorted and a much larger anisochrony is needed to elicit the same subjective sensation of anisochrony (Figures 7.7a-i).

Ultimately, we find some evidence that adapting to an anisochronous sequence does result in a subjective shift in perception and explicitly, makes a subsequently presented anisochronous sequence feel more regular. Additionally, that adapting to a pattern fails to make a regular sequence seem more irregular. Despite the weak pattern of results here, these findings may be explained through a dynamic updating of sequence processing – as explained earlier in this chapter, if temporal sequences are processed on an interval-by-interval basis, then the alternating long and short intervals should theoretically cancel out the effect of one another resulting in no – or weak observable after-effects – an outcome that may partly explain why the effects in the present experiments failed to reach statistical significance.

It is known that exposure to aperiodic/fully anisochronous sequences leads to poorer discrimination and sensitivity of subsequently presented sequences (Duarte

& Lemus, 2017; Horr & Di Luca, 2015; Ehrlé & Samson, 2005). The impact of individual differences and training on these thresholds is also noteworthy. Differences in beat detection between musical novices compared to those with musical expertise is of particular importance (Bouwer, Burgoyne, Odijk, Honing, & Grahn, 2018). Bouwer found that musical experience mediated qualitatively different processing for beats with accents imposed on them. The authors further stress the importance of designing ecological valid stimuli when assessing beat and rhythm perception as some individuals may require stronger accents on the beats presented to them to elicit the same percept of a beat that would be experienced by an individual with more musical expertise. Since these results suggest not all beats are processed uniformly across individuals, it further stresses the importance of carefully constructing rhythms in future experiments (Bouwer et al., 2018).

Interestingly, it has been found that the relationship between sensory jitter and detection thresholds suggests a human inability to detect perturbations in a sequence smaller than those possible for the motor production of such a sequence (Madison & Merker, 2004). Despite the present experiments employing a range of offsets in anisochronous sequences, future work may perhaps equate sensitivity across the sensory modalities and adapt subjects in accordance with their specific thresholds. In efforts to maintain consistency across experimentation and subjects, the present experiments employed identical stimuli features and testing conditions but equating sensitivity before testing subjects in line with their individual thresholds may perhaps be a more robust method to assess temporal pattern perception.

Here, we find that rhythm perception is partly influenced by the regularity of the intervals that compose the rhythm. Moreover, that regular rhythms invoke a much stronger sense of time than their irregular counterparts. The results presented here can be explained through a building up of the average (akin to serial dependencies), as alternating short and long intervals are being averaged, adaptation is failing to occur for a number of conditions. It is plausible that humans cannot adapt to more than one duration simultaneously, and the lack of duration information (as in our anisochronous conditions) prohibits a build-up of any resulting after-effect. This hypothesis however, does support the model of duration channels to process time

(Chapter 2), and will need to be explored further in future work. Future work may also extend these findings to clinical populations who have been documented to be impaired in beat-based rhythm discrimination (Grahn & Brett, 2009). Moreover, future studies could also aim to explore anisochrony within a temporal pattern, and more complicated temporal patterns, as these could provide insight into the perception of “expressive timing in music” which often uses very subtle temporal deviations in otherwise, regular temporal structures (Schulze, 1989).

8. Adapting to cross-modal rhythms in search of a centralised timer

When exploring sensory time perception, the sensory modes most commonly studied (audition, touch, vision) are often organised and studied separately. A growing body of evidence suggests substantial similarities between the modes of audition and touch (Soto-Faraco & Deco, 2009) and other lines of evidence also shed light on the unique nature of audio-tactile temporal perception compared to other sensory pairings (Occelli, Spence, & Zampini, 2011). To date, much of the work attempting to differentiate between centralised versus distributed theories of processing has focused on separate unimodal assessments of timing mechanisms. To truly gauge and activate the presence of centralised timing mechanisms however, cross-modal stimuli should be employed. Thus, an alternative method to explore the cross-modal nature of time is to assess the subjective perception of intervals defined by two different sensory modalities.

8.1. Internal representation of cross-modal markers

A contentious question is how cross-modal markers are represented subjectively. One postulation is the internal marker hypothesis (Grondin, 1993) suggesting that the window of internal representation is activated by the response function to the first signal and then to the response function generated by the second signal. This hypothesis suggests that the natural differences in response times between tactile and visual modalities is what may lead to biases in temporal estimation (Grondin, 1993). Statistical evidence for this hypothesis however, remains sparse. Empty interval discrimination has been found to be affected by both the signal (marker) length and also the sensory modality composing the signal (Grondin, Roussel, Gamache, Roy, & Ouellet, 2005). It has been documented that shorter intervals are more accurately discriminated in the auditory modality (Welch & Warren, 1980). Additionally, it has been suggested that audio-visual intervals are perceived to last longer than visuo-auditory intervals of the same duration (Grondin & Rousseau, 1991). As shorter durations are understood to be processed by automatic mechanisms, whereas longer durations are thought to be more cognitively mediated (Lewis & Miall, 2006; Lewis & Miall, 2003), differences between auditory and visual modes in the processing of sub and supra-second durations may actually arise from threshold differences for automatic and cognitive processing.

In an assessment of audio, tactile and visual empty duration judgements (Grondin & Rousseau, 1991), it was reported that the perceived duration of an interval depends on the sensory signals defining it but also that the level of prior certainty regarding the interval also affects judgements. Specifically, higher certainty regarding the sensory modalities defining an interval improves judgement. The critical finding has been that intramodal empty intervals (those defined by the same modality) were easier to discriminate than intermodal empty intervals (those defined by two separate modalities) (Grondin & Rousseau, 1991). The authors go on to suggest the presence of two distinct processing systems mediating duration discrimination; one with processing capacities that are modality-specific and the other that is modality-independent and responsible for intermodal interval discrimination (Grondin & Rousseau, 1991).

8.2. Influence of sensory modalities in empty interval duration discrimination

In a further investigation of marker influence (composing intermodal intervals) on duration discrimination, it was again reported that the modalities of markers defining an interval do indeed affect performance on duration discrimination tasks (Grondin, Ivry, Franz, Perreault, & Metthe, 1996). Using a range of durations employing both VA and AV intervals (‘A’ denotes auditory; ’V’ denotes visual), it was found that while the intensity of markers failed to influence duration judgements, the marker modality did influence duration discrimination and so did the length of the first and second markers. Specifically, for certain durations, VA intervals were better discriminated than AV intervals however, for durations ranging from 250-750ms, AV intervals were perceived as longer than VA intervals, a result also supported by Mayer et al., (2014). These results are unaffected by the intensity of the visual signal; and finally, that the length of the first and second markers both influence the perceived duration of the intermodal interval (Grondin et al., 1996). The finding that VA intervals are easier to discriminate than AV intervals (and that discrimination was impaired cross-modally (but not unimodally), again importantly demonstrates the presence of independent mechanisms governing timing within these modalities, as there is no reason why a centralised mechanism would allow such discrepancies between two different sensory modalities demarcating the same interval. A contentious explanation for this may be that auditory intervals are generally perceived as longer than visual intervals despite having the same physical duration (Wearden, Edwards, Fakhri, & Percival, 1998; Shi & Burr, 2016), thus, AV intervals being reported as longer than VA intervals may be explained by the preceding influence of the first auditory signal. Further evidence however, is needed to validate this claim. In a more intricate design, Kuroda and colleagues (Kuroda, Hasuo, Labonte, Laflamme, & Grondin, 2014) assessed discrimination of intra and inter-modal empty intervals marked by three successive stimuli. The possible intervals were either AVA, VAV, AAA or VVV; thus, the first and last signal were always from the same modality and the first interval remained fixed at 500ms, subjects then judged whether the second interval was shorter or longer than the first in a two-alternative forced-choice design. In a comparison of intra and inter-modal intervals, intermodal discrimination was impaired compared to intramodal discrimination, but critically, this effect was amplified for the VAV condition (compared to the AVA condition), despite both stimuli consisting of the same intervals. In other words, performance was markedly impaired in the VAV condition but not the AVA condition compared to AAA or VVV (Kuroda et al., 2014). This may again be explained with the processing of the first interval, which, if auditory, then drives the temporal processing of the remaining stimulus.

8.3. Comparing unimodal versus cross-modal intervals

To comprehensively compare how amodal our sense of interval time truly is, Grondin and McAuley (2009) ran four duration discrimination experiments across a range of unimodal and cross-modal conditions. All four experiments employed a two-alternative forced-choice procedure where subjects were presented with two sequences and asked to report whether the second sequence was shorter or longer than the first. The sequences ranged in the number of intervals they presented (ranging from one to four intervals) and in the sensory modalities presented (either auditory or visual). Markers in the first experiment were tones for the first sequence and flashes for the second sequence; experiment two ran the same modalities but in reverse (flashes for the first sequence and tones for the second). Experiment three present flashes for both sequences and experiment four presented tones for both sequences. In AV, VA and VV conditions, duration discrimination was improved by increasing the length of the sequence, providing further support for the ‘Multiple Look Model’. More generally, performance was most accurate for AA sequences, average for AV and VA sequences and considerably worse for VV sequences. Interestingly, performance was also best when a fixed standard interval was presented in the first sequence compared to presenting a variable interval (Grondin & McAuley, 2009). The authors assert that our perception of time is amodal in the sense that it can be conveyed through multiple sensory modes. Their results also support previous findings of temporal perception is indeed markedly better in the auditory than visual or tactile modalities and is significantly reduced when two different modalities demarcate a single empty interval (Grondin, 1993). It has also been noted that repetitions of an interval lead to better performance compared to subjects only being presented with one interval (Michon, 1964; Drake & Botte, 1993).

8.4. Neurophysiological investigations of cross-modal interval perception

Investigations using psychophysics paired with electroencephalography have also been conducted to investigate the mechanisms modulating the processing of unimodal and cross-modal intervals. Predictably, intervals marked purely by the auditory modality are better discriminated than the same intervals in the visual modality when the interval was demarcated unimodally rather than cross-modally (Gontier, Hasuo, Mitsudo, & Grondin, 2013). In Gontier et al.'s experiments, subjects were presented with short and long empty unimodal and cross-modal intervals demarcated by either the auditory or visual modality. Performance as indicated by behavioural data was improved for the unimodal auditory intervals compared to VV, AV or VA intervals. These results were corroborated by EEG data indicating a significant increase in the Contingent Negative Variation (CNV) amplitude recorded at fronto-central electrodes during the long unimodal auditory conditions, whereas no such changes were observed in the time course for this component during VV, AV or VA presentations. The time course for the CNV revealed that marked improvements with unimodal auditory intervals may be explained by higher effectiveness in the neural mechanisms underlying the temporal processing of ongoing intervals, a result that has previously been noted for auditory rhythms also (McAuley & Henry, 2010). It was also found that N1 and P2 amplitudes were higher for cross-modal (VA/AV) compared to unimodal intervals. The authors explain this through an attentional bias linked to the cognitive load associated with switching between sensory modalities (Gontier et al., 2013). This assertion has also been made by others who suggest that an increased attentional load also compromises temporal prediction abilities (Baker, Dexter, Hardwicke, Goldstone, & Kourtzi, 2014).

Differences in oscillatory mechanisms of auditory and visual stimulus processing have also been noted (Zoefel & VanRullen, 2017). Specifically, that the rhythmic component in signals is not crucial for the visual, but is crucial for the auditory system. Exposure to unpredictable stimuli also confer differences in oscillatory mechanisms (Zoefel & VanRullen, 2017). Conflicting evidence from psychophysical assessments of vibro-tactile and visual asynchronies however demonstrates consistency in Weber fractions across comparisons of unimodal and cross-modal intervals. This has been used to suggest that contrary to what has been suggested, cross-modal comparisons do not elicit higher levels of noise in judgement tasks (van Erp & Werkhoven, 2004).

8.5. Influences of training on cross-modal duration discrimination

In efforts to gauge the cross-modal dependence between senses in duration discrimination, Grondin and Ulrich employed a training period of discriminating brief auditory intervals and later tested the possibility of improved performance in brief visual intervals (Grondin & Ulrich, 2011). Two separate groups completed pre-test visual interval discrimination. One group was then subjected to auditory interval discrimination (experimental group), whereas the second had no such auditory training (control group). Both groups then completed the post-test visual interval discrimination task. More notably, it was found that whilst duration discrimination of visual intervals was improved in the post-testing phase, this effect was not statistically attributable to the auditory training. In other words, both groups demonstrated improvements in visual duration discrimination when tested post-training. It can be speculated that this is a result of familiarity with the task. The authors suggest that despite substantial perceptual learning, a failure of cross-modal transfer was evidenced (Grondin & Ulrich, 2011) providing further support for the distributed theory of timing (Becker & Rasmussen, 2007; Motala et al., 2018).

In four separate tasks investigating unimodal and cross-modal spatial attention and rhythm induced expectation (Jones, 2015), Jones (albeit predictably) found that responses were faster for stimuli that were spatially attended to compared to stimuli that were unattended. It was also found that targets presented in synchrony with a rhythm were responded to faster than stimuli that were presented prematurely in cross-modal tasks. Interestingly, however, they found that rhythmic stimuli in one modality influenced temporal expectancy in another modality; suggesting that rhythmically-induced temporal expectation is centralised and modality-independent.

Furthermore, spatial attention and rhythmic cueing were found to be largely dissociated from one another and did not provide evidence of an interaction – proposing that their influence on behaviour is independent from one another (Jones, 2015). More recent evidence on the discrimination of cross-modal rate has suggested optimal integration of random visual and auditory signals irrespective of their unisensory temporal correlations (Locke & Landy, 2017). Conflictingly, however, the authors suggest that optimal integration occurs when cross-modal stimuli are spatially coupled (Locke & Landy, 2017).

It is clear that whilst considerable effort has been made to elucidate the processing of cross-modal rate, more conclusive evidence is needed to determine the extent to which sensory systems are able to combine concurrently presented unimodal rhythms to form a single rhythm. To more directly assess the presence of a centralised timing mechanism, the following experiments adapted subjects to an interleaved and co-localised cross-modal rhythm (audio-visual) at a 3Hz frequency to investigate whether such a rhythm comprising two different sensory rhythms could be integrated and combined to form a single rhythmic percept. The results from this design can support one of two possibilities; the first, that subjects are only able to adapt to the two unimodal streams thereby producing unimodal after-effects where a unimodal test interval is underestimated whereas cross-modal intervals are either reproduced veridically, or slightly overestimated too. The second possibility is that subjects are able to integrate the two sensory streams as one rhythmic percept and therefore, any resulting after-effect will be cross-modal in nature and will cause cross-modal test stimuli to be overestimated and unimodal test stimuli to either be reproduced as veridical or slightly overestimated.

8.6. Methods

Subjects:

3 participants (2 female and 1 male, mean age = 36, standard deviation = 14 years) participated, with self-reported normal hearing and visual abilities. Following initial

practice sessions, a lengthy process of data collection began, in a series of sessions spread over several weeks. Two of the participants had previous experience of psychophysical data collection. The third participant (KB) had no such experience and was naïve to the purpose of the experiments. The experiments received ethical approval from the Research Ethics Committee at the School of Optometry and Vision Sciences, University of Cardiff and all experiments were performed in accordance with relevant guidelines and regulations. Informed consent was obtained for study participation.

Stimuli:

Brief (16msec duration) sensory stimuli were presented – either in the auditory or visual modality and all stimuli were grossly suprathreshold. Stimulus generation and presentation was controlled by an Intel ® Core ™ i5-4460 desktop computer running Microsoft Windows 7. The programming environment involved MATLAB 8.6 (Mathworks, USA) in combination with Psychophysics Toolbox 3 (http://www.psychtoolbox.org). Stimulus timing was verified using a dual-channel oscilloscope.

Visual:

Visual signals were bright flashes presented centrally against a uniform dark background (0.32cd/m2). Stimulus duration was a single frame (approximately 16ms at the monitor frame rate of 60Hz) and the viewing distance was kept constant at 60cm. Visual stimuli in the co-localised conditions were presented via a 20mm LED bulb presenting a highly luminous green/yellow flash at 568nm. These were bright flashes presented centrally against a uniform dark background. A single flash lasted 16ms. At the viewing distance of 60cm the circular flash subtended a diameter of approximately 2° of visual angle.

Auditory:

Auditory stimuli consisted of brief (16ms duration) bursts of white noise generated by a Xonar Essence STX (ASUS) soundcard (https://www.asus.com/us/Sound-Cards/Xonar_Essence_STX/) with a sampling rate of 44,100Hz. Stimuli were delivered using a loudspeaker. Co-localisation was achieved by attaching an LED to present visual stimuli to the speaker, and placed centrally in front of the participant. In non-co-localised conditions, auditory stimuli were presented over headphones as in Chapters 5, 6 and 7.

Experiment 1 – Does unimodal rhythm adaptation elicit distortions in cross-modal intervals?

The key question here is whether distortions as a result of rhythm adaptation can be demonstrated in cross-modal test intervals. The aim of these experiments was therefore to extend the result from Chapter 5 that rhythm after-effects can be evidenced within single cross-modal interval comparisons as demonstrated with the interval reproduction method.

To establish typical rebound after-effects to unimodal rhythms, subjects were adapted to 1.5Hz and 3Hz unimodal rhythms and tested with 333ms and 667ms empty cross-modal intervals (AA, VV, AV and VA) respectively. To clarify, they adapted to 8-10 seconds of a unimodal rhythm, followed by a pre-test attentional signal to limit perceptual grouping, and then were presented with the test interval to be reproduced via the response disk. The attentional signal would be of the same modality but altered slightly, for example, for the auditory condition, the signal was a ‘beep’ of a higher frequency than those of the adapting sequence. Similarly, in the visual condition, subjects adapted to a sequence of white flashes on the screen and the attentional signal was a blue flash to indicate the beginning of the test phase.

Subjects were required to reproduce the empty test interval by tapping with their right forefinger on a piezoelectric transducer (https://www.amazon.co.uk/Piezo-electric-disk-transducer-15mm/dp/B01K8X9E5K) to mark the beginning and end of the interval. The resulting voltage output was fed to the ‘audio in’ of the soundcard as a recording which was analysed within MATLAB to extract the duration reproduced. The transducer was enclosed in a sound-dampening environment and shielded from sight of the subject.

We also ran a series of pilot experiments on subject AM before concluding which experiments to proceed with. The pilot conditions are outlined below:

- Adapting to 1.5Hz unimodal (auditory and visual) rhythms and testing with 333ms and 667ms intervals.
- Adapting to 3Hz unimodal (auditory and visual) rhythms and testing with 333ms and 667ms intervals.
- Adapting to 3Hz cross-modal (AV) rhythms and testing with 333ms and 667ms intervals.

