

University of Bradford eThesis

This thesis is hosted in [Bradford Scholars](#) – The University of Bradford Open Access repository. Visit the repository for full metadata or to contact the repository team



© University of Bradford. This work is licenced for reuse under a [Creative Commons Licence](#).

Human sensory time perception

A psychophysical investigation of audio-visual timing in
the millisecond range

John Hotchkiss

Submitted for the degree of Doctor of Philosophy

School of Optometry and Vision Science

University of Bradford

2012

Keywords: Time perception; channels; adaptation; duration; spatial frequency.

Abstract

John Hotchkiss

Human sensory time perception

Time perception; channels; adaptation; duration; spatial frequency

The experiments described in this thesis use psychophysical techniques and human observers to investigate temporal processing in the millisecond range. The thesis contains five main sections. Introductory chapters provide a brief overview of the visual and auditory systems, before detailing our current understanding of duration processing. During the course of this review, several important questions are highlighted. The experiments detailed in Chapters 8-11 seek to address these questions using the psychophysical techniques outlined in Chapter 7. The results of these experiments increase our understanding of duration perception in several areas. Firstly, Experiments 1 and 2 (Chapter 8) highlight the role of low level stimulus features: even when equated for visibility stimuli of differing spatial frequency have different perceived durations. Secondly, a psychophysical hypothesis arising from the “duration channels” or “labelled lines” model of duration perception is given strong support by the adaptation experiments detailed in Chapter 9 and 10. Specifically, adaptation to durations of a fixed temporal extent induces repulsive duration aftereffects that are sensory specific and bandwidth limited around the adapted duration. Finally Chapter 11 describes the results of experiments designed to probe the processing hierarchy within

duration perception by measuring the interdependency of illusions generated via duration adaptation and via multisensory cue combination. The results of these experiments demonstrate that duration adaptation is a relatively early component of temporal processing and is likely to be sub served by duration selective neurons situated in early sections of the visual and auditory systems.

Acknowledgements

This thesis would not have been possible without the help and support of many people to whom I owe a vast debt of gratitude. Firstly, I would like to thank my supervisors Professor David Whitaker and Dr James Heron for their good natured tolerance of my many shortcomings. Starting a PhD in middle age is a daunting prospect and I have to thank them both for their help and support in making psychophysics comprehensible, helping me to think in a more scientific way and most of all, making the whole process really enjoyable. I also owe a huge debt to my unofficial third supervisor Dr Craig Aaen-Stockdale who strove manfully to initiate me into the mysteries of Matlab and was a constant source of help and advice.

I would like to thank everyone within the School of Optometry and Vision Science for making me feel so at home during my time in Bradford. I have thoroughly enjoyed helping out with teaching and find it difficult to imagine that there is a friendlier department anywhere in the country.

Finally, I would like to thank my wife Val for her encouragement and rock solid support throughout. Without her, I would never have started a PhD much less finished one.

Table of contents

Abstract	i
Acknowledgements.....	iii
Table of contents	iv
Chapter 1	1
Introduction	1
Chapter 2	5
The visual and auditory systems.....	5
2.1 The visual system	5
2.2 Spatial vision.....	17
2.3 The auditory system.....	29
Chapter 3	37
Models of time perception	37
3.1 Pacemaker/accumulator theories	39
3.1.1 Weber's law.....	46
3.2 Rival models for time perception.....	48
3.2.1 Distributed timing.....	49
3.2.2 Labeled lines/Channels	51
3.2.3 Intrinsic timing models.....	54
Chapter 4	58
Factors influencing time perception	58

4.1	Stimulus nature.....	58
4.2	Stimulus expectancy.....	61
4.3	Attention.....	67
4.4	Adaptation.....	69
4.4.1	Spatially specific adaptation	76
Chapter 5.....		81
Multisensory integration		81
5.1	Multisensory neurons.....	81
5.2	Audio-visual multisensory integration in humans	87
5.3	Cue combination.....	92
Chapter 6.....		99
Neural correlates of duration perception		99
6.1	Motor and sensory timing.....	105
Chapter 7		109
Psychophysical methods		109
7.1	Duration scaling	109
7.2	Duration discrimination	111
7.3	Methods used in this thesis.....	113
7.4	Apparatus	127
7.5	Gamma correction	128
Chapter 8.....		130

Spatial frequency and perceived duration	130
8.1 Introduction	130
8.2 General methods	131
8.2.1 Observers.....	131
8.2.2 Apparatus.....	131
8.2.1 Stimuli	133
8.3 Experiment 1.....	133
8.4 Procedures	133
8.5 Equating for visibility	134
8.6 Results.....	136
8.7 Experiment 2.....	140
8.8 Methods.....	140
8.9 Equating for visibility	142
8.10 Results.....	144
8.11 Discussion	148
Chapter 9	154
An investigation into duration adaptation	154
9.1 Introduction	154
9.2 Materials and Methods.....	155
9.2.1 Participants	155
9.2.2 Stimuli	155