Experiment 2 – Can we construct a multisensory rhythm?

The aim of this experiment was to establish whether a perfectly interleaved audio-visual sequence can be perceived as a single rhythm.

Subjects were presented with an audio-visual rhythm presented at 3Hz. This rhythm was perfectly interleaved by 333ms such that the unimodal signals were presented at 1.5Hz but when combined, the cross-modal rhythm was presented at 3Hz (Figure 7.1). After 10 seconds of the adapting sequence, a short pause ensued, followed by a 500ms empty test interval that could be one of four possible sensory combinations. The test signal duration was fixed at 500ms, however each signal defining the beginning and end of the interval could be demarcated by either the visual or auditory modality. This meant that the adapting sequence was followed by an empty interval that was either audio-audio, visual-visual, audio-visual or visual-audio. The task then required subjects to reproduce the interval by tapping on a response disk.

Data was gathered from 25 trials for each condition in each experiment and across each subject.

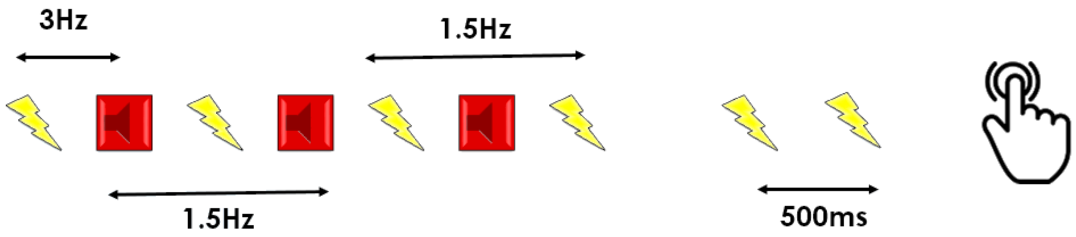


Figure 8.1. Schematic of the second (interval reproduction) task. The participant would adapt to 10 seconds of the AV rhythm. The test interval here is a unimodal empty visual interval lasting 500ms that the participant is required to reproduce. Response phase drawing accessed from www.iconsmind.com.

Experiment 3 – Can a co-localised bi-modal rhythm be adapted to as a single rhythm?

As literature exists indicating the importance of co-localising auditory and visual streams when assessing central timing mechanisms (Levitan et al., 2015), we also felt it was necessary present co-localised sensory streams and assess whether subjects were then able to combine the two streams into one and perceive a single rhythm. This was done by presenting auditory signals via a TEAC two-way speaker system with a superimposed LED presenting visual signals, both placed in the same testing booth as previous experiments.

The second cross-modal adaptation experiment exposed participants to auditory and visual sensory co-localised streams in order to present signals from the same spatial location. The subject then reproduced the duration of the presented interval by tapping on the response disk twice to indicate the beginning and end of the interval.

Data was gathered from 25 trials for each condition in each experiment and across each subject.

The purpose of this experiment was to assess to what extent participants would be able to adapt to a co-localised 3Hz rhythm demarcated by signals arriving from two different sensory modalities. By assessing reproduction values of a range of unimodal and cross-modal intervals, we may be able to deduce the strength (if any), of integrating these two signals to form a single rhythmic percept.

Experiment 4 – Can a unimodal rhythm perceptually match a bi-modal rhythm?

The purpose of this experiment was to allow participants to more strictly match a cross-modal rhythm to its unimodal equivalent. If cross-modal signals can be combined to form a single rhythm, we would expect participants to match this rhythm to a unimodal rhythm closer to 3 rather than 1.5Hz. On the contrary, if subjects match the cross-modal rhythm closer to a 1.5Hz unimodal frequency, this would suggest lower levels of integration of the two sensory streams in forming a single rhythmic percept.

We aimed to establish the unimodal equivalent rhythm of a cross-modal 3Hz rhythm. This was done by employing a two-alternative forced-choice design. Specifically, the subject was presented with a co-localised 3Hz AV rhythm for 10 seconds, followed by a unimodal rhythm (either auditory or visual) that ranged from 1-3.5Hz and that was also presented for 10 seconds, the subject was then instructed to respond to which sequence they perceived as faster, the first or the second via a keypress.

Data was gathered from 25 trials for each condition in each experiment and across each subject. Preliminary experiments (Experiment 1) were blocked according to sensory modality and condition. The order of blocks was randomised. For the main experiments (2 and 3) where subjects were adapting to unimodal and cross-modal streams simultaneously, the presentation of test stimuli was randomised. This was kept consistent throughout the main experiment and across subjects. For the final control experiment (Experiment 4), a random block design was used where the conditions presented within a block were randomised but separate blocks were formed by different sensory modalities (leaving two blocks – one visual and one auditory).

Other considerations

Careful considerations had to be made to elicit a general sense of rhythm across the two modalities. The time needed to deduce the fact that two successive stimuli were successive as opposed to simultaneous is dependent upon the modality that the stimuli are presented in. The time needed to deduce which of two events preceded the other is roughly 20ms and is stable across all modalities (Hirsh & Sherrick, 1961).

Furthermore, comprehensively perceiving stimuli at higher frequencies becomes harder the higher the temporal frequency of stimulus presentation; whereas at lower rates it almost becomes too easy to recognise components as unintegrated elements (Garner & Gottwald, 1968). During pilot testing it was found that a cross-modal frequency any faster than 3Hz would result in both sensory streams appearing to present simultaneously. This result has been noted previously (Zampini, Guest, Shore, & Spence, 2005) where sequences of brief clicks and flashes with disparities of up to 100ms (10Hz) were perceived as simultaneous (Zampini et al., 2005). For these reasons, we decided to use a 3Hz cross-modal rhythm, as any frequency below this was too slow and failed to elicit a sense of rhythm across two modalities whereas any sequence faster than 3Hz resulted in the audio and visual signals appear synchronous and again, failed to elicit the perception of a single rhythm presented through two different sensory modalities (Figure 8.1).

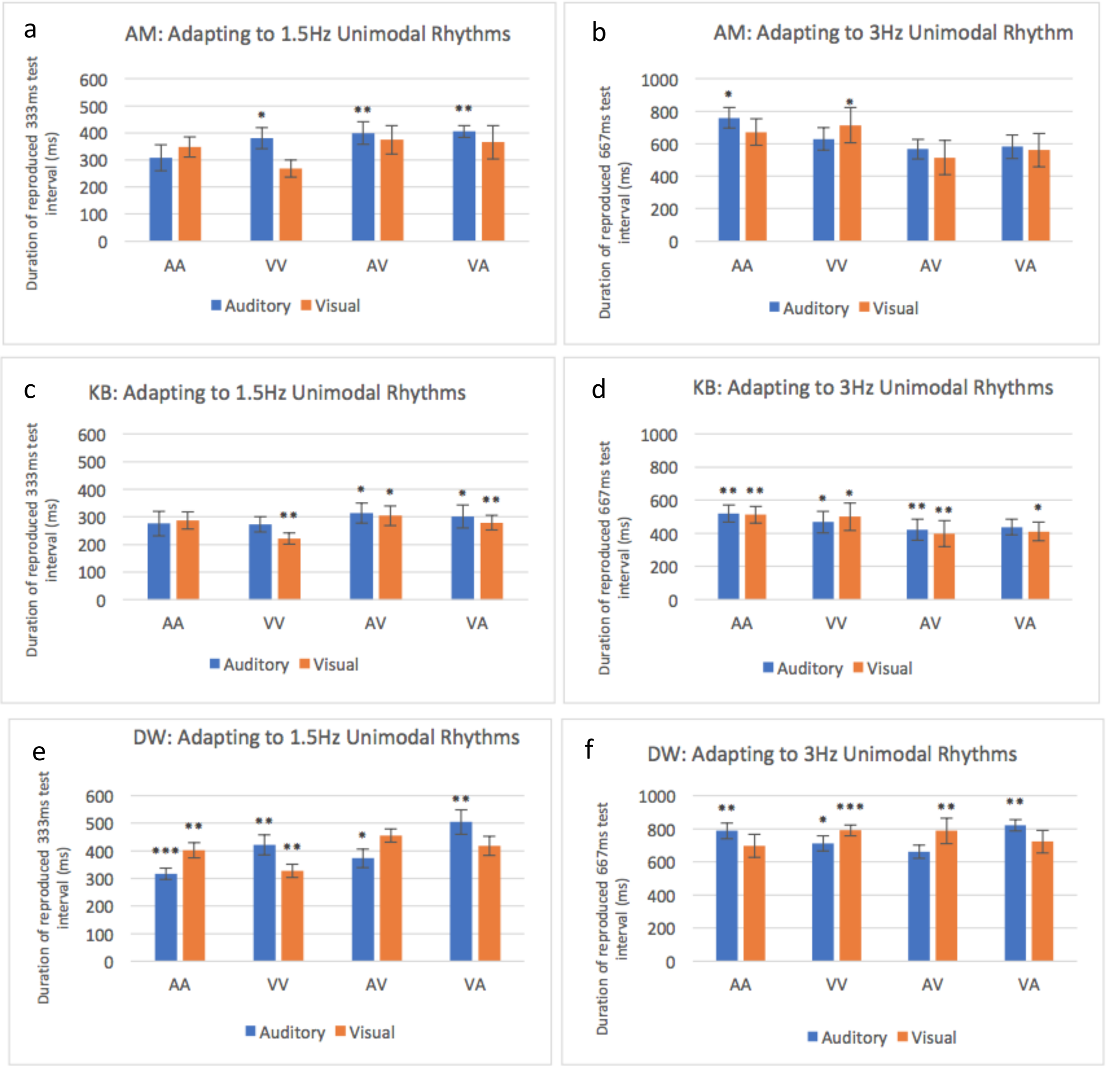
8.7. Results

There are two clear alternative outcomes for these experiments. The first is that subjects will combine the interleaved auditory and visual rhythms as one, resulting in cross-modal after-effects (such that AV and VA intervals will be overestimated). The alternative is that subjects will not be able to integrate the two different rhythms as one and will instead only present after-effects to unimodal AA and VV intervals (and will under produce these).

Experiment 1 – Does unimodal rhythm adaptation elicit distortions in cross-modal intervals?

The aim of these experiments was to establish whether adapting to a unimodal rhythm is able to elicit distortions in a cross-modal empty interval. Data presented below is from all three subjects.

Statistical significance for each subject was assessed on the mean reproduced values for the test interval between pre- and post-adaptation conditions. Using SPSS (IBM Corp. Released 2015. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.), paired-samples t-tests were conducted on reproduction data from each subject and each condition to test the hypothesis that adapting to a unimodal rhythm would still elicit after-effects with unimodal and cross-modal empty test intervals. Results of these tests are below (Figure 8.2).



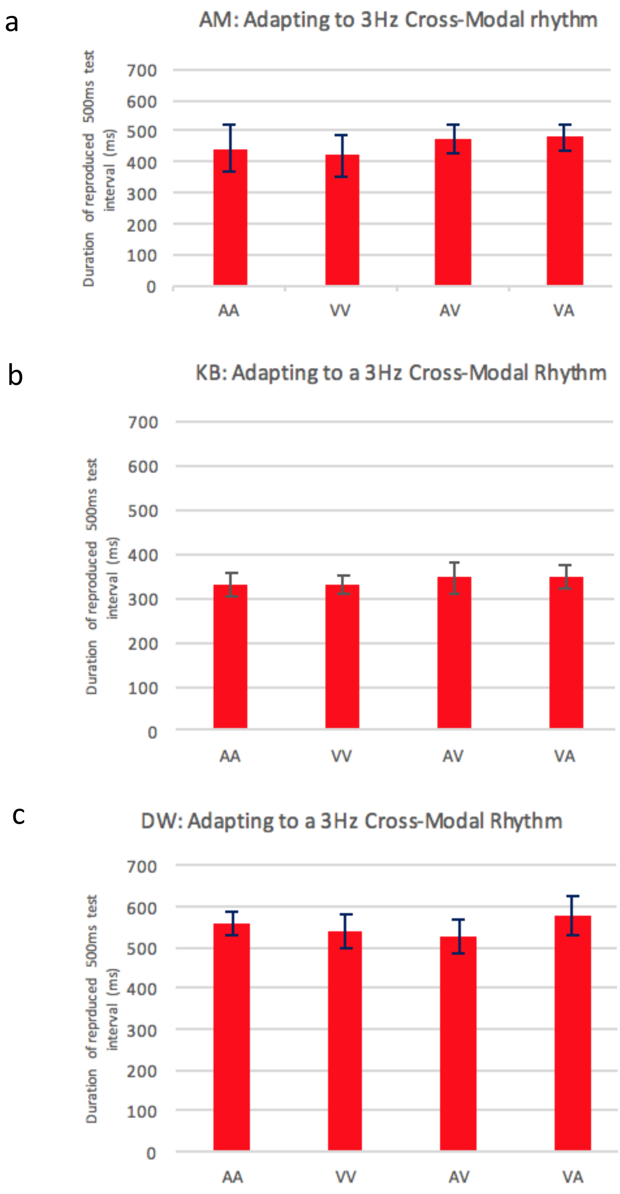
Figures 8.2a-f. Left panel represents the effects of adapting to a 1.5Hz unimodal rhythm (where the blue bars indicate auditory rhythm after-effects and orange bars represent visual rhythm after-effects) through interval reproduction of a 333ms interval. Right panel represents the after-effect of adapting to a 3Hz unimodal rhythm (where the blue bars indicate auditory rhythm after-effects and orange bars represent visual rhythm after-effects) through interval reproduction of a 667ms interval. The y axis demonstrates the range of values reproduced for the test interval whereas the x axis outlines the different intervals that were reproduced (either AA, VV, AV or VA). Three different subjects are plotted at the top, middle and bottom of the panel, corresponding to the previous presentation order. Asterisk above each bar represents statistically significant after-effects, *represents $p<0.05$, **represents $p<0.01$ and ***represents $p<0.001$.

Typically, participants are exposed to rhythm adaptation and are then also tested with similar rhythms in a variety of tasks. We show that typical rebound duration after-effects are demonstrated post rhythm adaptation with empty interval comparisons (Figures 8.2a, 8.2b, 8.2d, 8.2e and 8.2f). These effects are demonstrated as a consequence of adapting to various unimodal rhythms in the auditory and visual modalities, across several subjects and adapting temporal frequency. As these effects are only observed loosely, it is clear that individual differences posed a modulating influence on responses – for instance, subject AM only demonstrated an after-effect on 31% of conditions, KB on 75% of conditions and DW on 69% of conditions.

Despite these data loosely replicating results evidenced earlier (Chapter 7), the extent of influence of the modality of the first signal demarcating the test interval still remains to be explored. Here, we observe that the sensory modality of the first signal fails to define the pattern of resulting after-effects when the subject is exposed to a cross-modal interval. This is evidenced in no single individual demonstrating an after-effect to a specific cross-modal pairing. Individual differences are again apparent here as for subject KB rebound aftereffects are observed for 75% of cross-modal conditions, for 50% of cross-modal tests for subject DW. Contrastingly, this effect is only observed for 25% of cross-modal test intervals for subject AM. One plausible explanation for this may be that when presented with cross-modal intervals, subjects were unable to build a stable temporal representation as the test interval was transient. To explore this further, the next experiment will expose subjects to a sustained cross-modal interval presented in the form of a bi-modal rhythm.

Experiment 2 – Can we construct a multisensory rhythm?

The primary aim of these experiments was to establish whether a perfectly interleaved cross-modal rhythm can be perceived as a single unitary rhythmic percept.



Figures 8.3a-c. The after-effect of adapting to a 3Hz audio-visual rhythm demonstrated through interval reproduction of a 500ms empty interval demarcated by either auditory or visual signals. The y axis demonstrates the range of values reproduced for the 500ms test interval whereas the x axis outlines the different intervals that were reproduced (either AA, VV, AV or VA). Three different subjects are plotted at the top, middle and bottom of the panel, corresponding to the previous presentation order.

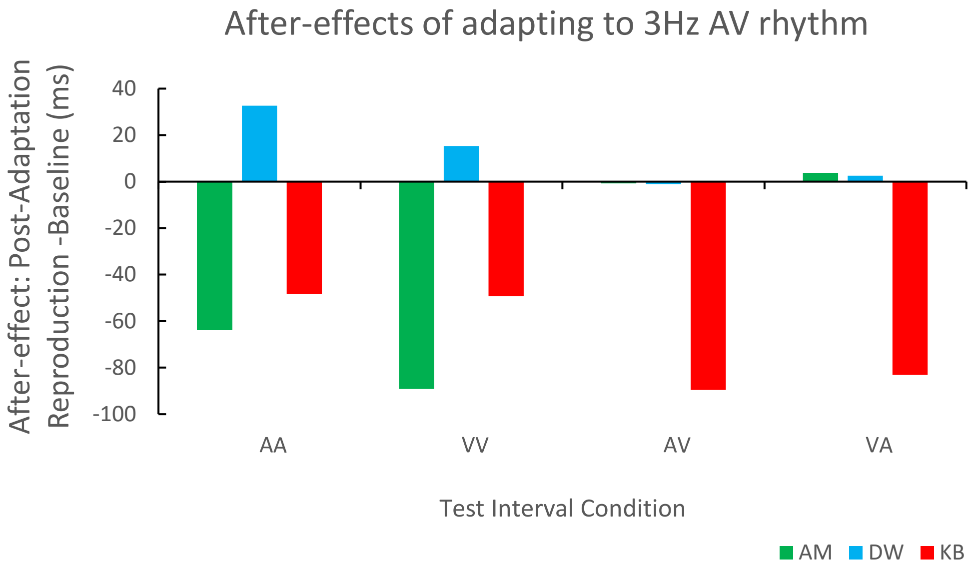


Figure 8.4. Comparisons of test interval reproduction for baseline and post-adaptation conditions. The bars demonstrate values of post-adaptation minus baseline values for each condition. The right legend indicates labelling for each subject.

The downwards extending bars in Figure 8.4 demonstrate conditions where the test interval was under produced as a result of adaptation (as the calculation is post-adaptation reproduction minus baseline reproduction). Whereas the upward extending arrows demonstrate the opposite effect – the test interval being overproduced as a result of adaptation. Results from paired-samples t-tests comparing the baseline (no adaptation) reproduction of a unimodal test interval to the same interval reproduced after a period of adaptation, are shown below (Table 8.1).

Condition	**Participant AM**	**Participant KB**	**Participant DW**
AA	.038*	.183	.191
VV	.031*	.136	.147
AV	.979	.052	.942

VA	.867	.090	.852

Table 8.1. Paired-samples t-tests comparing baseline versus post-adaptation of a 500ms empty test interval. * denotes conditions significant at p = .05.

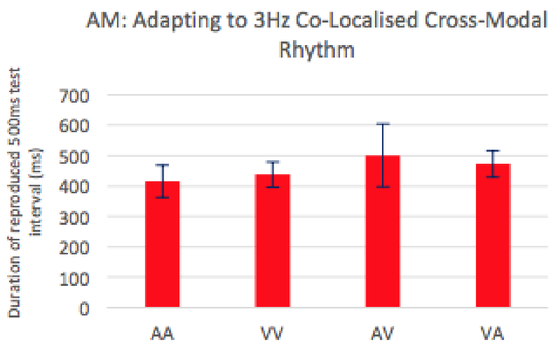
Subjects (AM and KB) loosely show the same pattern of results as the typical adaptation experiments when reproducing unimodal test intervals. It becomes clear however, that failing to co-localise stimuli results in a decrease in the amplitude of any resulting after-effect (Figure 8.4). A two-factor ANOVA (with replication) was then conducted on reproduced test intervals at a first level to compare a statistically significant difference between unimodal and cross-modal intervals, and then on whether the first signal of the test stimuli was significant. Subject AM demonstrated a significant difference in test interval reproduction between unimodal and cross-modal intervals (F(1,16) = 8.55, *p* = 0.01) and no significant difference in comparisons of whether the first modality signifying the test interval influenced the aforementioned result (F(1,16) = 0.55, *p* = 0.47). Subject KB demonstrated no significant difference between either unimodal or cross-modal intervals, nor between whether the first signal demarcating the test interval was unimodal or cross-modal (F(1,16) = 1.37, *p* = 0.26) and (F(1,16) = <0.01, *p* = 0.96) respectively. Subject DW demonstrated a similar pattern of results and showed no significant difference between either unimodal or cross-modal intervals or the sensory modality demarcating the first signal of that interval (F(1,16) = 0.06, *p* = 0.81) and (F(1,16) = 1.24, *p* = 0.28) respectively.

Experiment 3 – Can a co-localised bi-modal rhythm be adapted to as a single rhythm?

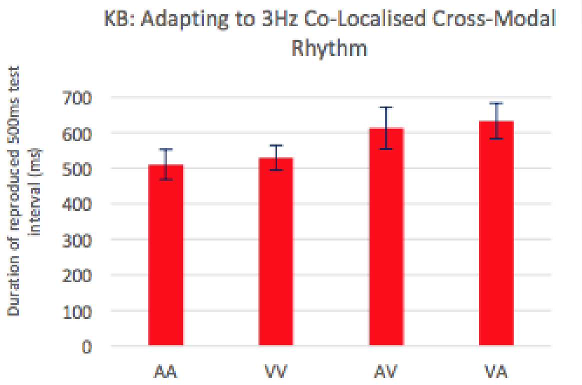
As there exists evidence to suggest that the processing of sensory time is spatially-specific, Experiment 2 was repeated using the same sensory pairing however in these set of experiments, the auditory and visual sensory streams were co-localised in their spatial locations.

Subjects exposed to a 3Hz co-localised rhythm were presented with unimodal intervals occurring every 666ms (resulting in 1.5Hz unimodal adaptation) and simultaneously to a cross-modal interval every 333ms (resulting in 3Hz cross-modal adaptation). Our hypothesis therefore suggests that a 500ms test interval should be compressed for unimodal compared to cross-modal intervals in the response phase.

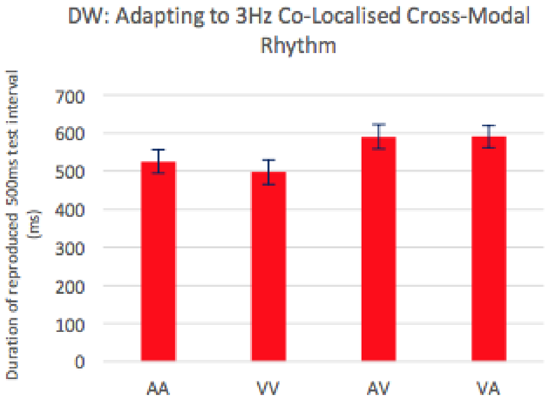
a



b



c



Figures 8.5a-c. The after-effect of adapting to a co-localised 3Hz audio-visual rhythm demonstrated through interval reproduction of a 500ms empty interval demarcated by either auditory or visual signals. The y axis demonstrates the range of values reproduced for the 500ms test interval whereas the x axis outlines the different intervals that were reproduced (either AA, VV, AV or VA). Three different subjects are plotted at the top, middle and bottom of the panel, corresponding to the previous presentation order.

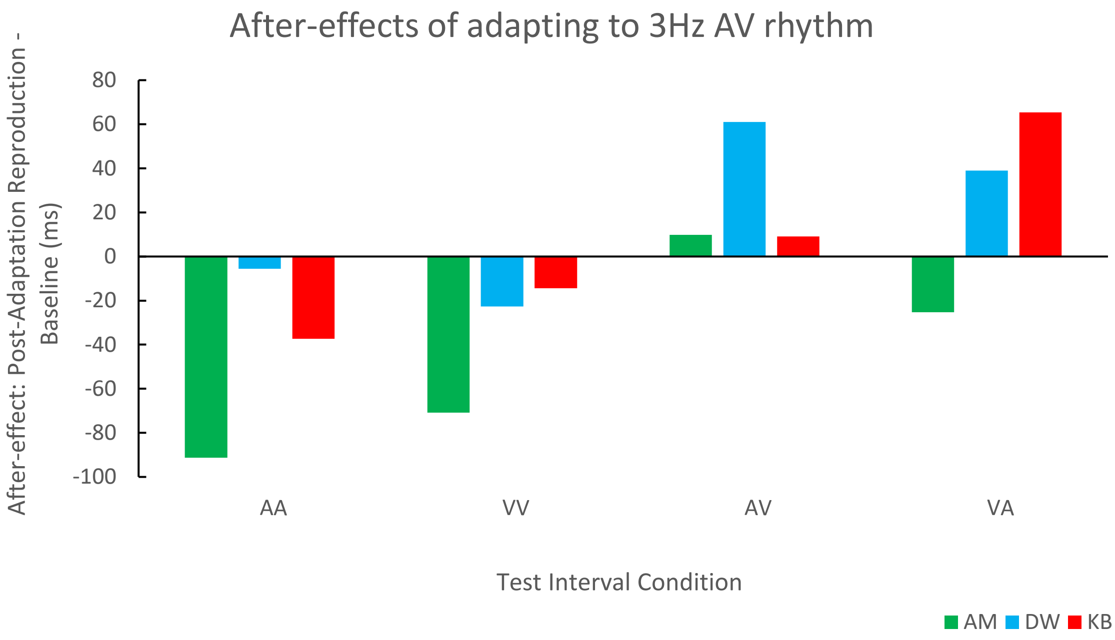


Figure 8.6. Comparisons of test interval reproduction for spatially co-localised baseline and post-adaptation conditions. The bars demonstrate values of post-adaptation minus baseline values for each condition. The right legend indicates labelling for each subject.

As for Figure 8.4, downward bars indicate an underproduction of the test interval whereas upward bars indicate an overproduction of the test intervals. The importance of co-localising auditory and visual streams becomes immediately apparent as the results demonstrate a more consistent pattern across subjects.

Objectively, participants generally underestimate unimodal intervals compared to their cross-modal counterparts, after adapting to a cross-modal rhythm, suggesting evidence for unimodal behavioural after-effects. Compellingly, they also show an overestimation of cross-modal signals suggesting cross-modal after-effects to adapting to regular intervals separated by gaps of 333ms. Data from subject AM (Figure 8.5a) suggests that compression of unimodal intervals post cross-modal adaptation is a result of unimodal stimuli being perceived as shorter. Contrastingly, the same data for subject DW (Figure 8.5c) suggests that the underestimation of unimodal intervals is a result of cross-modal intervals feeling longer. Results from paired-samples t-tests comparing the baseline (no adaptation) reproduction of a unimodal test interval to the same interval reproduced after a period of adaptation, are shown below (Table 8.2).

Condition	**Participant AM**	**Participant KB**	**Participant DW**
AA	.009*	.063	.224
VV	.025*	.542	.038*
AV	.545	.741	.006*
VA	.429	.055	.002*

Table 8.2. Paired-samples t-tests comparing baseline versus post-adaptation of a 500ms empty test interval. * denotes conditions significant at $p = .05$.

As with experiment 2, a two-factor ANOVA (with replication) was then conducted on data gathered by each participant. Subject AM demonstrated a significant difference between unimodal and cross-modal interval reproduction ($F(1,16) = 18.20$, $p = <0.01$) but this effect was not a result of the modality marking the first signal of the test interval ($F(1,16) = 0.07$, $p = 0.80$). Subject KB showed a similar result with a significant difference between unimodal and cross-modal interval reproduction ($F(1,16) = 45.13$, $p = <0.01$), but again, no significant difference between whether the first signal demarcating the test interval was auditory or visual (F(1,16) = 1.69, *p* = 0.21). Subject DW also demonstrated a significant difference between unimodal and cross-modal conditions (F(1,16) = 123.02, *p* = <0.01) and, also demonstrated no significant difference between the sensory modalities signifying the first marker of the test interval (F(1,16) = 3.58, *p* = 0.08).

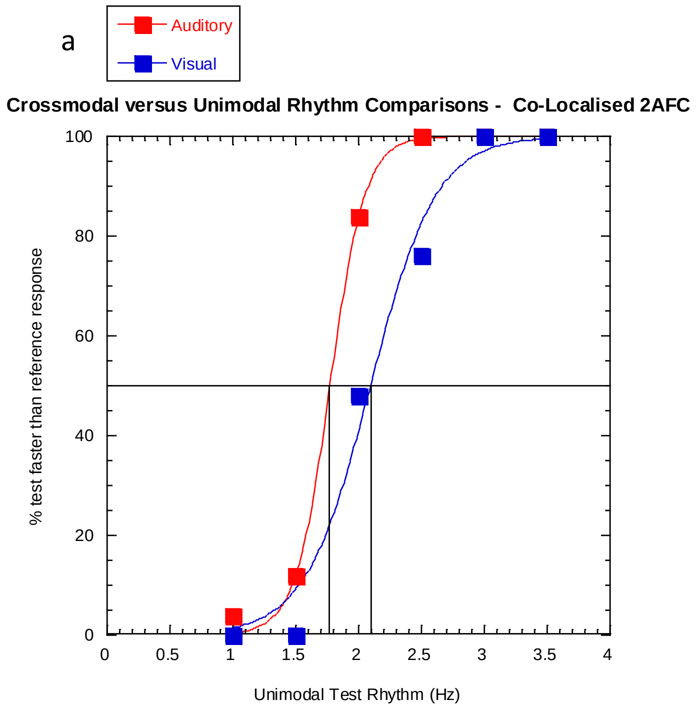
Experiment 4 – Can a unimodal rhythm perceptually match a bi-modal rhythm?

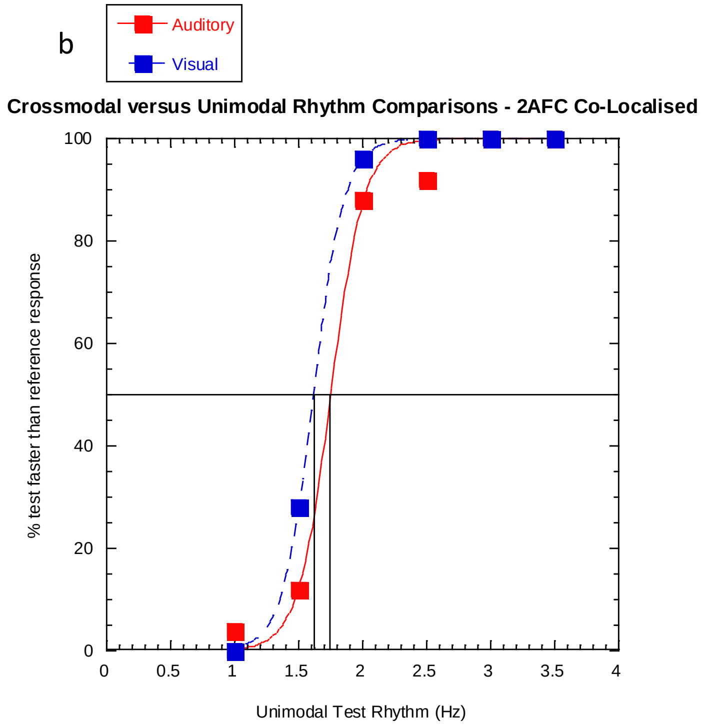
In the final experiment, subjects were required to match the same interleaved 3Hz audio-visual rhythm to its unimodal perceptual counterpart in a 2AFC design.

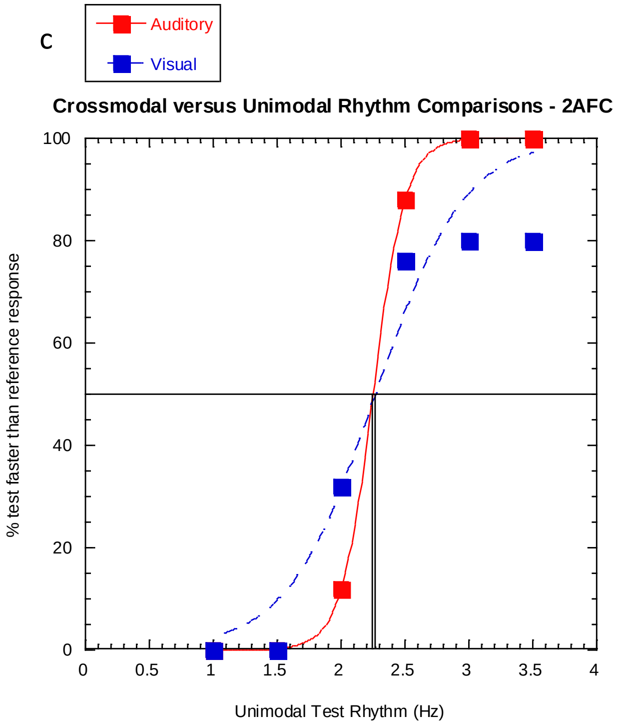
Duration discrimination judgments from the 2AFC experiments were fitted with a psychometric function using the subject's interval discrimination judgement of the proportion of responses of 'test longer than reference'. The functions were then fitted with a logistic of the form

$$\boldsymbol{y} = \frac{100}{1 + e^{\left(x - \left(\frac{\alpha}{\theta}\right)\right)}}$$

where '$\boldsymbol{\alpha}$' denotes the point of subjective equality (PSE – the 50% response level on a psychometric function) and '$\boldsymbol{\theta}'$' denotes an estimate of the interval discrimination threshold.


a
Auditory
Visual
Crossmodal versus Unimodal Rhythm Comparisons - Co-Localised 2AFC
% test faster than reference response
0
20
40
60
80
100
0
0.5
1
1.5
2
2.5
3
3.5
4
Unimodal Test Rhythm (Hz)


b
Auditory
Visual
Crossmodal versus Unimodal Rhythm Comparisons - 2AFC Co-Localised
% test faster than reference response
0
20
40
60
80
100
0
0.5
1
1.5
2
2.5
3
3.5
4
Unimodal Test Rhythm (Hz)



Figures 8.5a-c. 2AFC cross-modal versus unimodal rhythm results for each subject. PSE (point of subjective equality) comparisons to a unimodal visual rhythm are demonstrated with the blue curve whereas judgements to unimodal auditory rhythms are presented with the red curve. Three different subjects are plotted at the top (AM), middle (KB) and bottom (DW) of the panel.

Figure 8.5a corresponding to subject AM demonstrates that this subject matches a 3Hz AV rhythm to approximately 1.8Hz in the auditory modality and approximately 2.1Hz in the visual modality. Similarly, for subject KB, a 3Hz cross-modal rhythm is matched to a 1.8Hz in the auditory modality and 1.6Hz in the visual modality. Lastly, subject DW matches the same 3Hz AV rhythm to 2.26Hz in the auditory modality and 2.28Hz in the visual modality.

The pattern observed across subjects is that they match the cross-modal rhythm to an intermediate frequency in between both unimodal and cross-modal rhythm presentation.

8.8. Discussion

The present experiments indicate two key findings; one, that rhythm adaptation can produce evidence of duration after-effects, replicating the results from Chapter 6 and extending these findings to the reproduction of cross-modal intervals too. The second and perhaps most critical finding, is that these typical duration after-effects are also evidenced after being exposed to a 3Hz cross-modal rhythm. The after-effects however, show that the 500ms unimodal test interval is under-produced after having adapted to a cross-modal 3Hz rhythm (when compared to a baseline, no adaptation control) (Figure 8.4a). At the same time however, we also see an overestimation of the same 500ms test interval when it is demarcated by cross-modal signals, suggesting partial adaptation to a cross-modal rhythm also (Figure 8.4c). These results collectively suggest that even when two different sensory streams come together to act as one single rhythm, our perceptual systems are able to dynamically process them in a simultaneously integrated and segregated manner. Results from the 2AFC task corroborate these findings as each subject matched the 3Hz cross-modal rhythm to an intermediate unimodal rhythm that stood between 1.5-3Hz suggesting an interference of the co-occurring unimodal and cross-modal intervals (8.5a, 8.5b and 8.5c).

Despite all subjects demonstrating the same pattern of results, i.e. that matching a cross-modal rhythm to a unimodal rhythm results in an intermediate matching frequency, we observe marked differences in the tasks. Specifically, one subject (DW) demonstrated stronger cross-modal integration (Figure 8.5c) whereas another (subject AM) demonstrated a stronger unimodal bias (Figure 8.5a). A third subject (KB) demonstrated results in between the two (8.5b). In terms of responses to cross-modal intervals, the responses varied greatly depending on individual differences. For instance, in Experiment 3, subject DW showed consistency between reproduced values for AV and VA intervals, subject AM on the other hand, reproduced the AV interval almost veridically however interval VA was under-produced, suggesting a slight cross-modal after-effect. Contrastingly, subject KB over-produced all intervals but interval VA was most over-estimated in comparison to unimodal intervals and the AV test interval.