9.3	Procedure	156
9.4	Results.....	158
9.4.1	Experience dependent duration plasticity.....	159
9.4.2	Tuned duration after-effects	162
9.4.3	Scaled, self-similar duration channels.....	164
9.4.4	Modelling the effects of duration adaptation.....	168
9.5	Duration adaptation or anchoring?.....	170
9.6	Duration adaptation or temporal frequency adaptation?	171
9.7	Introduction	171
9.8	Methods.....	172
9.8.1	Subjects	172
9.8.2	Stimuli	172
9.8.3	Procedures.....	172
9.8.4	Results	173
9.9	Discussion	175
9.9.1	Psychophysical context.....	176
9.9.2	Neural basis	178
9.10	Conclusions.....	183
Chapter 10	184
Stimulus specificity of duration adaptation	184
10.1	Is duration adaptation truly intra modal?.....	184

10.1.1	Introduction	184
10.2	Methods.....	186
10.2.1	Subjects	186
10.2.2	Stimuli	186
10.2.3	Procedures.....	186
10.3	Results.....	188
10.4	Discussion	190
10.5	The role of visibility	192
10.6	Introduction.....	192
10.7	Methods.....	193
10.7.1	Subjects	193
10.7.2	Procedure	193
10.8	Results.....	194
10.9	Discussion	195
10.10	Adaptation or aftereffect?	196
10.11	Introduction.....	196
10.12	Methods.....	197
10.13	Subjects.....	197
10.14	Stimuli.....	197
10.15	Procedures	198
10.16	Results.....	199

10.17	Discussion	200
10.18	Adaptation and spatial frequency	202
10.19	Introduction.....	202
10.20	Methods.....	203
10.20.1	Subjects	203
10.20.2	Stimuli	203
10.21	Procedures	204
10.21.1	Adaptation phase	204
10.21.2	Test phase	204
10.22	Results.....	205
10.23	Discussion	210
10.24	Is temporal adaptation Retinotopic or Spatiotopic?	212
10.25	Introduction.....	212
10.26	Methods.....	215
10.26.1	Apparatus and stimuli	215
10.26.2	Subjects	215
10.26.3	Procedure	215
10.27	Results.....	217
10.28	Discussion	221
Chapter 11	224
Duration adaptation precedes sensory integration	224

11.1	Introduction.....	224
11.2	General methods.....	227
11.2.1	Participants.....	227
11.2.2	Stimuli and apparatus.....	227
11.3	Multisensory integration.....	228
11.3.1	Procedure.....	228
11.3.2	Results.....	231
11.4	Adaptation.....	236
11.4.1	Procedure.....	236
11.4.2	Results.....	238
11.4.3	Discussion.....	241
Chapter 12	246
Conclusions	246
References	250

Chapter 1

Introduction

The way humans perceive time appears to be influenced by a number of factors. Research into these factors and the way they influence our experience of time has produced a mass of empirical data which researchers have attempted to explain by means of a variety of models.

It is generally accepted that different strategies are adopted by animals to process different time scales (see Figure 1) and these time scales may be divided into 4 main categories (microseconds, milliseconds, seconds/minutes and circadian rhythms) which require differing mechanisms in order to process the passage of time (Buonomano et al., 2002, Buhusi et al., 2005, Lewis et al., 2003). Microsecond timing enables the differential arrival of sound between two ears to be perceived. This in turn permits an animal or human to detect the source of sounds, an ability which is essential to both the hunter and the hunted.

Millisecond timing is the range used in the production and understanding of speech in humans as well as many motor actions. Second/minute timing tends to be used in order to anticipate and recall events and it is this range which we are probably most consciously aware of in our everyday lives. Finally circadian rhythms regulate physical and behavioural activity over

longer periods. These are internally generated rhythms which permit an organism to synchronize behavioural and physical cycles such as hormone levels and periods of sleep with the cycles of its environment. For example human core body temperature varies over a 24 hour cycle and is at its lowest at around 3-4 o'clock in the morning rising to its highest in the late afternoon. Circadian timing depends upon the phase of the 24 hour oscillation. The circadian clock is able to set its phasing through the use of environmental cues such as the presence or absence of light and the phases of the circadian clock are continuously updated.