Individual differences therefore, appear to be a critical modulator of cross-modal thresholds. It has recently been found that the strength of an audio-visual correlation (temporally-correlated sensory signals originating from a single event) modulates multisensory perception (Nidiffer, Diederich, Ramachandran, & Wallace, 2018). Using a combination of psychophysics on human participants, it was reported that unimodal sensory temporal features were likely represented, at least in part, by individual differences within participants. After controlling for individual differences, it was found that cross-modal perception was varied linearly according to the correlation strength of the cross-modal signals, suggesting that the decisional system uses stimulus correlation as sensory evidence. Ultimately, the authors suggest that cross-modal (AV) correlation acts as an important cue to audio-visual integration and that the strength of this correlation acts as an influential determinant of the flexibility of these processes (Nidiffer et al., 2018). Individual differences also exist in the processing of beat perception and it has been suggested that these differences modulate motor and auditory area activity (Grahn & McAuley, 2009). fMRI results have also reflected the same influence of individual differences; suggesting that activation patterns reflected the patterns reproduced by the participants rather than actual veridical pattern presentation, suggesting that fMRI activation was indicative of the internal rhythm representation (Sakai et al., 1999). These results support the present evidence of individual differences modulating thresholds for cross-modal rhythm perception.

Future work could gather data on a wider participant group to address whether these inter-individual differences still prevail. Furthermore, the underlying neural mechanisms determining these individual differences still remain elusive. In a novel exploration of human brain anatomy and duration estimation of supra-second time scales for multi modal stimuli, it was found that inter-individual differences indeed exist however, specifically the authors found a link between discrimination of longer durations to “self-initiated rhythm maintenance mechanisms” (Gilaie-Dotan, Kanai, & Rees, 2011). The authors found that for either longer or shorter durations, duration estimation abilities correlated across modalities; demonstrating that across sensory modalities, common neural mechanisms exist. Furthermore, the structure of right auditory and somatosensory cortices (indicated by grey matter volume) and also para-hippocampal gyri during longer durations of stimulus presentation (12 seconds) predicted discrimination ability of longer durations in auditory and visual modalities. Despite providing cortical evidence for individual differences in time perception, what is also key here, is that classically, these structures have been thought of as modality-specific however, the current findings propose that they process temporal information in a modality-independent way. An avenue to explore in the future is to distinguish the nature and order of processing of sensory and temporal information; early research in this area currently suggests that duration perception precedes multisensory integration (Heron, Hotchkiss, Aaen-Stockdale, Roach, & Whitaker, 2013).

Previous work has also noted the attentional constraints placed when an individual is required to shift between modalities in the case of assessing intermodal intervals (Desmond & Moore, 1991). This may partly explain the differences we observed when subjects reproduced cross-modal intervals (AV and VA), and particularly the results of our naïve subject in non-co-localised tasks. It is plausible that intermodal timing may simply require a higher cognitive capacity as the participant is required to shift between modalities thereby contributing to an overall more difficult task, compared to the unimodal counterparts (Mauk & Buonomano, 2004). According to this view however, our results (for Experiment 3) should have presented constant error for the AV and VA conditions as the same modality-shift in cognitive resources was required, and under this presumption we should also observe the same effect for these two conditions. This result however, was not observed, suggesting the presence of distributed timing mechanisms that modulate incoming signals from different sensory modalities independently. Nevertheless, it has been suggested that multimodal stimuli in general evoke faster responses than unimodal stimuli alone. For instance, trimodal stimuli were responded to faster than their bimodal counterparts which were responded to faster than their unimodal counterparts

(Diederich & Colonius, 2004), presenting the additional benefit that multimodal stimuli afford human temporal and sensory processing.

An important question to note is whether sounds and visual input need to arise from the same spatial location in order to activate a centralised timing mechanism. Research has found that even intervals demarcated by the same modality can produce large levels of variability in discrimination (Grondin, 1998), and that performance is markedly improved when signals arise from the same source compared to a large distance between visual signals (Grondin, 1998). Temporal ventriloquism refers to the observation that a sound presented in close proximity to a visual signal can modify the perception of the light (Vroomen & Keetels, 2006; Bausenhart, de la Rosa, & Ulrich, 2014). In a series of experiments, subjects responded to the temporal order of several visual signals and reported which occurred first, whilst in some conditions the visual signals were also accompanied by sounds before the first and after the second visual signal. Vroomen & Keetels results challenge the notion that co-localisation in a spatial sense must occur for intersensory interactions (Vroomen & Keetels, 2006). Results from the present experiment however conflict with this conclusion. Explicitly we find marked amplification in after-effects when auditory and visual streams are co-localised compared to when they are not. Whilst non-co-localised auditory and visual streams are still able to elicit limited evidence of after-effects post-cross-modal adaptation, these after-effects are much stronger when the two streams are co-localised. Future work investigating multimodal rhythm perception should be equally considerate of spatial influences on sensory perception and aim to co-localise stimuli if the hypothesis aims to assess cross-modal time perception.

Differences between sensory modalities have been presented in tasks using a concurrent (and non-temporal) distractor task. In one such case, a temporal discrimination task was performed either alone or in another condition alongside a self-paced finger tapping task. Results suggest that the processing of time within a brief time period is likely to be automatic but only for the auditory modality, and not the visual (Mioni et al., 2016). Moreover, despite several conflicting reports of reciprocal influences between audition and vision, recent evidence has implicated a more balanced and reciprocal relationship between audition and touch (Villanueva & Zampini, 2018). It was found that an irrelevant and distractor sensory modality was able to modulate subjective duration of a target stimulus in another modality in a 2AFC design. Results demonstrated that the distractor modality influenced duration perception of both auditory and visual modalities. Interestingly, they also found evidence of a tactile influence in enhancing auditory perception, but audition did not facilitate such an enhancement on tactile perception of duration (Villanueva & Zampini, 2018). These results imply some sort of communicative mechanism between audition and touch to efficiently process duration information – a mechanism that is apparently (in their investigation) not as conducive to processing in vision. Similarly, a speeding up effect (duration compression) is found with an auditory-motor pairing but not with the visuo-motor counterpart in temporal recalibration. It is further suggested that changes in auditory processing speed mediate these changes in audio-motor synchrony and the resulting window of temporal recalibration (Sugano, Keetels, & Vroomen, 2017). Contrastingly, we find evidence for the opposite – that equally communicative channels exist between audition and vision that result in sensory signals from both these modes being combined to process rhythmic timing information. A possible explanation for why others may have found opposing results is due to differing task foci for example, a task requiring participants to respond accurately may then result in the participant placing more weighting on the auditory signals over their visual counterparts whereas a task requiring judgement of the signal source may cause the same participant to shift their focus on the location of visual as opposed to auditory signals.

It has been shown that perceived synchrony of audio-visual pairs influences temporal interval discrimination (van Eijk, Kohlrausch, Juola, & van de Par, 2009). The authors presented audio-visual pairs with a range of relative delays with audio leading to visual leading and found that pairs are regularly judged as "synchronous". It was for this reason that we were not able to repeat our cross-modal rhythm conditions with alternative temporal frequencies as any faster than 3Hz would have resulted in the two streams appearing synchronous whereas any slower would have prohibited an integration of the two sensory streams as one. Large-scale lesion studies have implicated several areas in neural networks that support audio-visual integration in the context of speech (Hickok et al., 2018). Specifically, lesions incorporating the superior temporal auditory, lateral occipital visual and multisensory areas of the superior temporal sulcus are thought to be most damaging to AV speech integration. Compellingly, the location of the lesion is able to predict failure in AV speech integration, for example, whether this is an auditory or visual capture failure. It is concluded that both, unimodal auditory and visual cortices support AV speech integration, but also that these are recruited alongside more multimodal regions such as the superior temporal sulcus (Hickok et al., 2018). Similar evidence has been presented in estimating visual distance, particularly that AV cues improve the precision in visual distance estimation (Jaekl, Seidlitz, Harris, & Tadin, 2015). Indeed, it has even been reported that the awareness of visual events can be improved with sounds despite the lack of cross-modal integration (Pápai & Soto-Faraco, 2017) suggesting the continuous unconscious recruitment of our senses despite no conscious call for recruitment. Studies presenting audio-visual rhythms through movement observation via human point-light figures have shown that rhythmic movement assists the perception of auditory rhythms and that these may be subserved by the internal motor system that may be coupled with a perceptually integrated audio-visual beat (Su, 2014). An increasing body of work is being gathered to highlight the mechanisms by which our sensory systems aid the processing of time and other features of external environments. The present work is significant in establishing that our sensory timing system adapts an organisational system with considerable complexity to process sensory input. We observe that when a participant is presented with a highly complex rhythm such as a perfectly interleaved AV rhythm with no suggestion regarding which signals to focus on, the participant organises the sensory input into both unimodal and cross-modal rhythms and processes both simultaneously (Figure 8.5). This is an incredible feat considering the speed and efficiency with which these processes are carried out. Future work could aim to investigate the underlying neural processes that coincide to carry out these functions.

It is important to question whether the results evidenced here are a result of subjects equally integrating and processing the two different sensory streams. For instance, it could be possible that interleaving auditory and visual rhythms could result in better integration as demonstrated here. Alternatively, it could be possible that the two streams weren't better integrated but a higher-order processing system exists that is able to override the impairment in integration and still process all sensory input that the subject is exposed to. One solution to this would be to interleave the two streams (yet not perfectly interleave them) and alter the intervals between the streams to observe whether the subject fails to integrate these into a single perceptual rhythm, i.e. how instrumental is regularity in the formation of combining two sensory rhythms as a single percept?

It is clear that much work remains to be considered in understanding why and how the timing system integrates signals from multiple sensory modalities. Factors such as attention, rhythm structure and physical location will be particularly relevant for future exploration. The body of work conducted here provides an important addition to this field and helps set up the landscape for future studies to build upon. We find a range of distortions in sensory time perception when subjects are exposed to a perfectly interleaved and co-localised audio-visual rhythm presented at 3Hz. We importantly find that when subjects are required to match the cross-modal rhythm to a unimodal rhythm, subjects match the 3Hz AV rhythm to an intermediate unimodal rhythm falling between 1.5-3Hz. In the adaptation tasks, we find that after being exposed to a 3Hz AV rhythm, subjects reproduce a 500ms empty test interval demarcated by unimodal signals as shorter than 500ms. The same interval marked by cross-modal signals is reproduced as longer than 500ms. Collectively these results suggest that exposing subjects to a cross-modal rhythm without any prior knowledge or information results in a simultaneous interference of cross-modal and unimodal intervals. We observe that subjects are able to process unimodal signals efficiently but that they are also able to combine unimodal and cross-modal signals to process the global overarching rhythm. Overall, these findings provide a valuable contribution in aiding our understanding of how the human perceptual system organises timing signals arising from multiple senses to carry out vital functions in everyday life.

9. Discussion

There is no doubt that the brain is striking in its ability to not only integrate, but also organise and utilise the multitude of spatial and temporal signals it constantly receives. The finer details underlying how it performs these actions are currently being explored. The experiments presented within this thesis have employed sensory adaptation to explore the mechanisms underlying human sensory time perception. The specific focus has been on whether the human system relies on centralised or distributed processing mechanisms.

The hypothesis that centralised timers mediate temporal processing implies that all neural circuits could naturally process timing information and in principle, any neural circuit could be recruited. The nature of the task paired with the sensory modality of the presented information, would then direct the location of the spatial processing of this information. In simple terms, a hypothesis promoting the centralised theory of timing suggests the presence of an over-arching temporal mechanism overseeing the processing of incoming sensory signals from multiple modalities. In contrast, the distributed theory of timing posits multiple dedicated timing mechanisms. Each timer dictates the processing of a specific sensory modality and functions independently of timers corresponding to other sensory modalities.

In Experiment 1 (Chapter 5), subjects were adapted to varying temporal frequencies ranging between 1.05-8.46Hz presented to the auditory, tactile and visual modalities. The task then required the same subjects to reproduce a 3Hz test frequency that was crucially either presented within the same or a different modality. They reproduced this test stimulus by tapping the temporal frequency with their index finger on a response disk. Each subject was adapted and tested with each possible sensory combination which totalled to 9 sensory pairings. Control experiments were also run where the subject was unaware of the sensory modality that the test stimulus would be presented to and also a second control experiment using the audio-visual pairing where stimuli presentation was co-localised in space. All other experimental details were kept the same as the main experiments.

The results from these experiments provide important insights into the processing of unimodal rhythms and the factors that govern these mechanisms. In particular, of great significance is the finding that rhythm after-effects exist for the auditory, touch and visual sensory modalities. Hitherto, considerable evidence has been mounting to demonstrate the presence of duration after-effects and the presence of neural channels dedicated to the processing of specific durations of time in a band-limited manner. As it is observed that these after-effects (in Chapter 5) dissipate as the discrepancy between the adapting and test modalities increase, the results described in Chapter 5 demonstrate clear evidence of similarly tuned channels in the perception of rhythm too.

Furthermore, subsequent experimentation demonstrated that these effects prevail for each modality even when subjects are unaware of the sensory modality of the test stimulus. Importantly, this demonstrates that the strength of this after-effect is undeterred by attentional factors. The second control experiment again demonstrated the presence of rhythm after-effects that are modality-specific even when the adapting and test modalities are spatially co-localised.

Collectively, the series of experiments make critical advances regarding temporal after-effects and demonstrate the first evidence of modality-specific rhythm after-effects in audition, touch and vision evidenced with the method of rate reproduction. These after-effects prevail despite controlling for cognitive factors such as attentional focus on specific aspects of the task and also prevail even after co-localising audio and visual sensory streams in space. As no observation was made of a cross-modal rhythm after-effect, these results provide unambiguous support for distributed theories of timing.

To more comprehensively understand the processing of rhythm, the aim of the second experiment (Chapter 6) was to identify how the processing of rhythm explicitly differentiates from the processing of the single intervals that compose these rhythms. To do this, an identical rhythm adaptation method was used as in Chapter 5 however in this experiment, the test phase presented subjects with a single empty interval of 333ms which they were required to reproduce by tapping the response disk. In a second experiment, the subjects were again adapted to rhythms as before, but were presented with a two-alternative forced-choice test where they were required to compare a reference interval of 333ms presented to the modality not adapted to, to a test interval ranging from 282-382ms and were asked to respond to which was longer. In a control experiment, the interval reproduction experiments were replicated with a *filled* test interval.

The key finding was that subjects are able to adapt to rhythms and demonstrate these repulsive adaptation-induced after-effects when presented with empty and filled test intervals. This result was also replicated with filled intervals and reproduced in both methods of interval reproduction and two-alternative forced-choice. Importantly, these results provoke us to critically consider rate perception and how exactly, the brain processes signals received during exposure to a rhythm. Immediately, these results may encourage the reader to consider that the processing of rhythm is underlain by a duration processing mechanism. This assertion, however, lies at odds with the fundamental uniqueness of rhythm, for example, it is commonly reported that momentary exposure to a rhythm is sufficient to entrain an individual to the beat and respond by (often unintentionally) tapping or otherwise, moving along to a beat. At the same time, it is also important to consider that any duration-specific mechanism would in theory, be able to process rhythms as any regular rhythm would in essence, be the presentation of a duration in a loop. Contrastingly, any rhythm-specific mechanism would be tuned to rhythms alone and therefore, would have to exist alongside a separate duration-specific mechanism. It may be plausible that both mechanisms can co-exist and future work could aim to investigate the thresholds that determine which processing mechanism is adopted (see below).

As a result of the findings in Chapter 5 and 6, a critical addition has been made to the field in linking the processing of duration and rate and the modality-specificity of both. Much of the work conducted hitherto on duration perception and duration channels asserts the modality-specificity of duration perception. Prior to the experiments conducted in Chapter 5, there were ongoing contentious discussions regarding the modality-specificity of rate also. Considering the debate presented in Chapter 6 regarding the controversial link between interval-based and beat-based timers, the results from these experiments elucidate several elements of these debates. Explicitly, they present the evidence to suggest that the processing of duration can be extended to the processing of rate also, and that the processing of both types of temporal signals are modality-specific.

The concept of rhythms was further decomposed in experiment three (Chapter 7) in order to establish whether semi-isochronous rhythms were processed differently to rhythms adhering to stricter isochronicity. Specifically, subjects were presented with temporally-varied rhythms in the auditory, tactile and visual modalities, and were asked to compare these to perfectly isochronous rhythms in a two-alternative forced-choice task used to establish a baseline threshold for regularity. In the proceeding experiment, the same subjects were adapted to unimodal patterned rhythms in the auditory, tactile and visual modalities and were then presented with a single rhythmic test stimulus that varied in its isochronicity. The task required subjects to respond to whether the test stimulus was regular or not in a single stimulus design.

The results of this experiment were more ambiguous than the experiments preceding it and it was found that some subjects (but not all), were able to adapt to anisochronous sequences in certain conditions. Moreover, the results were unable to provide a consistent pattern of effects, either within a subject or across sensory modalities.

In Chapter 6, it was found that rhythms were processed in terms of the single durations composing the overarching rhythmic sequence, suggesting the possibility that the rhythm after-effect is supported by the duration after-effect acting in a repeating loop. The sequences used to explore anisochrony in Chapter 7 were quasi-regular patterned sequences alternating as long-short-long (and so on) intervals. One explanation for the results may be that alternating short and long intervals discourage the build-up of consistent duration signals thereby resulting in a failure to adapt (akin to the lack of a duration after-effect, as each interval presents alternating information).

The final experiment (Chapter 8) examined the presence of centralised timing mechanisms more directly. Specifically, I asked whether a perfectly interleaved and spatially co-localised audio-visual rhythm could be combined as one percept to invoke the presence of such a mechanism.

As with previous experiments, this experiment also employed an adaptation paradigm. Subjects were exposed to a 3Hz audio-visual rhythm that was perfectly interleaved. This meant that unimodally, the signals were presented at 1.5Hz however, as a single cross-modal rhythm they were presented at 3Hz. After 10 seconds of adaptation to this co-localised rhythm, subjects were presented with a 500ms test interval that was demarcated by either unimodal (auditory/visual) or cross-modal (AV/VA) signals. The subject’s task required them to reproduce this empty interval by tapping on a response disk as before.

A second experiment presented subjects with the same cross-modal rhythm, followed by a unimodal rhythm (ranging from 1-4Hz) in a two-alternative forced-choice design and required subjects to match the cross-modal rhythm to its unimodal counterpart by indicating which rhythm was faster using a keypress.

Results demonstrated that subjects matched a 3Hz AV rhythm to an intermediate unimodal rhythm ranging between 1.5-3Hz. The same pattern of results was evidenced when using the adaptation and interval reproduction method in spatially co-localised conditions as each subject demonstrated support for unimodal and cross-modal adaptation when reproducing unimodal and cross-modal empty intervals. The extent to which subjects were able to integrate unimodal audio and visual signals into a combined bi-modal rhythm appeared to be influenced by individual thresholds.