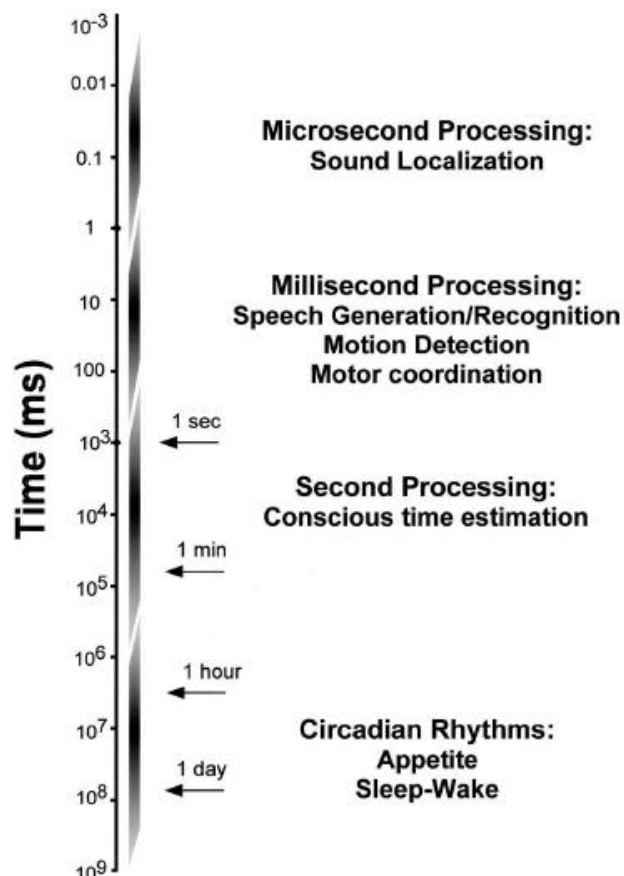


Figure 1 Time scales which are thought to require differing temporal processing mechanisms (Buonomano et al., 2002).

This thesis is chiefly concerned with millisecond timing, which in common with circadian timing, but unlike second/minute timing, is thought to be beyond cognitive control and automatic. In the case of circadian timing we can point to neural structures and have an understanding of how these structures mediate the process (Figure 2).

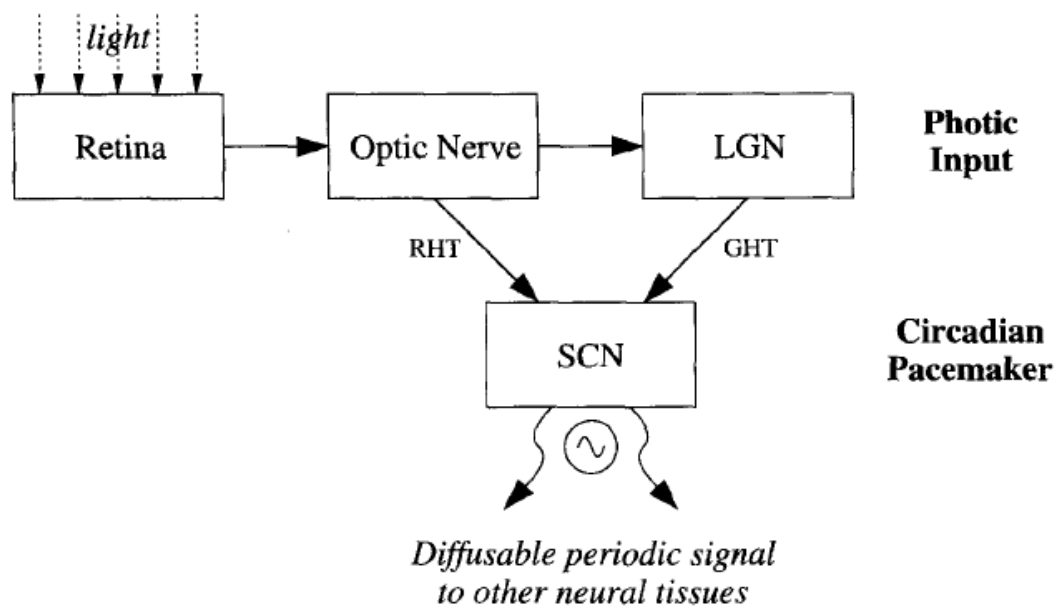


Figure 2 Circadian timing. The suprachiasmatic nuclei (SCN) generate an approximately 24 hour oscillation. The phase of this rhythm is synchronized with the light/dark cycle of the external environment. This is achieved when light falls on the retina and a neural signal is sent down a subset of neurons in the optic nerve known as the retinohypothalamic tract (RHT). The RHT connects with the SCN in the hypothalamus. At a later stage of the visual pathway, the lateral geniculate nucleus (LGN) sends another signal to the SCN via the geniculohypothalamic tract (GHT). Variations in light levels are then communicated to other brain areas. In this way, cyclical fluctuations in the external light/dark cycle are accommodated and external and internal cycles synchronised (Hinton et al., 1997).