Whilst the primary focus of this thesis has been to clarify the role of centralised and distributed mechanisms underlying sensory time, a recurring theme emerges in the distinction between duration-based and rate-based processing. Particularly, evidence that the processing of duration and rate are underlain by the same, or at least by heavily interlinked mechanisms recurs consistently across Chapter 6. This finding is further reinforced in Chapter 8 and extended to stimuli cross-modal in structure.

9.1. Conclusions

It is unequivocally clear that the processing of time is relentlessly dynamic and in a constant state of recalibration. The collection of experiments presented here have provided invaluable additions to understanding human timing. Importantly, I have demonstrated the presence of unimodal auditory, touch and visual rhythm after-effects akin to those evidenced with duration. These effects prevail despite controlling for attentional and spatial influences. Moreover, in deconstructing the perception of rhythms, I have highlighted the processing similarities between duration and rate and demonstrated the modality-specificity of both. In work exploring irregular rhythms, the significance of consistent temporal signals has been revealed. In the final experiment, it was found that exposing subjects to a perfectly interleaved audio-visual rhythm results in subjects simultaneously being able to adapt to unimodal and cross-modal signals. Compellingly, it has been revealed that when asked to match an interleaved cross-modal rhythm to its unimodal counterpart, subjects match the cross-modal rhythm to an intermediate frequency in between both unimodal and cross-modal rhythms.

In conclusion, the results presented within this thesis have made important contributions to the understanding of human sensory time perception. The extent to which sensory signals are processed in a distributed manner, and whether these eventually extend to centralised mechanisms, will be an important avenue for future work to explore.

9.2. Future Work

In this thesis, I have made several important advances in our understanding of human sensory time perception. These findings, paired with the inherent complexity of timing mechanisms provide many exciting opportunities for future work.

An immediate opportunity for future work may be to replicate these experiments with a wider band of temporal frequencies. This approach is key in identifying the limits of the perceptual system and also enabling experimenters to comprehensively match temporal sensitivity in each different sensory modality. In terms of experimental limits particularly in relation to the questions investigated in Chapter 6, it is known that there are considerable perceptual phenomena driven by temporal frequencies too high for the human system to be able to reasonably extract the component durations. For instance, in situations where the temporal frequency judgements are far less variable than the variability associated with their component durations (for example, a 60Hz flicker discrimination JND versus a 17ms visual discrimination JND). If future work can improve on the current temporal constraints of apparatus commonly used in psychophysical experiments, we may more robustly be able to examine just how important component durations are in the perception of rhythms and rhythmic stimuli.

Moreover, further investigation will be pivotal in explaining the questions that have arisen, particularly from Chapter 7 and 8. In Chapter 7, there was considerable lack of consistency even within each subject, demonstrating the impact of individual differences in thresholds for regularity. One reason for this may have been that patterned temporal sequences in the form of alternating long-short-long-short durations are not commonly experienced in the natural world or other social contexts. The use of more ecologically valid stimuli, such as speech may enable us to more robustly measure the importance of isochronicity when investigating non-regular rhythms.

Several avenues for future exploration have arisen as a result of the experiments conducted in Chapter 8. One such question arising from these findings is why some subjects were able to demonstrate evidence of adaptation to unimodal and cross-modal rhythms simultaneously (in spatially co-localised conditions). Moreover, why another subject was unable to demonstrate this effect despite equal hours of adaptive exposure and testing. Further experimentation investigating the extent to which individual thresholds govern the integration of unimodal and cross-modal signals will importantly elucidate these differences.

The evidence presented from Chapter 8 posits the presence of centralised timing mechanisms alongside distributed timing mechanisms. This highlights the complexities a single model faces in attempting to encompass the multitude of facets and features of time. Additionally, it indicates the need to further develop the current models to incorporate modality-specific and modality-independent features of time. Neural evidence has suggested the presence of multimodal and modality-specific neurons and therefore, future work could spatially map the range within which these different neurons lie. This work could then clarify the conflicting evidence presented here.

Alternative implications for this work in settings outside of a psychophysics laboratory are also becoming increasingly more relevant. Recent research in robotics has identified that automating machines to demonstrate sensory-guided motor behaviour is perhaps one of the most challenging aspects of building intelligent machines in society. This work provides some of the groundwork to the structure of human time perception and will allow future work to explore these functions more coherently.

Lastly, it is imperative to question why there are no individuals with core difficulties with time processing (akin to memory and individuals with amnesia), and why humans do not have a dedicated timing organ. The answers to these questions lie in the fundamental importance of timing and time perception. As the results found here are only the beginning of a long and exciting journey in human sensory time perception, future work will provide a more comprehensive picture of the relationship between our senses and time. A particularly exciting feat will be disentangling the extent to which sensory systems are subserved by distributed timing mechanisms, and under which circumstances (if any), they can extend to a centralised system of processing.

10. References

Abel, S. M. (1972). Discrimination of temporal gaps. Journal of the Acoustical Society of America, 52(2B), 519-524.

Acoustics—Normal Equal-Loudness Contours [ISO 226:1987]. (1987). International Organization for Standardization, Geneva.

Akdoğan, B., Balcı, F., & van Rijn, H. (2016). Temporal expectation indexed by pupillary response. Timing & Time Perception, 4(4), 354-370.

Allan, L. G. (1979). Perception of time. Perception & Psychophysics, 26, 340-354.

Allman, M. J., & Meck, W. H. (2011). Pathophysiological distortions in time perception and timed performance. Brain, 135(3), 656-677.

Amesbury, E. C., & Schallhorn, S. C. (2003). Contrast sensitivity and limits of vision. International Ophthalmology Clinics, 43(2), 31-42.

Athanasopoulos, G., & Moran, N. (2013). Cross-cultural representations of musical shape. Empirical Musicology Review, 8(3-4), 185-199.

Bailey, I. L., Bullimore, M. A., Raasch, T. W., & Taylor, H. R. (1991). Clinical grading and the effects of scaling. Investigative Ophthalmology & Visual Science, 32(2), 422-432.

Baker, R., Dexter, M., Hardwicke, T. E., Goldstone, A., & Kourtzi, Z. (2014). Learning to predict: Exposure to temporal sequences facilitates prediction of future events. Vision Research, 99, 124-133.

Barlow, H. B., & Hill, R. M. (1963). Evidence for a physiological explanation of the waterfall phenomenon and figural after-effects. Nature, 200, 1345-1347.

Bartholomew, A. J., Meck, W. H., & Cirulli, E. T. (2015). Analysis of genetic and non-genetic factors influencing timing and time perception. *PLoS One*, *10*(12), e0143873.

Bausenhart, K. M., Bratzke, D., & Ulrich, R. (2016). Formation and representation of temporal reference information. Current Opinion in Behavioral Sciences, 8, 46-52.

Bausenhart, K. M., de la Rosa, M. D., & Ulrich, R. (2014). Multimodal integration of time. Experimental Psychology, 61(4), 310-322.

Bausenhart, K. M., Dyjas, O., & Ulrich, R. (2015). Effects of stimulus order on discrimination sensitivity for short and long durations. Attention, Perception & Psychophysics, 77(4), 1033-1043.

Becker, M. W., & Rasmussen, I. P. (2007). The rhythm aftereffect: Support for time sensitive neurons with broad overlapping tuning curves. Brain and Cognition, 64(3), 274-281.

Bedford, F. L. (1989). Constraints on learning new mappings between perceptual dimensions. Journal of Experimental Psychology: Human Perception and Performance, 15(2), 232-248.

Behar, I., & Bevan, W. (1961). The perceived duration of auditory and visual intervals: cross-modal comparison and interaction. American Journal of Psychology, 74, 17-26.

Berglund, B., Berglund, U., Ekman, G., & Frankehaeuser, M. (1969). The influence of auditory stimulus intensity on apparent duration. Scandinavian Journal of Psychology, 10(1), 21-26.

Berson, D. M., Dunn, F. A., & Takao, M. (2002). Phototransduction by retinal ganglion cells that set the circadian clock. Science, 295(5557), 1070-1073.

Binetti, N., Lecce, F., & Doricchi, F. (2012). Time-dilation and time-contraction in an anisochronous and anisometric visual scenery. Journal of Vision, 12(7), 8.

Blakemore, C. T., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. Journal of Physiology, 203(1), 237-260.

Borsook, D., Becerra, L., Fishman, S., Edwards, A., Jennings, C. L., Stojanovic, M., Papinicolas, L., Ramachandran, V. S., Gonzalez, R. G. & Breiter, H. (1998). Acute plasticity in the human somatosensory cortex following amputation. Neuroreport, 9(6), 1013-1017.

Bouwer, F. L., Burgoyne, J. A., Odijk, D., Honing, H., & Grahn, J. A. (2018). What makes a rhythm complex? The influence of musical training and accent type on beat perception. PLoS ONE, 13(1), e0190322.

Brannon, E. M., Libertus, M. E., Meck, W. H., & Woldorff, M. G. (2008). Electrophysiological measures of time processing in infant and adult brains: Weber's Law holds. Journal of Cognitive Neuroscience, 20(2), 193-203.

Brebner, J. T. a. A. T. W. (1980). Introduction: an historical background sketch. In A. T. Welford (Ed.), Reaction Times (pp. 1-23). New York: Academic Press.

The Editors of Encyclopaedia Britannica (30/05/2016). Weber’s Law. Encyclopædia Britannica, inc. Retrieved from https://www.britannica.com/science/Webers-law. Access date: 19/03/2019

Brochard, R., Touzalin, P., Despres, O., & Dufour, A. (2008). Evidence of beat perception via purely tactile stimulation. Brain Research, 1223, 59-64.

Bruno, A., & Cicchini, G. M. (2016). Multiple channels of visual time perception. Current Opinion in Behavioral Sciences, 8, 131-139.

Bruno, A., Ayhan, I., & Johnston, A. (2015). Changes in apparent duration follow shifts in perceptual timing. Journal of Vision, 15(6), 2.

Buffardi, L. (1971). Factors affecting the filled-duration illusion in the auditory, tactual, and visual modalities. Perception & Psychophysics, 10(4), 292-294.

Burr, D., Tozzi, A., & Morrone, M. C. (2007). Neural mechanisms for timing visual events are spatially selective in real-world coordinates. Nat Neurosci, 10(4), 423-425.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. Journal of physiology, 197(3), 551-566.

Carlile, S. (1996). The physical and psychophysical basis of sound localization virtual auditory space: Generation and applications (pp. 27-78). Berlin, Heidelberg: Springer Berlin Heidelberg.

Celma-Miralles, A., de Menezes, R. F., & Toro, J. M. (2016). Look at the beat, feel the meter: Top-down effects of meter induction on auditory and visual modalities. Frontiers in Human Neuroscience, 10, 108.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. Cerebral Cortex, 18(12), 2844-2854.

Chen, K. M., & Yeh, S. L. (2009). Asymmetric cross-modal effects in time perception. Acta Psychologica, 130(3), 225-234.

Chen, L., & Vroomen, J. (2013). Intersensory binding across space and time: a tutorial review. Attention, Perception, & Psychophysics, 75(5), 790-811.

Chen, L., & Zhou, X. (2014). Fast transfer of crossmodal time interval training. Experimental Brain Research, 232(6), 1855-1864.

Cirelli, L. K., Trehub, S. E., & Trainor, L. J. (2018). Rhythm and melody as social signals for infants. Annals of the New York Academy of Sciences, 66-72.

Conard, N. J., Malina, M., & Münzel, S. C. (2009). New flutes document the earliest musical tradition in southwestern Germany. Nature, 460(7256), 737-740.

Craig, J. C. (1973). A constant error in the perception of brief temporal intervals. Perception & Psychophysics, 13(1), 99-104.

Craig, J. C. (1999). Grating orientation as a measure of tactile spatial acuity. Somatosensory & Motor Research, 16(3), 197-206.

Cravo, A. M., Rohenkohl, G., Wyart, V., & Nobre, A. C. (2013). Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. Journal of Neuroscience, 33(9), 4002-4010.

Bueti, D. & Macaluso, E. (2010). Auditory temporal expectation modulates activity in visual cortex. Neuroimage, 51, 1168-1183.

Dale, C. L., Gratton, G., & Gibbon, J. (2001). Event-related brain potentials isolate the motor component in a tapping task. Neuroreport, 12(14), 3015-3018.

Dalla Bella, S., Farrugia, N., Benoit, C. E., Begel, V., Verga, L., Harding, E., & Kotz, S. A. (2017). BAASTA: Battery for the Assessment of Auditory Sensorimotor and Timing Abilities. Behavior Research Methods, 49(3), 1128-1145.

Deneve, S., & Pouget, A. (2004). Bayesian multisensory integration and cross-modal spatial links. Journal of Physiology, 98(1), 249-258.

Deroy, O., & Spence, C. (2016). Crossmodal correspondences: Four challenges. Multisensory Research, 29(1-3), 29-48.

Desmond, J. E., & Moore, J. W. (1991). Altering the synchrony of stimulus trace processes: Tests of a neural-network model. Biological Cybernetics, 65(3), 161-169.

Di Luca, M., & Rhodes, D. (2016). Optimal perceived timing: Integrating sensory information with dynamically updated expectations. Scientific Reports, 6, 28563.

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

Doehrmann, O., & Naumer, M. J. (2008). Semantics and the multisensory brain: How meaning modulates processes of audio-visual integration. Brain Research, 1242, 136-150.

Drake, C., & Botte, M. C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. Perception & Psychophysics, 54(3), 277-286.

Duarte, F., & Lemus, L. (2017). The time is up: Compression of visual time interval estimations of bimodal aperiodic patterns. Frontiers in Integrative Neuroscience, 11, 17.

Efron, R. (1970). The minimum duration of a perception. Neuropsychologia, 8(1), 57-63.

Ehrlé, N., & Samson, S. (2005). Auditory discrimination of anisochrony: Influence of the tempo and musical backgrounds of listeners. Brain and Cognition, 58(1), 133-147.

Ehrle, N., Samson, S., & Baulac, M. (2001). Processing of rapid auditory information in epileptic patients with left temporal lobe damage. Neuropsychologia, 39(5), 525-531.

Eisler, A. D., & Eisler, H. (2001). Subjective time in a patient with neurological impairment. Psychologica, 28, 193-206.

Elliott, D. B., & Flanagan, J. (2007). 3 – Assessment of visual function: Clinical Procedures in Primary Eye Care (Third Edition) (pp. 29-81). Edinburgh: Butterworth-Heinemann.

Ellis, M. C. (1991). Research note: Thresholds for detecting tempo change. Psychology of Music, 19(2), 164-169.

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

Ernst, M. O., & Bulthoff, H. H. (2004). Merging the senses into a robust percept. Trends Cognitive Sciences, 8(4), 162-169.

Erp, J. B. F. v., & Werkhoven, P. J. (2004). Vibro-tactile and visual asynchronies: Sensitivity and consistency. Perception, 33(1), 103-111.

Large, E. M., & Jones, M. R. (1999). The dynamics of attending: How we track time-varying events. Psychological Review, 106, 119-1159.

Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Processe. Archiv für die gesamte Physiologie des Menschen und der Tiere, 11(1), 581-602.

Fechner, G. T. (1860). Elements of Psychophysics. Leipzig: Breitkopf & Harterl.

Ferreira, V. F. M., Paiva, G. P., Prando, N., Graça, C. R., & Kouyoumdjian, J. A. (2016). Time perception and age. *Arquivos de neuro-psiquiatria*, *74*(4), 299-302.

Fidopiastis, C., Fuhrman, C., Meyer, C., & Rolland, J. (2005). Methodology for the iterative evaluation of prototype head-mounted displays in virtual environments: Visual acuity metrics. Presence, 14(5), 550-562.

Fraisse, P. (1984). Perception and estimation of time. Annual Review of Psychology, 35(1), 1-37.

Freeman, J. S., Cody, F. W., O'Boyle, D. J., Craufurd, D., Neary, D., & Snowden, J. S. (1996). Abnormalities of motor timing in Huntington's disease. Parkinsonism & Related Disorders, 2(2), 81-93.

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. Journal of the Acoustical Society of America, 98(5), 2524-2531.

Fulcher, C. (2017). The role of sensory history and stimulus context in human time perception. (Doctor of Philosophy), University of Bradford, Bradford.

Gamache, P. L., & Grondin, S. (2010). Sensory-specific clock components and memory mechanisms: Investigation with parallel timing. European Journal of Neuroscience, 31(10), 1908-1914.

García-Pérez, M. A. (2014). Does time ever fly or slow down? The difficult interpretation of psychophysical data on time perception. Frontiers in Human Neuroscience, 8, 415.

Gardner. (2019). Signal Detection Simulation. Retrieved 03/09/2019, from http://gru.stanford.edu/doku.php/tutorials/sdt

Garner, W. R., & Gottwald, R. L. (1968). The perception and learning of temporal patterns. Quarterly Journal of Experimental Psychology, 20(2), 97-109.

Geringer, J. M., & Madsen, C. K. (1984). Pitch and Tempo Discrimination in Recorded Orchestral Music among Musicians and Nonmusicians. Journal of Research in Music Education, 32(3), 195-204.

Giersch, A., Poncelet, P. E., Capa, R. L., Martin, B., Duval, C. Z., Curzietti, M., Hoonacker, M., van Assche, M. & Lalanne, L. (2015). Disruption of information processing in schizophrenia: The time perspective. Schizophrenia Research: Cognition, 2(2), 78-83.

Gil, S., & Droit-Volet, S. (2011). "Time flies in the presence of angry faces"... depending on the temporal task used! Acta Psychologica, 136(3), 354-362.

Gilaie-Dotan, S., Kanai, R., & Rees, G. (2011). Anatomy of human sensory cortices reflects inter-individual variability in time estimation. Frontiers in Integrative Neuroscience, 5, 76.

Glenberg, A. M., & Jona, M. (1991). Temporal coding in rhythm tasks revealed by modality effects. Memory & Cognition, 19(5), 514-522.

Cicchini, G. M., Arrighi, R., Cecchetti, L., Giusti, M. & Burr, D. C. (2012). Optimal encoding of interval timing in expert percussionists. Journal of Neuroscience, 32, 1056.

Goldstein, B. E. (2007). Introduction to Vision. In J. Hague & A. Perkins (Eds.), Sensation and Perception (Eighth Edition ed., pp. 44-133). California, USA: Wadsworth Cengage Learning.

Goldstone, S., & Goldfarb, J. L. (1963). Judgment of filled and unfilled durations: Intersensory factors. Perceptual and Motor Skills, 17(3), 763-774.

Goldstone, S., & Goldfarb, J. L. (1964). Auditory and visual time judgement. Journal of General Psychology, 70, 369-387.

Goldstone, S., Lhamon, W. T., & Sechzer, J. (1978). Light intensity and judged duration. Bulletin of the Psychonomic Society, 12(1), 83-84.