As yet no equivalent neural structures have been attributed to the processing of millisecond timing or second/minute timing, although many models involving much faster oscillators than circadian timing have been suggested (Creelman, 1962, Treisman, 1963, Gibbon, 1977, Zakay et al., 1997). It should also be noted that the division between automatic millisecond timing and cognitively controlled interval estimation (second/minute timing) is by no means clear. The dividing line between these categories is thought to be somewhere between 0.50 and 2.0 seconds, but since these differing timing mechanisms are thought to run in parallel, a certain amount of overlap is probably to be expected (Buhsu et al., 2005, Buonomano et al., 2002).

The experiments described here all investigate sensory timing and involve vision and audition. Therefore I begin with an overview of the visual and auditory systems, paying particular attention to spatial vision.

Chapter 2

The visual and auditory systems

2.1 The visual system

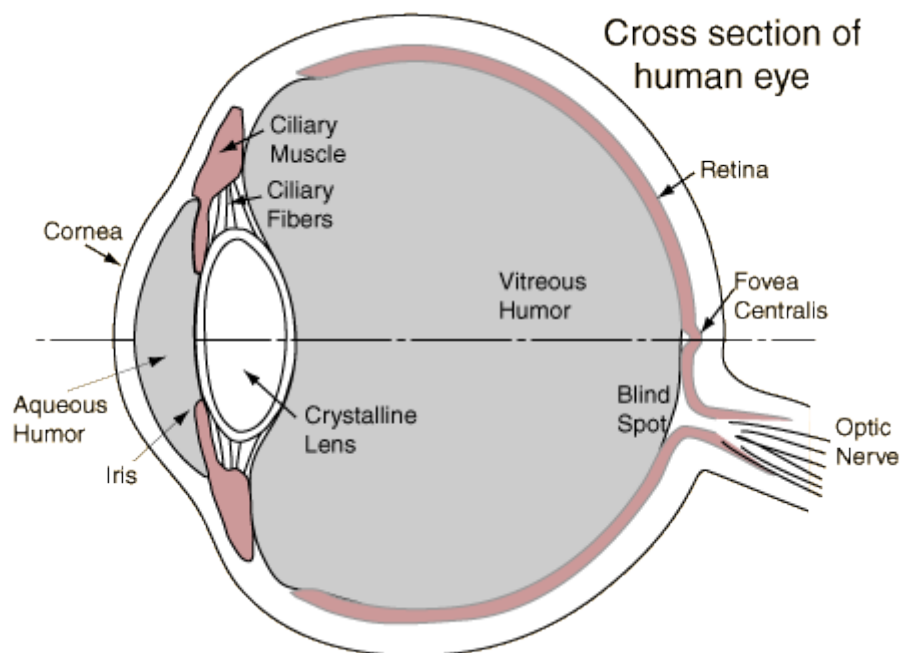


Figure 3 A schematic of the human eye. Light rays are refracted by the cornea, aqueous humour, crystalline lens, vitreous humour before reaching the retina. <http://hyperphysics.phy-astr.gsu.edu/hbase/vision/imgvis/eyesection.gif>

In order to construct a visual representation of the outside world our eyes are required to focus light from external objects onto a light sensitive surface. The light rays are refracted during their passage through the transparent

media of the cornea, aqueous humour, crystalline lens and vitreous humour before arriving at the retina (Figure 3).

Light falling on the retina is then transduced into electrical impulses which progress onwards to the brain. The retina has a layer of light sensitive cells known as photoreceptors which form the penultimate layer of the eye, and lie next to the retinal pigment epithelium (the outer layer (See Figure 5)). These photoreceptors consist of inner and outer segments and two receptor types - rods and cones (Figure 4).

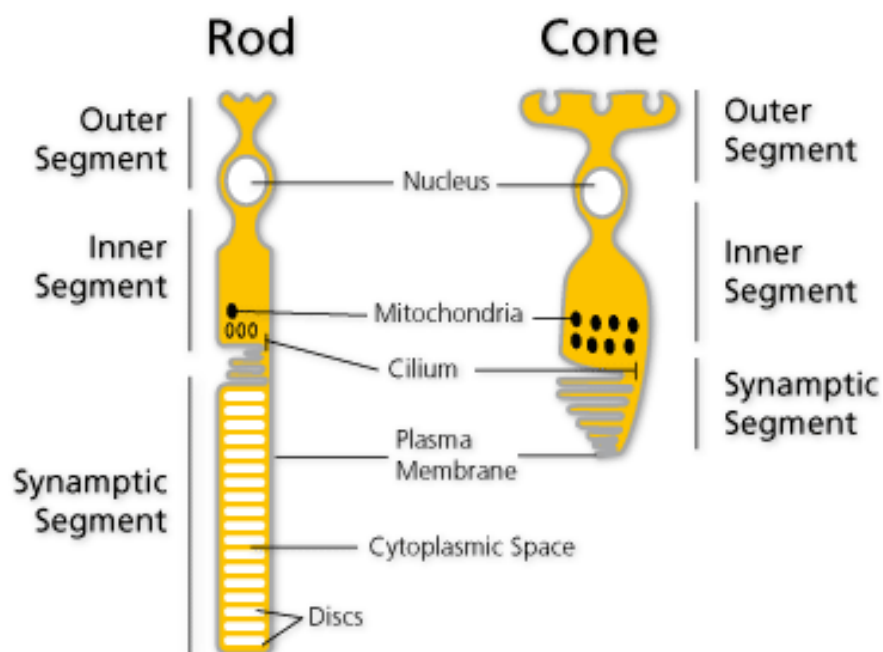


Figure 4 Rod and cone structure

http://www.sharpsighted.org/images/stories/sight/rods_cones/rods_cones.gif

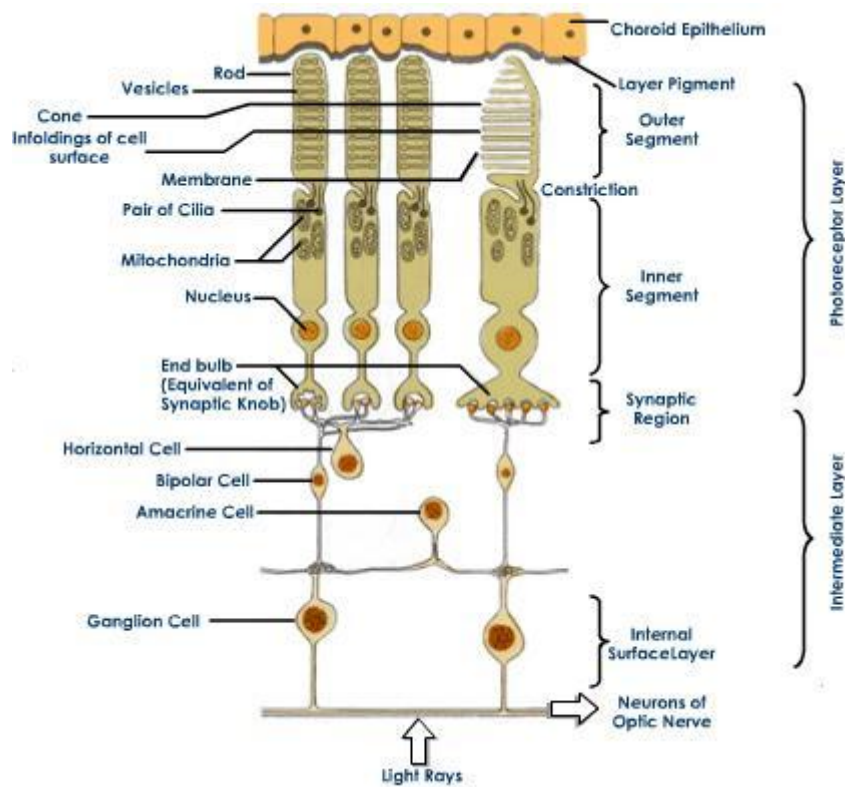


Figure 5 The layers and cell types of the human retina. <http://images.tutorvista.com/content/nervous-coordination/retinal-part-layers-structure.jpeg>

Rods and cones both transduce photons of light energy into electrical impulses, but there are important differences between the two receptor types. Cones operate under conditions of relatively high light intensity (photopic conditions) and three different cone types are receptive to three different spectral ranges of the electromagnetic spectrum. L, M & S cone types have their respective peak sensitivities at 560nm, 530nm and 420nm (Svaetichin, 1956, Dartnall et al., 1983).

This enables the visual system to analyse an objects' colour. Cones are plentiful in and around the macula but become increasingly sparse with rising retinal eccentricity.

Rods on the other hand, are absent from the fovea but become more densely distributed as we move away from the macular reaching a peak at around 20 degrees of eccentricity. Rods and cones have different outer segment structures. Rods possess a thin 'rod-like' outer segment in contrast to their more stumpy 'cone-like' counterparts. These structural differences are in part, related to differences in photoreceptor function. Both types of outer segment contain numerous intricate folds of their plasma membrane – known as 'discs' (Figure 4) which are surrounded by molecules of visual pigment. Differences in the pigment found in rods, and the three different cone types, allow the photoreceptor to be sensitive to the visible spectrum of electromagnetic radiation. Rods contain more than 1000 free-floating discs within their outer segment. This gives them a surface area which is much larger than that of the cones. Pigment molecules undergo a biochemical chemical reaction which is responsible for the photosensitivity of rods and cones.