Gontier, E., Hasuo, E., Mitsudo, T., & Grondin, S. (2013). EEG investigations of duration discrimination: The intermodal effect is induced by an attentional bias. PLoS ONE, 8(8), e74073.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. Trends in Neuroscience, 15(1), 20-25.

Grahn, J. & Rowe, J. (2013). Finding and feeling the musical beat: Striatal dissociations between detection and prediction of regularity. Cerebral Cortex, 23, 913-921.

Grahn, J. A. (2012). Neural mechanisms of rhythm perception: Current findings and future perspectives. Topics in Cognitive Science, 4(4), 585-606.

Grahn, J. A. (2012). See what I hear? Beat perception in auditory and visual rhythms. Experimental Brain Research, 220(1), 51-61.

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. Neuroimage, 47, 1894-1903.

Grondin, S. (1993). Duration discrimination of empty and filled intervals marked by auditory and visual signals. Perception & Psychophysics, 54(3), 383-394.

Grondin, S. (1998). Judgments of the duration of visually marked empty time intervals: Linking perceived duration and sensitivity. Perception & Psychophysics, 60(2), 319-330.

Grondin, S. (2001). Discriminating time intervals presented in sequences marked by visual signals. Perception & Psychophysics, 63(7), 1214-1228.

Grondin, S., & McAuley, J. D. (2009). Duration discrimination in crossmodal sequences. Perception, 38(10), 1542-1559.

Grondin, S., & Rousseau, R. (1991). Judging the relative duration of multimodal short empty time intervals. Perception & Psychophysics, 49(3), 245-256.

Grondin, S., & Ulrich, R. (2011). Duration discrimination performance: No cross-modal transfer from audition to vision even after massive perceptual learning. Multidisciplinary Aspects of Time and Time Perception, 92-100.

Grondin, S., & Ulrich, R. (2011). Duration Discrimination Performance: No Cross-Modal Transfer from Audition to Vision Even after Massive Perceptual Learning. In A. Vatakis, A. Esposito, M. Giagkou, F. Cummins, & G. Papadelis (Eds.), Multidisciplinary Aspects of Time and Time Perception: COST TD0904 International

Workshop, Athens, Greece, October 7-8, 2010, Revised Selected Papers (pp. 92-100). Berlin, Heidelberg: Springer Berlin Heidelberg.

Grondin, S., Ivry, R. B., Franz, E., Perreault, L., & Metthe, L. (1996). Markers' influence on the duration discrimination of intermodal intervals. Perception & Psychophysics, 58(3), 424-433.

Grondin, S., Meilleur-Wells, G., Ouellette, C., & Macar, F. (1998). Sensory effects on judgments of short time-intervals. Psychological Research, 61(4), 261-268.

Grondin, S., Roussel, M. E., Gamache, P. L., Roy, M., & Ouellet, B. (2005). The structure of sensory events and the accuracy of time judgments. Perception, 34(1), 45-58.

Grube, M., Cooper, F. E., Chinnery, P. F., & Griffiths, T. D. (2010). Dissociation of duration-based and beat-based auditory timing in cerebellar degeneration. Proceedings of the National Academy of Sciences, 107(25), 11597-11601.

Gu, B. M., Jurkowski, A. J., Lake, J. I., Malapani, C., & Meck, W. H. (2015). Bayesian models of interval timing and distortions in temporal memory as a function of Parkinson’s disease and dopamine-related error processing. In *Time Distortions in Mind* (pp. 281-327). Brill.

Halpern, A. R., & Darwin, C. J. (1982). Duration discrimination in a series of rhythmic events. Perception & Psychophysics, 31(1), 86-89.

Hancock, P. A. & Rausch, R. (2010). The effect of age and sex on the perception of time in life. The American Journal of Psychology, 133(1), 170-179.

Handel, S., & Buffardi, L. (1968). Pattern perception: Integrating information presented in two modalities. Science, 162(3857), 1026-1028.

Hanson, J. V., Heron, J., & Whitaker, D. (2008). Recalibration of perceived time across sensory modalities. Experimental Brain Research, 185(2), 347-352.

Hartcher-O’Brien, J., Brighouse, C., & Levitan, C. A. (2016). A single mechanism account of duration and rate processing via the pacemaker–accumulator and beat frequency models. Current Opinion in Behavioral Sciences, 8, 268-275.

Hartcher-O'Brien, J., Di Luca, M., & Ernst, M. O. (2014). The duration of uncertain times: Audiovisual information about intervals is integrated in a statistically optimal fashion. PLoS ONE, 9(3), e89339.

Hass, J., Blaschke, S., & Herrmann, J. M. (2012). Cross-modal distortion of time perception: Demerging the effects of observed and performed motion. PLoS ONE, 7(6), e38092.

Hasuo, E., Nakajima, Y., Tomimatsu, E., Grondin, S., & Ueda, K. (2014). The occurrence of the filled duration illusion: a comparison of the method of adjustment with the method of magnitude estimation. Acta Psychologica, 147, 111-121.

Hearing Haven. Key Structures of the outer, middle and inner ear. Retrieved 21 September 2016, from http://www.hearinghaven.com/how-do-we-hear/ear-diagram/

Hayashi, M. J., Ditye, T., Harada, T., Hashiguchi, M., Sadato, N., Carlson, S., Walsh, V. & Kanai, R. (2015). Time adaptation shows duration selectivity in the human parietal cortex. PLOS Biology, 13(9), e1002262.

Hejtmancik, J. F., & Nickerson, J. M. (2015). Chapter one - Overview of the visual system. In J. F. Hejtmancik & J. M. Nickerson (Eds.), Progress in Molecular Biology and Translational Science (Vol. 134, pp. 1-4): Academic Press.

Hellstrom, A. (2003). Comparison is not just subtraction: Effects of time- and space-order on subjective stimulus difference. Perception & Psychophysics, 65(7), 1161-1177.

Hellstrom, A., & Rammsayer, T. H. (2004). Effects of time-order, interstimulus interval, and feedback in duration discrimination of noise bursts in the 50- and 1000-ms ranges. Acta Psychologica, 116(1), 1-20.

Hellstrom, A., & Rammsayer, T. H. (2015). Time-order errors and standard-position effects in duration discrimination: An experimental study and an analysis by the sensation-weighting model. Attention, Perception & Psychophysics, 77(7), 2409-2423.

Heron, J. (2006). Audiovisual Perceptual Interaction in Humans. (Doctor of Philosophy), University of Bradford, Bradford.

Heron, J., Aaen-Stockdale, C., Hotchkiss, J., Roach, N. W., McGraw, P. V., & Whitaker, D. (2012). Duration channels mediate human time perception. Proceedings of the Royal Society of London B: Biological Sciences, 279(1729), 690-698.

Heron, J., Hanson, J. V., & Whitaker, D. (2009). Effect before cause: Supramodal recalibration of sensorimotor timing. PLoS ONE, 4(11), e7681.

Heron, J., Hotchkiss, J., Aaen-Stockdale, C., Roach, N. W., & Whitaker, D. (2013). A neural hierarchy for illusions of time: Duration adaptation precedes multisensory integration. Journal of Vision, 13(14), 4.

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. Journal of Experimental Psychology, 63(3), 289-293.

Herzog, M. H., Kammer, T., & Scharnowski, F. (2016). Time slices: What is the duration of a percept? PLOS Biology, 14(4), e1002433.

Hickok, G., Rogalsky, C., Matchin, W., Basilakos, A., Cai, J., Pillay, S., Ferrill, M., Mickelson, S., Anderson, S. W., Love, T., Binder, J., & Fridriksson, J. (2018). Neural networks supporting audiovisual integration for speech: A large-scale lesion study. Cortex, 103, 360-371.

Hickok, G., Rogalsky, C., Matchin, W., Basilakos, A., Cai, J., Pillay, S., Ferrill, M., Mickelsen, S., Anderson, S. W., Love, T., Binder, J. & Fridriksson, J. (2018). Neural networks supporting audiovisual integration for speech: A large-scale lesion study. Cortex, 103, 360-371.

Hirsh, I. J. (1959). Auditory perception of temporal order. Journal of the Acoustical Society of America, 31(6), 759-767.

Hirsh, I. J., & Sherrick Jr, C. E. (1961). Perceived order in different sense modalities. Journal of Experimental Psychology, 62(5), 423-432.

Hirsh, I. J., Monahan, C. B., Grant, K. W., & Singh, P. G. (1990). Studies in auditory timing: 1. Simple patterns. Perception & Psychophysics, 47(3), 215-226.

Hocherman, S., & Ben-Dov, G. (1979). Modality-specific effects on discrimination of short empty time intervals. Perceptual & Motor Skills, 48(3), 1, 807-814.

Hofstetter, H. W., Griffin, J. R., & Cline, D. (1997). Dictionary of Visual Science (4th Edition ed.). Boston: Butterworth-Heinemann.

Horr, N. K., & Di Luca, M. (2015). Taking a long look at isochrony: Perceived duration increases with temporal, but not stimulus regularity. Attention, Perception & Psychophysics, 77(2), 592-602.

Horr, N. K., & Di Luca, M. (2015). Timing rhythms: Perceived duration increases with a predictable temporal structure of short interval fillers. PLoS ONE, 10(10), e0141018.

Horr, N. K., & Di Luca, M. (2015a). Filling the blanks in temporal intervals: the type of filling influences perceived duration and discrimination performance. Frontiers in Psychology, 6, 114.

Horr, N. K., & Di Luca, M. (2015b). Taking a long look at isochrony: Perceived duration increases with temporal, but not stimulus regularity. Attention, Perception & Psychophysics, 77(2), 592-602.

Horr, N. K., Wimber, M., & Di Luca, M. (2016). Perceived time and temporal structure: Neural entrainment to isochronous stimulation increases duration estimates. Neuroimage, 132, 148-156.

Horvath, R. A., Schwarcz, A., Aradi, M., Auer, T., Feher, N., Kovacs, N., Tenyi, T., Szalay, C., Perlaki, G., Orsi, G., Komoly, S., Doczi, T., Woermann, F. G., Gyimesi, C & Janszky, J. (2011). Lateralisation of non-metric rhythm. Laterality, 16(5), 620-635.

Horváth, J., Czigler, I., Winkler, I., & Teder-Sälejärvi, W. A. (2007). The temporal window of integration in elderly and young adults. Neurobiology of aging, 28(6), 964-975.

İlhan, B., & VanRullen, R. (2012). No counterpart of visual perceptual echoes in the auditory system. PLoS ONE, 7(11), e49287.

Indraccolo, A., Spence, C., Vatakis, A., & Harrar, V. (2016). Combined effects of motor response, sensory modality, and stimulus intensity on temporal reproduction. Experimental Brain Research, 234(5), 1189-1198.

Ivry, R. B., & Hazeltine, R. E. (1995). Perception and production of temporal intervals across a range of durations: evidence for a common timing mechanism. Journal of Experimental Psychology: Human Perception & Performance, 21(1), 3-18.

Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. Journal of Cognitive Neuroscience, 1(2), 136-152.

Ivry, R. B., & Richardson, T. C. (2002). Temporal control and coordination: The multiple timer model. Brain and Cognition, 48(1), 117-132.

Grahn, J. A. & Rowe, J. B. (2012). Finding and feeling the musical beat: striatal dissociations between detection and prediction of regularity. Cerebral Cortex, 23, 913-921.

Grahn, J. A. & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. Neuroimage, 47, 1894-1903.

Grahn, J. A. & Brett, M. (2009). Impairment of beat-based rhythm discrimination in Parkinson's disease. Cortex, 45, 54-61.

Jaekl, P., Seidlitz, J., Harris, L. R., & Tadin, D. (2015). Audiovisual delay as a novel cue to visual distance. PLoS ONE, 10(10), e0141125.

Jin, D. Z., Fujii, N., & Graybiel, A. M. (2009). Neural representation of time in cortico-basal ganglia circuits. Proceedings of the National Academy of Sciences, 106(45), 19156-19161.

Johnson, K. O., & Phillips, J. R. (1981). Tactile spatial resolution. I. Two-point discrimination, gap detection, grating resolution, and letter recognition. Journal of Neurophysiology, 46(6), 1177-1192.

Johnston, A., Arnold, D. H., & Nishida, S. (2006). Spatially localized distortions of event time. Current Biology, 16(5), 472-479.

Jokiniemi, M., Raisamo, R., Lylykangas, J., & Surakka, V. (2008). Crossmodal rhythm perception. Paper presented at the Proceedings of the 3rd international workshop on Haptic and Audio Interaction Design, Jyvaskyla, Finland.

Jones, A. (2015). Independent effects of bottom-up temporal expectancy and top-down spatial attention. An audiovisual study using rhythmic cueing. Frontiers in Integrative Neuroscience, 8, 96.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. Psychological Review, 83(5), 323-355.

Kalloniatis, M., & Luu, C. (2011). Psychophysics of vision. The organisation of the retina and visual system. Retrieved 11/03/2019, 2019, from https://webvision.med.utah.edu/book/part-viii-gabac-receptors/psychophysics-of-vision/

Kanai, R., Paffen, C. L. E., Hogendoorn, H., & Verstraten, F. A. J. (2006). Time dilation in dynamic visual display. Journal of Vision, 6(12), 1421-1430.

Karmarkar, U. R., & Buonomano, D. V. (2003). Temporal specificity of perceptual learning in an auditory discrimination task. Learning & Memory, 10(2), 141-147.

Kawashima, S., Ueki, Y., Kato, T., Matsukawa, N., Mima, T., Hallett, M., Ito, K & Ojika, K. (2012). Changes in striatal dopamine release associated with human motor-skill acquisition. PloS one, 7(2), e31728.

Kaya, U., Yildirim, F. Z., & Kafaligonul, H. (2017). The involvement of centralized and distributed processes in sub-second time interval adaptation: An ERP investigation of apparent motion. European Journal of Neuroscience, 46(8), 2325-2338.

Keele, S. W., Nicoletti, R., Ivry, R. I., & Pokorny, R. A. (1989). Mechanisms of perceptual timing: Beat-based or interval-based judgements? Psychological Research, 50(4), 251-256.

Keele, S. W., Pokorny, R. A., Corcos, D. M., & Ivry, R. (1985). Do perception and motor production share common timing mechanisms: A correlational analysis. Acta Psychologica, 60(2), 173-191.

Keetels, M., & Vroomen, J. (2012). Perception of synchrony between the senses. Frontiers in the Neural Basis of Multisensory Processes, 147-177.

Kerlinger, F. N. (1999). Foundations of behavioral research. Belmont: Wadsworth Publishing.

Kirman, J. H. (1973). Tactile communication of speech: A review and an analysis. Psychological Bulletin, 80(1), 54-74.

Klink, P. C., Montijn, J. S., & van Wezel, R. J. A. (2011). Crossmodal duration perception involves perceptual grouping, temporal ventriloquism, and variable internal clock rates. Attention, Perception, & Psychophysics, 73(1), 219-236.

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The evolution of rhythm processing. Trends in Cognitive Sciences, 22(10), 896-910.

Kraemer, P. J., Randall, C. K., & Brown, R. W. (1997). The influence of stimulus attributes on duration matching-to-sample in pigeons. Animal Learning & Behavior, 25(2), 148-157.

Krampe, R. T., Engbert, R., & Kliegl, R. (2002). The effects of expertise and age on rhythm production: Adaptations to timing and sequencing constraints. Brain and Cognition, 48(1), 179-194.

Kuhn, T. L. (1974). Discrimination of modulated beat tempo by professional musicians. Journal of Research in Music Education, 22(4), 270-277.

Kuroda, T., Hasuo, E., Labonte, K., Laflamme, V., & Grondin, S. (2014). Discrimination of two neighboring intra- and intermodal empty time intervals marked by three successive stimuli. Acta Psychologica, 149, 134-141.

Landy, M. (2006). Low and high spatial frequency sine wave gratings. Retrieved 19/03/2019 from http://www.cns.nyu.edu/~david/courses/perception/lecturenotes/channels/channels.html

Lapid, E., Ulrich, R., & Rammsayer, T. (2008). On estimating the difference limen in duration discrimination tasks: a comparison of the 2AFC and the reminder task. Perception & Psychophysics, 70(2), 291-305.

Larsson, J. Psychophysics, psychophysical methods, and signal detection theory. Retrieved 04/03/2019, from https://slideplayer.com/slide/4573100/15/images/12/Stimulus+no+longer+detected+Threshold%3A+average+stimulus+intensity.jpg

Lawrance, E. L. A., Harper, N. S., Cooke, J. E., & Schnupp, J. W. H. (2014). Temporal predictability enhances auditory detection. Journal of the Acoustical Society of America, 135(6), EL357-EL363.

Lerens, E., Araneda, R., Renier, L., & De Volder, A. G. (2014). Improved beat asynchrony detection in early blind individuals. Perception, 43(10), 1083-1096.

Levitan, C. A., Ban, Y. H. A., Stiles, N. R. B., & Shimojo, S. (2015). Rate perception adapts across the senses: Evidence for a unified timing mechanism. Scientific Reports, 5, 6, 8857.

Lewis, P. A., & Miall, R. C. (2003). Distinct systems for automatic and cognitively controlled time measurement: Evidence from neuroimaging. Current Opinion in Neurobiology, 13, 250-255.

Lewis, P. A., & Miall, R. C. (2006). Remembering the time: A continuous clock. Trends in Cognitive Sciences, 10(9), 401-406.

Lewkowicz, D. J. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. Psychological Bulletin, 126(2), 281-308.

Allan, L. G. (1979). Perception of time. Perception & Psychophysics, 26, 340-354.

Li, B. L., Yuan, X. Y., & Huang, X. T. (2015). The aftereffect of perceived duration is contingent on auditory frequency but not visual orientation. Scientific Reports, 5, 9.

Li, B., Chen, Y., Xiao, L., Liu, P., & Huang, X. (2017). Duration adaptation modulates EEG correlates of subsequent temporal encoding. Neuroimage, 147, 143-151.

Li, B., Xiao, L., Yin, H., Liu, P., & Huang, X. (2017). Duration aftereffect depends on the duration of adaptation. Frontiers in Psychology, 8, 491.

Li, B., Yuan, X., Chen, Y., Liu, P., & Huang, X. (2015). Visual duration aftereffect is position invariant. Frontiers in Psychology, 6, 1536.

Li, M. S., & Di Luca, M. (2018). Musical scales in tone sequences improve temporal accuracy. Frontiers in Psychology, 9, 105.

Locke, S. M., & Landy, M. S. (2017). Temporal causal inference with stochastic audiovisual sequences. PLoS ONE, 12(9), e0183776.