Photons of incident light energy cause changes in molecular shape, which, in turn produce changes in the electric current around the molecule. This results in the transmission of an electric current along the outer membrane towards the synapse with the next neuron. This electrical activity is transferred to the next neuron in the retinal circuit via chemical transmitters present within the synaptic junctions of the photoreceptors. Because of their greater surface area, incident photons are much more likely to strike a molecule of visual pigment on a rod than a cone and therefore rods are much more sensitive to

light. However, because they are sensitive to only one wavelength band (peaking at around 500nm) they are insensitive to colour. Cones have a smaller number of discs attached to the outer segment membrane. This gives a reduced surface area and subsequently cones have a relatively lower sensitivity to light energy.

After this process, pigment molecules are said to be 'bleached' and must be restored to their original molecular shape. This is achieved via the action of enzymes present within the retinal pigment epithelium (RPE). This layer removes spent discs from photoreceptors via phagocytosis, thus maintaining the health of the inner retinal layers.

At the fovea a single cone is connected to a single midget bipolar cell. This is a connecting neuron which has two axons, one extending to the posterior and one to the anterior. Foveal bipolar cells are connected to a midget ganglion cell. These are also neurons but have only one axon. The axons from ganglion cells form the retinal nerve fibre layer which is the innermost retinal layer. Retinal nerve fibres go on to form the optic nerve (Figure 5). They exit the eye via the lamina cribrosa of the optic nerve head. This area of the retina has no photoreceptors and corresponds to the 'blind spot'. As we move away from the fovea, a number of cones may connect to a single diffuse bipolar cell and in turn, several diffuse bipolar cells connect to a single parasol ganglion cell. Thus the area of the visual field to which a ganglion cell responds, which is known as its 'receptive field' increases the further away from the fovea it is. In consequence, the larger the area of summation, the lower the spatial resolution within the corresponding area of the visual field will be. It is in part due to this, that resolution is high near fixation because

receptive fields here are small. In the peripheral visual field on the other hand, receptive fields increase in size as we move away from fixation and resolution diminishes. Several rods project to a single rod bipolar cell however there are no direct connections between rod bipolars and ganglion cells. Instead, electrical activity from rods reaches ganglion cells indirectly via horizontal connecting units known as amacrine cells and so, in this way, signals from the rods enter the cone pathway (Figure 5).

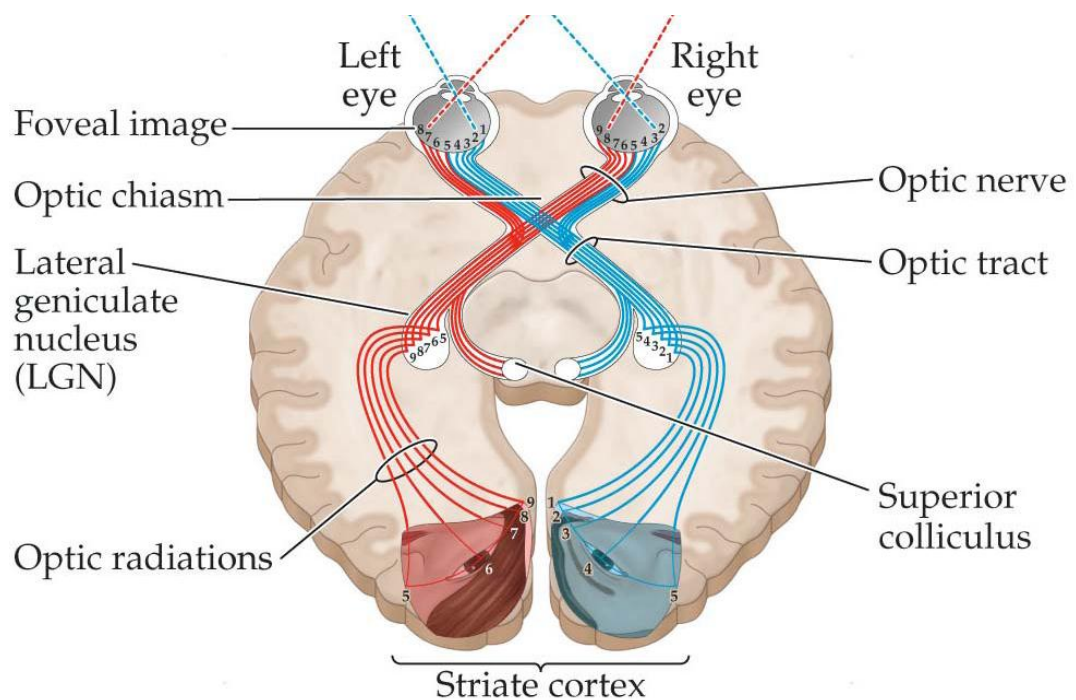


Figure 6 An axial view of the visual pathway.

http://homepage.psy.utexas.edu/homepage/faculty/pillow/courses/perception09/slides/Lec05_V1.pdf

Retinal ganglion cell axons progress down the optic nerve toward the optic chiasm (Figure 6). The fibres originating from the nasal half of the retina

cross the midline to combine with those from the temporal half of the opposite eye at the chiasm. This allows objects in one half of the visual field to be represented by ganglion cells from both left and right eyes. In this way the optic tract carries information from both retinæ from the contralateral visual field.

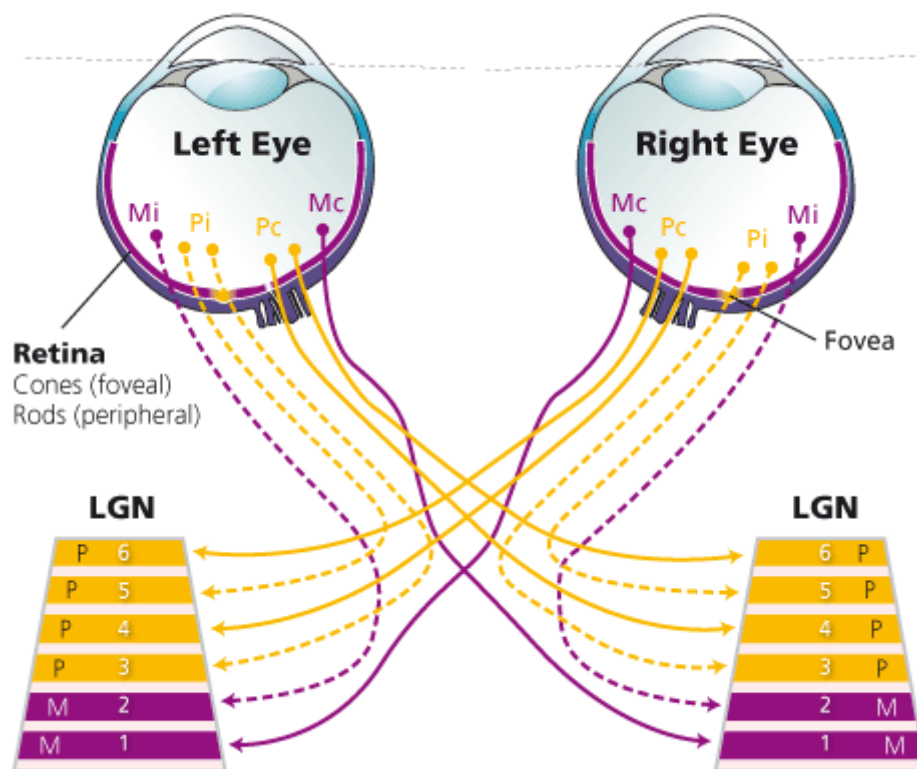


Figure 7 M and P pathway layers within the LGN.

http://www.sharpsighted.org/images/stories/sight/magno/m_p_eyes_1.gif

Just before reaching the lateral geniculate nucleus (LGN) a small percentage of fibres leave the optic tract and synapse in the superior colliculus. This

structure is involved in stimulus localisation, coordination of eye movements and the integration of information across the senses (Wallace et al., 1997, Lunenburger et al., 2001). The first synapse is encountered at the LGN. The LGN's main role is believed to be the relaying and segregation of ganglion cell signals into different information streams. Retinal ganglion cells may be classified as Magnocellular (M) or Parvocellular (P). The Magnocellular and Parvocellular pathways project to different layers within the LGN. The LGN has a laminar arrangement and keeps left and right eye inputs separate (see Figure 7). The more dorsal layers contain alternate left/right eye Parvo cells with smaller cell bodies. The more ventral layers show alternate layers containing left/right eye Magno cells.

Magno type cells have a relatively high sensitivity to contrast; transient cell responses; large RF sizes, and relatively low colour sensitivity. By contrast, Parvo type cells have relatively low sensitivity to contrast; sustained cell responses; smaller RF sizes and relatively high colour sensitivity.

From the LGN, fibres travel to the striate cortex via the optic radiations. This cortical region is responsible for the initial stage of cortical processing and forms part of the occipital lobe. It is split into left and right hemispheres which represent the left and right halves of the visual field due to the crossing of the fibres at the chiasm. The striate cortex, which is also known as "V1" begins to decipher the encoded signals from the retina and also further encodes the signals before they are processed in neighbouring 'higher' visual areas (V2-V5, see Figure 8). The cortical surface is an intricately folded sheet of neurons which has a large surface area and a correspondingly vast number of neurons. Visual cortical areas all demonstrate some form of retinotopic

representation of space which corresponds to that of the retina. Thus any particular point within the visual field will stimulate a particular point on the retina, which will stimulate a particular region of neurons within the visual brain region. Neighbouring areas in the visual field will stimulate neighbouring cortical neurones and so a neuronal representation of the outside world may be produced. Because this neuronal map is linked to retinal coordinates it is said to be retinotopic. However, not all regions of space have an equal share of cortical space. The closer a region is to fixation, the larger will be its allocation of cortical space (the number of neurones processing signals from that region of the visual field).