Maarseveen, J., Hogendoorn, H., Verstraten, F. A. J., & Paffen, C. L. E. (2018). Attention gates the selective encoding of duration. Scientific Reports, 8(1), 2522.

Maarseveen, J., Hogendoorn, H., Verstraten, F. A. J., & Paffen, C. L. E. (2017). An investigation of the spatial selectivity of the duration after-effect. Vision Research, 130, 67-75.

Maarseveen, J., Paffen, C. L. E., Verstraten, F. A. J., & Hogendoorn, H. (2019). The duration aftereffect does not reflect adaptation to perceived duration. PLoS ONE, 14(3), e0213163.

Macinnis, M. L. (2007). Do rats time filled and empty intervals of equal duration differently? Behavioural Processes, 75(2), 182-187.

Madison, G. (2004). Detection of linear temporal drift in sound sequences: Empirical data and modelling principles. Acta Psychologica, 117(1), 95-118.

Madison, G., & Merker, B. (2004). Human sensorimotor tracking of continuous subliminal deviations from isochrony. Neuroscience Letters, 370(1), 69-73.

Malapani, C., Rakitin, B., Levy, R., Meck, W. H., Deweer, B., Dubois, B., & Gibbon, J. (1998). Coupled temporal memories in Parkinson's disease: a dopamine-related dysfunction. Journal of Cognitive Neuroscience, 10(3), 316-331.

Marchant, J. L., & Driver, J. (2013). Visual and audiovisual effects of isochronous timing on visual perception and brain activity. Cerebral Cortex, 23(6), 1290-1298.

Mattes, S., & Ulrich, R. (1998). Directed attention prolongs the perceived duration of a brief stimulus. Perception & Psychophysics, 60(8), 1305-1317.

Matthews, W. J. (2013). How does sequence structure affect the judgment of time? Exploring a weighted sum of segments model. Cognitive Psychology, 66(3), 259-282.

Mauk, M. D., & Buonomano, D. V. (2004). The neural basis of temporal processing. Annual Review of Neuroscience, 27, 307-340.

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

McAuley, J. D., & Henry, M. J. (2010). Modality effects in rhythm processing: Auditory encoding of visual rhythms is neither obligatory nor automatic. Attention, Perception & Psychophysics, 72, 1377-1389.

McAuley, J. D., & Kidd, G. R. (1998). Effect of deviations from temporal expectations on tempo discrimination of isochronous tone sequences. Journal of Experimental Psychology: Human Perception & Performance, 24(6), 1786-1800.

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Lifespan development of timing and event tracking. Journal of Experimental Psychology: General, 135, 348-367.

Meck, W. H. (1996). Neuropharmacology of timing and time perception. Cognitive Brain Research, 3, 227-233.

Meck, W. H. (2005). Neuropsychology of timing and time perception. Brain & Cognition, 58, 1-8.

Meck, W. H. (2006). Frontal cortex lesions eliminate the clock speed effect of dopaminergic drugs on interval timing. Brain research, 1108(1), 157-167.

Merchant, H., Perez, O., Bartolo, R., Mendez, J. C., Mendoza, G., Gamez, J., Yc, K. & Prado, L. (2015). Sensorimotor neural dynamics during isochronous tapping in the medial premotor cortex of the macaque. European Journal of Neuroscience, 41(5), 586-602.

Michon, J. A. (1964). Studies on subjective duration: I. Differential sensitivity in the perception of repeated temporal intervals. Acta Psychologica, 22, 441-450.

Miller, L. K., & Eargle, A. (1990). The contributions of development versus music training to simple tempo discrimination. Journal of Research in Music Education, 38(4), 294-301.

Miller, N. S., & McAuley, J. D. (2005). Tempo sensitivity in isochronous tone sequences: The multiple-look model revisited. Perception & Psychophysics, 67(7), 1150-1160.

Mioni, G., Grassi, M., Tarantino, V., Stablum, F., Grondin, S., & Bisiacchi, P. S. (2016). The impact of a concurrent motor task on auditory and visual temporal discrimination tasks. Attention, Perception & Psychophysics, 78(3), 742-748.

Mioni, G., Stablum, F., McClintock, S. M., & Grondin, S. (2014a). Different methods for reproducing time, different results. Attention, Perception & Psychophysics, 76(3), 675-681.

Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. Behavioural Brain Research, 6(1), 57-77.

Monahan, C. B., & Hirsh, I. J. (1990). Studies in auditory timing: 2. Rhythm patterns. Perception & Psychophysics, 47(3), 227-242.

Motala, A., Heron, J., McGraw, P. V., Roach, N. W., & Whitaker, D. (2018). Rate after-effects fail to transfer cross-modally: Evidence for distributed sensory timing mechanisms. Scientific Reports, 8(1), 924.

Nagarajan, S. S., Blake, D. T., Wright, B. A., Byl, N., & Merzenich, M. M. (1998). Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality. Journal of Neuroscience, 18(4), 1559-1570.

Nemes, V., Whitaker, D., Heron, J., & McKeefry, D. (2011). Multiple spatial frequency channels in human visual perceptual memory. Vision research, 51(23), 2331-2339.

Nettl, B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. In M. B. Wallin NL, Brown S, editors (Ed.), The Origins of Music (pp. 463-472.). Cambridge: The MIT Press.

Nichelli, P. (1996). Time perception measurements in neuropsychology. Advances in Psychology, 115, 187-204.

Nidiffer, A. R., Diederich, A., Ramachandran, R., & Wallace, M. T. (2018). Multisensory perception reflects individual differences in processing temporal correlations. Scientific Reports, 8(1), 14483.

Nieto, M. P. (2015). Human Visual Pathway. Retrieved 8th November, 2018, from https://en.wikipedia.org/wiki/File:Human_visual_pathway.svg

Noulhiane, M., Pouthas, V., & Samson, S. (2009). Is time reproduction sensitive to sensory modalities? European Journal of Cognitive Psychology, 21(1), 18-34.

O'Boyle, D. J., Freeman, J. S., & Cody, F. W. (1996). The accuracy and precision of timing of self-paced, repetitive movements in subjects with Parkinson's disease. Brain, 119(1), 51-70.

Occelli, V., Spence, C., & Zampini, M. (2011). Audiotactile interactions in temporal perception. Psychonomic Bulletin & Review, 18(3), 429-454.

Ono, F., & Kitazawa, S. (2009). The effect of marker size on the perception of an empty interval. Psychonomic Bulletin & Review, 16(1), 182-189.

Ono, F., & Kitazawa, S. (2011). Shortening of subjective visual intervals followed by repetitive stimulation. PLoS ONE, 6(12), e28722.

Osterberg, G. (1935). Topography of the layer of rods and cones in the human retina. Acta Opthamologica Supplement, 6, 1-103.

Lewis, P. A. & Miall, R. C. (2003). Distinct systems for automatic and cognitively controlled time measurement: evidence from neuroimaging. Current Opinion in Neurobiology, 13, 250-255.

Pápai, M. S., & Soto-Faraco, S. (2017). Sounds can boost the awareness of visual events through attention without cross-modal integration. Scientific Reports, 7, 41684.

Pashler, H. (2001b). Perception and production of brief durations: Beat-based versus interval-based timing. Journal of Experimental Psychology: Human Perception & Performance, 27(2), 485-493.

Pasinski, A. C., McAuley, J. D., & Snyder, J. S. (2016). How modality specific is processing of auditory and visual rhythms? Psychophysiology, 53(2), 198-208.

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. Experimental Brain Research, 163(2), 226-238.

Penfield, W., & Rasmussen, T. (1950). The cerebral cortex of man; A clinical study of localization of function. Oxford, England: Macmillan.

Pinel, J. P. J. (2006). Basics of Biopsychology: Pearson Allyn and Bacon.

Plant, R. R., & Quinlan, P. T. (2013). Could millisecond timing errors in commonly used equipment be a cause of replication failure in some neuroscience studies? Cognitive, Affective & Behavioural Neuroscience, 13(3), 598-614.

Provasi, J., Anderson, D. I., & Barbu-Roth, M. (2014). Rhythm perception, production, and synchronization during the perinatal period. Frontiers in Psychology, 5, 1048.

Radeau, M. (1994). Auditory-visual spatial interaction and modularity. Current Psychology of Cognition, 13(1), 3-51.

Radeau, M., & Bertelson, P. (1977). Adaptation to auditory-visual discordance and ventriloquism in semirealistic situations. Perception & Psychophysics, 22(2), 137-146.

Rammsayer, T. H. (1999). Neuropharmacological evidence for different timing mechanisms in humans. Quarterly Journal of Experimental Psychology: Section B, 52(3), 273-286.

Rammsayer, T. H. (2010). Differences in duration discrimination of filled and empty auditory intervals as a function of base duration. Attention, Perception, & Psychophysics, 72(6), 1591-1600.

Rammsayer, T. H. (2014). The effects of type of interval, sensory modality, base duration, and psychophysical task on the discrimination of brief time intervals. Attention, Perception & Psychophysics, 76(4), 1185-1196.

Rammsayer, T. H., & Brandler, S. (2004). Aspects of temporal information processing: A dimensional analysis. Psychological research, 69(1), 115-123.

Rammsayer, T. H., Lima, S. D., & Vogel, W. H. (1993). Aging and temporal discrimination of brief auditory intervals. Psychological Research, 55(1), 15-19.

Rammsayer, T. H., & Lima, S. D. (1991). Duration discrimination of filled and empty auditory intervals: Cognitive and perceptual factors. Perception & Psychophysics, 50(6), 565-574.

Rammsayer, T. H., & Troche, S. J. (2014). Elucidating the internal structure of psychophysical timing performance in the sub-second and second range by utilizing confirmatory factor analysis. Advances in Experimental Medicine & Biology, 829, 33-47.

Rammsayer, T. H., Borter, N., & Troche, S. J. (2015). Visual-auditory differences in duration discrimination of intervals in the subsecond and second range. Frontiers in Psychology, 6, 1626.

Rammsayer, T., & Pichelmann, S. (2018). Visual-auditory differences in duration discrimination depend on modality-specific, sensory-automatic temporal processing: Converging evidence for the validity of the Sensory-Automatic Timing Hypothesis. Quarterly Journal of Experimental Psychology, 71(11), 2364-2377.

Ravignani, A., Honing, H., & Kotz, S. A. (2017). Editorial: The evolution of rhythm cognition: Timing in music and speech. Frontiers in Human Neuroscience, 11, 303.

Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. Journal of Neurophysiology, 89(2), 1078-1093.

Repp, B. H., & Marcus, R. J. (2010). No sustained sound illusion in rhythmic sequences. Music Perception: An Interdisciplinary Journal, 28(2), 121-134.

Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. Journal of Experimental Psychology-Human Perception and Performance, 28(5), 1085-1099.

Rhodes, D., & Di Luca, M. (2016). Temporal regularity of the environment drives time perception. PLoS ONE, 11(7), e0159842.

Ridgway, N., Milders, M., & Sahraie, A. (2008). Redundant target effect and the processing of colour and luminance. Experimental Brain Research, 187(1), 153-160.

Robinson, D. W., & Sutton, G. J. (1979). Age effect in hearing - a comparative analysis of published threshold data. Audiology, 18(4), 320-334.

Romei, V., De Haas, B., Mok, R. M., & Driver, J. (2011). Auditory stimulus timing influences perceived duration of co-occurring visual stimuli. Frontiers in Psychology, 2, 215.

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Tamada, T., Iwata, N. K., & Nielsen, M. (1999). Neural representation of a rhythm depends on its interval ratio. Journal of Neuroscience, 19(22), 10074-10081.

Samson, S., Ehrle, N., & Baulac, M. (2001). Cerebral substrates for musical temporal processes. Annals of the New York Academy of Sciences, 930, 166-178.

Santi, A., Miki, A., Hornyak, S., & Eidse, J. (2006). The perception of empty and filled time intervals by rats. Behavioural Processes, 71(2-3), 144-156.

Schubotz, R. I., Friederici, A. D., & Von Cramon, D. Y. (2000). Time perception and motor timing: a common cortical and subcortical basis revealed by fMRI. Neuroimage, 11(1), 1-12.

Schulze, H. H. (1989). The perception of temporal deviations in isochronic patterns. Perception & Psychophysics, 45(4), 291-296.

Shalom, S., & Zaidel, A. (2018). Better than Optimal. Neuron, 97(3), 484-487.

Shi, Z., & Burr, D. (2016). Predictive coding of multisensory timing. Current Opinion in Behavioral Sciences, 8, 200-206.

Shima, S., Murai, Y., Yuasa, K., Hashimoto, Y., & Yotsumoto, Y. (2018). Periodic Fluctuation of Perceived Duration. i-Perception, 9(2), 2041669518760625.

Shinn-Cunningham, B., Varghese, L., Wang, L., & Bharadwaj, H. (2017). Individual differences in temporal perception and their implications for everyday listening. In N. Kraus, S. Anderson, T. White-Schwoch, R. R. Fay, & A. N. Popper (Eds.), The frequency-following response: A window into human communication (pp. 159-192). Cham: Springer International Publishing.

Shipley, T. (1964). Auditory flutter-driving of visual flicker. Science, 145(3638), 1328-1330.

Skinner, B. (1938). The behavior of organisms: An experimental analysis. Appleton-Century. New York.

Smith, P. L., & Little, D. R. (2018). Small is beautiful: In defense of the small-N design. Psychonomic Bulletin & Review, 6, 2083-2101.

Sorkin, R. D. (1990). Perception of temporal patterns defined by tonal sequences. Journal of the Acoustical Society of America, 87(4), 1695-1701.

Sorkin, R. D., Boggs, G. J., & Brady, S. L. (1982). Discrimination of temporal jitter in patterned sequences of tones. Journal of Experimental Psychology: Human Perception and Performance, 8(1), 46-57.

Soto-Faraco, S., & Deco, G. (2009). Multisensory contributions to the perception of vibrotactile events. Behavioural Brain Research, 196(2), 145-154.

Spence, C. (2007). Audiovisual multisensory integration. Acoustical Science and Technology, 28(2), 61-70.

Spence, C., & Driver, J. (1997). Cross-modal links in attention between audition, vision, and touch: Implications for interface design. International Journal of Cognitive Ergonomics, 1(4), 351-373.

Spence, C., & Squire, S. (2003). Multisensory integration: Maintaining the perception of synchrony. Current Biology, 13(13), R519-521.

Spence, C., Nicholls, M. E. R., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. Perception & Psychophysics, 63(2), 330-336.

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. Journal of Experimental Psychology: General, 130(4), 799-832.

Stauffer, C. C., Haldemann, J., Troche, S. J., & Rammsayer, T. H. (2012). Auditory and visual temporal sensitivity: Evidence for a hierarchical structure of modality-specific and modality-independent levels of temporal information processing. Psychological research, 76(1), 20-31.

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

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

Su, Y.H. (2016). Visual enhancement of illusory phenomenal accents in non-isochronous auditory rhythms. PLoS ONE, 11(11), e0166880.

Sugano, Y., Keetels, M., & Vroomen, J. (2017). Audio-motor but not visuo-motor temporal recalibration speeds up sensory processing. PLoS ONE, 12(12), e0189242.

Tal, I., Large, E. W., Rabinovitch, E., Wei, Y., Schroeder, C. E., Poeppel, D., & Zion Golumbic, E. (2017). Neural entrainment to the beat: the “missing-pulse” phenomenon. Journal of Neuroscience, 37(26), 6331-6341.

Tanaka, M., Tsuzaki, M., & Kato, H. (1994). Discrimination of empty duration in the click sequence simulating a mora structure. Journal of the Acoustical Society of Japan (E), 15(3), 191-192.

Teki, S., Grube, M., & Griffiths, T. (2012). A unified model of time perception accounts for duration-based and beat-based timing mechanisms. Frontiers in Integrative Neuroscience, 5, 90.

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. Journal of Neuroscience, 31, 3805-3812.

ten Hoopen, G., Boelaarts, L., Gruisen, A., Apon, I., Donders, K., Mul, N., & Akerboom, S. (1994). The detection of anisochrony in monaural and interaural sound sequences. Perception & Psychophysics, 56(1), 110-120.

ten Oever, S., Schroeder, C. E., Poeppel, D., van Atteveldt, N., & Zion-Golumbic, E. (2014). Rhythmicity and cross-modal temporal cues facilitate detection. Neuropsychologia, 63, 43-50.

Tomassini, A., Vercillo, T., Torricelli, F., & Morrone, M. C. (2018). Rhythmic motor behaviour influences perception of visual time. Proceedings of the Royal Society B: Biological Sciences, 285.

Tong, J., Mao, O., & Goldreich, D. (2013). Two-point orientation discrimination versus the traditional two-point test for tactile spatial acuity assessment. Frontiers in Human Neuroscience, 7, 579.

Treisman, M., Faulkner, A., Naish, P. L. N., & Brogan, D. (1990). The internal clock: Evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. Perception, 19, 705-743.

Triplett, D. (1931). The relation between the physical pattern and the reproduction of short temporal intervals: A study in the perception of filled and unfilled time. Psychological Monographs, 41(4), 201-265.

Tse, C.-Y., & Penney, T. B. (2006). Preattentive timing of empty intervals is from marker offset to onset. Psychophysiology, 43(2), 172-179.

Ulbrich, P., Churan, J., Fink, M., & Wittmann, M. (2007). Temporal reproduction: Further evidence for two processes. Acta Psychologica, 125(1), 51-65.

van Atteveldt, N., Musacchia, G., Zion-Golumbic, E., Sehatpour, P., Javitt, D. C., & Schroeder, C. (2015). Complementary fMRI and EEG evidence for more efficient neural processing of rhythmic vs. unpredictably timed sounds. Frontiers in Psychology, 6, 1663.

van Boven, R. W., & Johnson, K. O. (1994). A psychophysical study of the mechanisms of sensory recovery following nerve injury in humans. Brain, 117(1), 149-167.

van Eijk, R. L., Kohlrausch, A., Juola, J. F., & van de Par, S. (2009). Temporal interval discrimination thresholds depend on perceived synchrony for audio-visual stimulus pairs. Journal of Experimental Psychology: Human Perception & Performance, 35(4), 1254-1263.

van Erp, J. B., & Werkhoven, P. J. (2004). Vibro-tactile and visual asynchronies: Sensitivity and consistency. Perception, 33(1), 103-111.

van Nes, S. I., Faber, C. G., Hamers, R. M., Harschnitz, O., Bakkers, M., Hermans, M. C., Meijer, R. J., van Doorn, P. A. & Merkies, I. S. (2008). Revising two-point discrimination assessment in normal aging and in patients with polyneuropathies. Journal of Neurology, Neurosurgergy & Psychiatry, 79(7), 832-834.

Vatakis, A., & Spence, C. (2007). Crossmodal binding: Evaluating the “unity assumption” using audiovisual speech stimuli. Perception & Psychophysics, 69(5), 744-756.

Vatakis, A., & Spence, C. (2008). Evaluating the influence of the ‘unity assumption’ on the temporal perception of realistic audiovisual stimuli. Acta Psychologica, 127(1), 12-23.

Vatakis, A., & Spence, C. (2010). Audiovisual temporal integration for complex speech, object-action, animal call, and musical stimuli. In M. J. Naumer & J. Kaiser (Eds.), Multisensory object perception in the primate brain (pp. 95-121). New York, NY, US: Springer Science + Business Media.

Vega-Bermudez, F., & Johnson, K. O. (2001). Differences in spatial acuity between digits. Neurology, 56(10), 1389-1391.

Villanueva, L., & Zampini, M. (2018). Reciprocal interference between audition and touch in the perception of duration. Multisensory Research, 31(5), 351-371.

Villanueva, L., & Zampini, M. (2018). Reciprocal interference between audition and touch in the perception of duration. Multisensory Research, 31(5), 351-371.

Vroomen, J., & Keetels, M. (2006). The spatial constraint in intersensory pairing: No role in temporal ventriloquism. Journal of Experimental Psychology: Human Perception & Performance, 32(4), 1063-1071.

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. Attention, Perception & Psychophysics, 72(4), 871-884.

Walker, J. T., & Scott, K. J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones and gaps. Journal of Experimental Psychology: Human Perception & Performance, 7(6), 1327-1339.

Walker, J. T., Irion, A. L., & Gordon, D. G. (1981). Simple and contingent aftereffects of perceived duration in vision and audition. Perception & Psychophysics, 29(5), 475-486.

Watanabe, J., Amemiya, T., Nishida, S., & Johnston, A. (2010). Tactile duration compression by vibrotactile adaptation. Neuroreport, 21(13), 856-860.

Watson, C. S., Foyle, D. C., & Kidd, G. R. (1990). Limits of auditory pattern discrimination for patterns with various durations and numbers of components. Journal of the Acoustical Society of America, 88(6), 2631-2638.

Wearden, J. H. (2003). Applying the scalar timing model to human time psychology: Progress and challenges. Time and mind II: Information processing perspectives, 21-39.

Wearden, J. H., & Bray, S. (2001). Scalar timing without reference memory? Episodic temporal generalization and bisection in humans. Quarterly Journal of Experimental Psychology B, 54(4), 289-309.

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. Quarterly Journal of Experimental Psychology: Section B, 51(2), 97-120.

Wearden, J. H., Goodson, G., & Foran, K. (2007). Subjective shortening with filled and unfilled auditory and visual intervals in humans? Quarterly Journal of Experimental Psychology (Hove), 60(12), 1616-1628.

Wearden, J., & Lejeune, H. (1993). Across the great divide: Animal psychology and time in humans. Time & Society, 2(1), 87-106.

Webster, M. A. (2015). Visual adaptation. Annual Review of Vision Science, 1(1), 547-567.

Weisstein, E. W. Least Squares Fitting. Retrieved 19/03/2019 from http://mathworld.wolfram.com/LeastSquaresFitting.html

Welch, R. B. (1999). Meaning, attention, and the “unity assumption” in the intersensory bias of spatial and temporal perceptions. Advances in Psychology, 129, 371-387.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. Psychological Bulletin, 88(3), 638-667.

Welch, R. B., DutionHurt, L. D., & Warren, D. H. (1986). Contributions of audition and vision to temporal rate perception. Perception & Psychophysics, 39(4), 294-300.

Wiener, M., Turkeltaub, P., & Coslett, H. B. (2010). The image of time: A voxel-wise meta-analysis. Neuroimage, 49(2), 1728-1740.

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. Proceedings of the National Academy of Sciences, 106(7), 2468-2471.

Wing, A. M. & Kristofferson, A. B. (1973a). Response delays and the timing of discrete motor responses. Perception & Psychophysics, 14(1), 5-12.

Wing, A. M. & Kristofferson, A. B. (1973b). The timing of interresponse intervals. Perception & Psychophysics, 13(3), 455-460.

Wittmann, M., & Lehnhoff, S. (2005). Age effects in perception of time. Psychological Reports, 97(3), 921-935.

Wittmann, M. (1999). Time perception and temporal processing levels of the brain. Chronobiology International, 16(1), 17-32.

Wright, B. A., Buonomano, D. V., Mahncke, H. W., & Merzenich, M. M. (1997). Learning and generalization of auditory temporal–interval discrimination in humans. Journal of Neuroscience, 17(10), 3956-3963.

Xuan, B., Zhang, D., He, S., & Chen, X. (2007). Larger stimuli are judged to last longer. Journal of Vision, 7(10), 2 1-5.

Yuasa, K., & Yotsumoto, Y. (2015). Opposite Distortions in Interval Timing Perception for Visual and Auditory Stimuli with Temporal Modulations. PLoS ONE, 10(8), e0135646.

Shi, Z., Ganzenmüller, S. & Müller, H. J. (2013). Reducing bias in auditory duration reproduction by integrating the reproduced signal. PLoS ONE, 8, e62065.

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

Zelaznik, H. N., Spencer, R. M., & Doffin, J. G. (2000). Temporal precision in tapping and circle drawing movements at preferred rates is not correlated: Further evidence against timing as a general-purpose ability. Journal of Motor Behaviour, 32(2), 193-199.

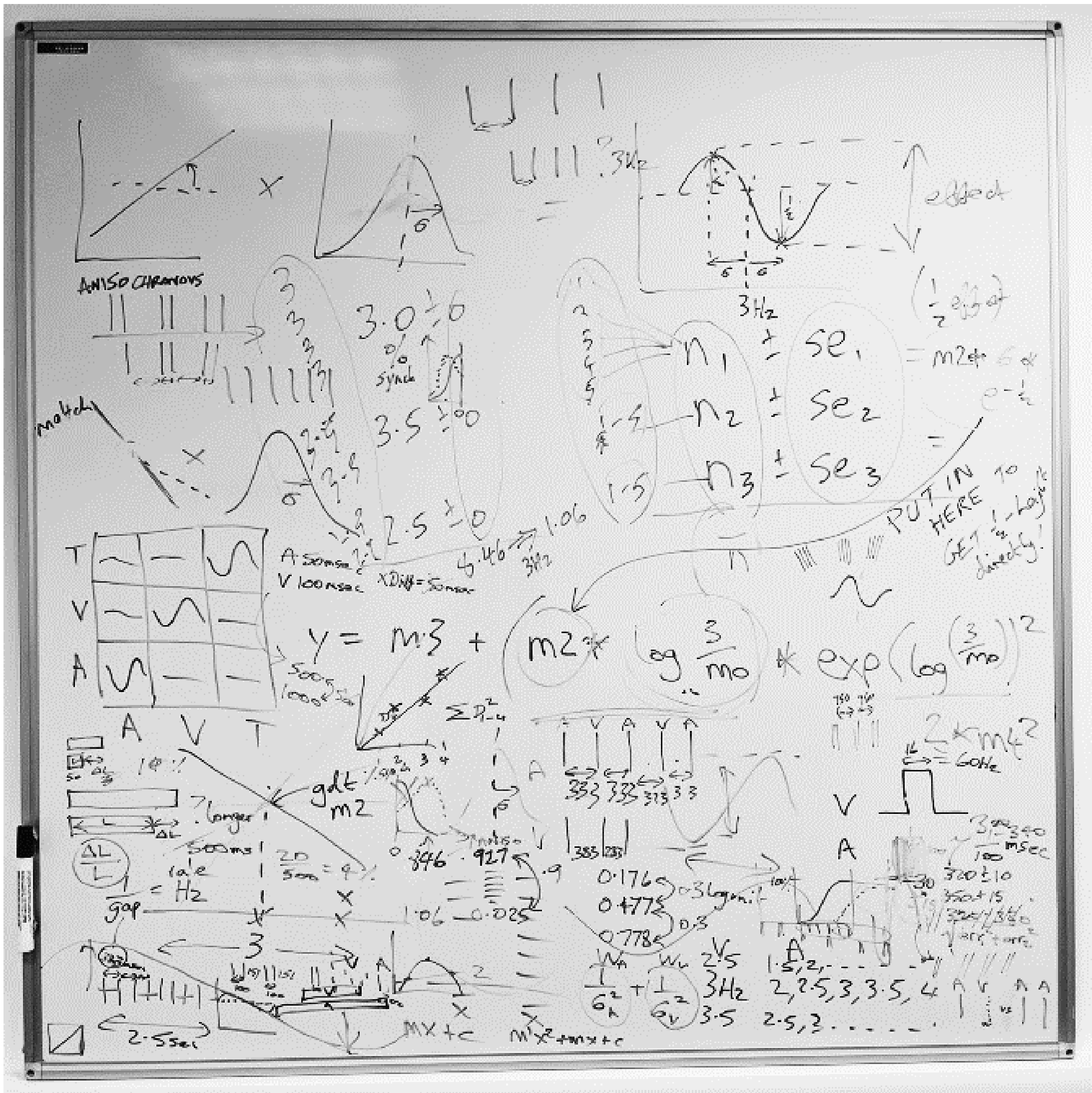
Zelkind, I. (1973). Factors in time estimation and a case for the internal clock. Journal of General Psychology, 88(2), 295-301.

Zeni, S., & Holmes, N. P. (2018). The effect of a regular auditory context on perceived interval duration. Frontiers in Psychology, 9, 1567.

Zhang, H., Chen, L., & Zhou, X. (2012). Adaptation to visual or auditory time intervals modulates the perception of visual apparent motion. Frontiers in integrative neuroscience, 6, 100.

Zoefel, B., & Van Rullen, R. (2017). Oscillatory mechanisms of stimulus processing and selection in the visual and auditory systems: State-of-the-art, speculations and suggestions. Frontiers in Neuroscience, 11, 296.

Appendix A: Lab whiteboard



Appendix B: Adjusted p-values for all subjects across all conditions for Experiment 1 in Chapter 5.

The first letter denotes the adapting modality and the second letter denotes the testing modality, ‘A’ refers to the auditory modality, ‘T’ to the tactile modality and ‘V’ to the visual modality, *(Values marked with an asterisk (*) signify results of statistical significance).*

Condition	**Subject AM**	**Subject DW**	**Subject YL**
AA	.002*	<.001*	<.001*
TT	.010*	<.001*	.002*
VV	<.001*	<.001*	.009*
AT	.096	1	1
TA	1	1	1
AV	.740	1	1
VA	1	1	1
TV	1	1	1
VT	1	1	1

Appendix C: Average amplitudes of effect across all unimodal conditions for each subject for Experiment 1 in Chapter 5.

Respective standard error values are denoted in brackets.

	Amplitude of effect – AM (Hz)	**Amplitude of effect – DW (Hz)**	**Amplitude of effect – YL (Hz)**
AA	0.65 *(.09)*	0.21 *(.03)*	0.35 *(.03)*
TT	0.8 *(.14)*	0.25 *(.03)*	0.25 *(.03)*
VV	0.53 *(.06)*	0.51 *(.04)*	0.26 *(.05)*

Appendix D: Average spread of effect across all unimodal conditions for each subject in Experiment 1 in Chapter 5.

Respective standard error values are denoted in brackets.

	Spread of effect – AM (in log units)	**Spread of effect – DW (in log units)**	**Spread of effect – YL (in log units)**
AA	0.31 *(.06)*	0.18 *(.02)*	0.45 *(.09)*
TT	0.40 *(.15)*	0.25 *(.03)*	0.43 *(.12)*
VV	0.36 *(.08)*	0.21 *(.01)*	0.43 *(.16)*

Appendix E: Holm-Bonferroni-adjusted p-values for all subjects across all conditions for the control experiment in which subjects were unaware of the test modality.

The first letter denotes the adapting modality and the second letter denotes the testing modality, ‘A’ refers to the auditory modality and ‘V’ to the visual modality. Values marked with an asterisk (*) signify results of statistical significance.

Condition	Subject AM	Subject DW	Subject YL
AA	.007*	.014*	<.001*
VV	.020*	.003*	.036*
AV	1	1	.963
VA	1	1	.963

Appendix F: Average amplitudes of effect across both unimodal conditions for each subject from first control experiment in Chapter 5.

Respective standard error values are denoted in brackets.

	Amplitude of effect – AM (Hz)	**Amplitude of effect – DW (Hz)**	**Amplitude of effect – YL (Hz)**
AA	.40 *(.08)*	.16 *(.04)*	.19 *(.01)*
VV	.69 *(.17)*	.29 *(.05)*	.36 *(.10)*

Appendix G: Average spread of effect across both unimodal conditions for each subject from the first control experiment in Chapter 5.

Respective standard error values are denoted in brackets.

	Spread of effect – AM (in log units)	**Spread of effect – DW (in log units)**	**Spread of effect – YL (in log units)**
AA	.24 (.04)	.23 (.05)	.29 (.01)
VV	.39 (.20)	.23 (.03)	.49 (.27)

Appendix H: Table of results from Control Experiment 2 in Chapter 5.

Amplitudes of adaptation effect (μ), spread (σ in log units) of adaptation effect and Holm-Bonferroni-adjusted p-values across all unimodal conditions for each subject for the control experiment in which visual and auditory stimuli were spatially and temporally overlapped. The first letter denotes the adapting modality and the second letter denotes the testing modality, 'A' refers to the auditory modality and 'V' to the visual modality. All cross-modal conditions were found to be not statistically significant (p>0.05).

	Subject AM			**Subject DW**		
	Amplitude (μ)	**p-value**	**Spread (σ)**	**Amplitude (μ)**	**p-value**	**Spread (σ)**
AA	0.50±.04	<0.001	0.29±.03	0.17±.05	0.030	0.23±.05
VV	0.58±.08	<0.001	0.29±.04	0.24±.06	0.022	0.25±.06

Appendix I: Table of results from Experiment 1 in Chapter 6.

Results from experiment 1 for subject AM and DW with 95% confidence intervals; $\bar{x}_1 - \bar{x}_2$ represents the mean difference between data plotted using logistic functions and $\mu_1 - \mu_2$ represents the difference between mean squared error; (A) denotes the auditory modality and (V) denotes the visual modality.

Experimental Condition	$\bar{x}_1 - \bar{x}_2$	95% CI for $\mu_1 - \mu_2$
(AM) Baseline-Adapt 160ms (A)	11.6	(5.2, 18)
(AM) Baseline-Adapt 160ms (V)	23.54	(10.7, 36.4)
(DW) Baseline-Adapt 160ms (A)	16.96	(15.3, 18.6)
(DW) Baseline-Adapt 160ms (V)	30.02	(20.9, 39.2)

Appendix J: Table of results from Experiment 1 in Chapter 7.

95% Confidence interval values. Specifically, the values in the table describe the PSE for each condition for each subject, ± the confidence interval for that condition.

Auditory			
	AM	**DW**	**KH**
Baseline	25.6 ± 0.9	34.7 ± 1.3	57.7 ± 6.9
Adapt 0ms	31.3 ± 8.8	37.1 ± 2.9	40.4 ± 2.9
Adapt 25ms	59.5 ± 11.6	42.6 ± 3.0	62.9 ± 8.3
Adapt 50ms	64.6 ± 7.8	44.5 ± 4.1	72.0 ± 5.6
Adapt 100ms	90.4 ± 10.5	47.3 ± 0.7	70.1 ± 5.1
Tactile			
	AM	**DW**	**KH**
Baseline	64.7 ± 11.0	67.2 ± 10.5	51.1 ± 10.5
Adapt 0ms	70.4 ± 19.3	51.7 ± 5.3	63.5 ± 14.8
Adapt 25ms	68.1 ± 38.2	49.9 ± 8.6	70.3 ± 10.7
Adapt 50ms	86.8 ± 29.5	62.1 ± 13.2	80.4 ± 37.0
Adapt 100ms	111.9 ± 68.8	42.4 ± 55.0	87.2 ± 8.6
Visual			
	AM	**DW**	**KH**
Baseline	65.0 ± 5.0	65.4 ± 6.7	86.3 ± 8.3
Adapt 0ms	57.6 ± 12.9	41.2 ± 9.4	88.1 ± 10.6
Adapt 25ms	61.7 ± 12.7	44.7 ± 4.0	43.4 ± 28.6
Adapt 50ms	67.7 ± 7.0	60.1 ± 9.2	70.3 ± 9.3
Adapt 100ms	80.8 ± 14.3	70.5 ± 4.1	93.2 ± 7.5

Appendix K: Table of results from Experiment 2 in Chapter 7.

95% Confidence interval values. Specifically, the values in the table describe the PSE for each condition for each subject, ± the confidence interval for that condition.

Auditory			
	AM	**DW**	**KH**
Baseline	25.6 ± 0.8	34.7 ± 1.3	57.7 ± 6.9
Adapt 0ms	39.8 ± 12.0	23.8 ± 4.3	40.4 ± 10.5
Adapt 100ms	39.3 ± 6.1	33.6 ± 8.6	54.5 ± 8.6
Tactile			
	AM	**DW**	**KH**
Baseline	64.7 ± 11.0	55.5 ± 5.0	51.1 ± 10.5
Adapt 0ms	88.5 ± 14.3	48.1 ± 7.7	69.3 ± 45.0
Adapt 100ms	78.6 ± 11.7	59.5 ± 12.5	84.3 ± 32.6
Visual			
	AM	**DW**	**KH**
Baseline	65.0 ± 5.0	65.4 ± 6.7	92.8 ± 16.5
Adapt 0ms	75.1 ± 23.5	39.6 ± 14.7	71.1 ± 25.6
Adapt 100ms	81.7 ± 15.1	49.6 ± 16.2	78.2 ± 14.5

Appendix L: Table of results from Experiment 8.2.

Table 8.1. Paired-samples t-tests comparing baseline versus post-adaptation of a 500ms empty test interval. * denotes conditions significant at p = .05.

Condition	**Participant AM**	**Participant KB**	**Participant DW**
AA	.038*	.183	.191
VV	.031*	.136	.147
AV	.979	.052	.942
VA	.867	.090	.852

Appendix M: Table of results from Experiment 8.3.

Table 8.2. Paired-samples t-tests comparing baseline versus post-adaptation of a 500ms empty test interval. * denotes conditions significant at p = .05.

Condition	**Participant AM**	**Participant KB**	**Participant DW**
AA	.009*	.063	.224
VV	.025*	.542	.038*
AV	.545	.741	.006*
VA	.429	.055	.002*