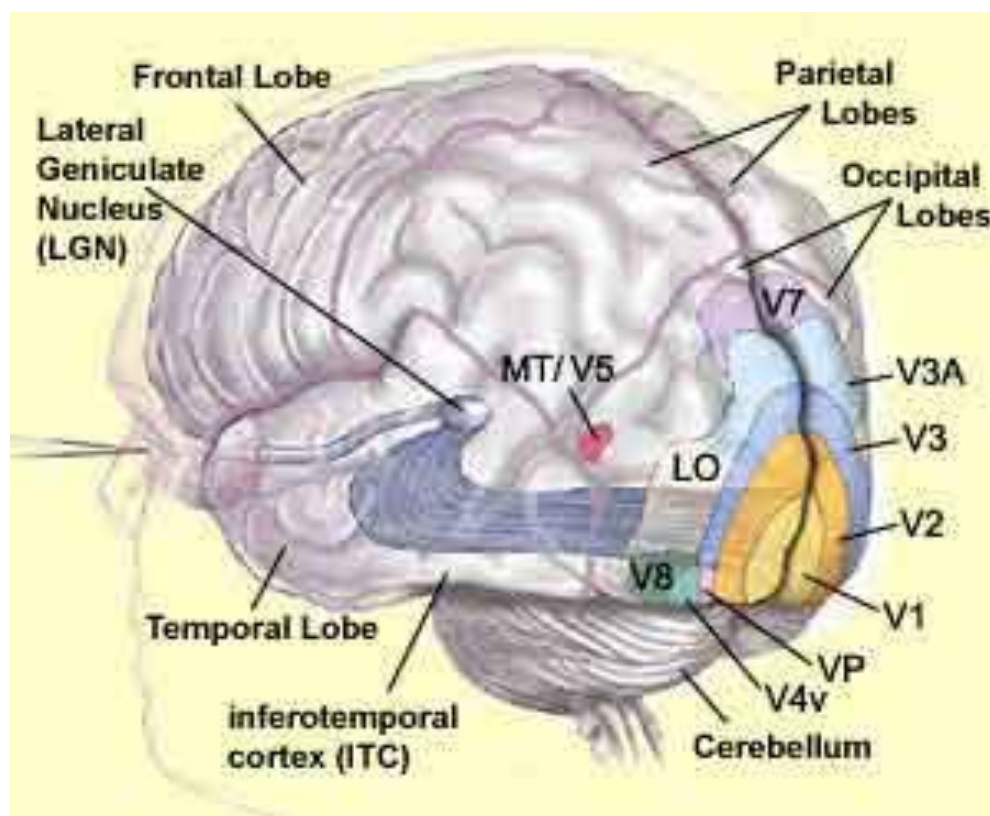


Figure 8 The different areas of the visual cortex <http://mybrainnotes.com/memory-language-brain.ht>

A large proportion of visual information processing is carried out in the occipital lobe, but neighbouring regions also influence visual performance. A somewhat controversial theory suggests the existence of two cortical pathways – the dorsal ‘where’ and the ventral ‘what’ judgments (Figure 9) (Mishkin et al., 1983). The vast majority of LGN fibres enter the visual cortex via V1, but subsequent processing pathways do not follow serial processing (V1 through to V5). Cortical regions have numerous layers (superficial through to deep) that are interconnected via feed forward projections (ascending projections from lower areas of the nervous system) and feedback projections (descending projections from higher areas).

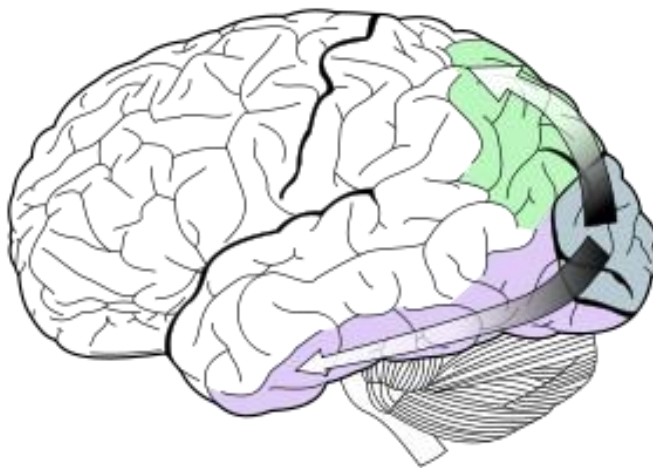


Figure 9 The proposed segregation of ‘where’ and ‘what’ information in the brain. The dorsal stream (green) and ventral stream (purple). http://en.wikipedia.org/wiki/Two-streams_hypothesis

These connections allow the flow of information forward and backward between two regions. The direction of these connections appears to be dependent upon the cortical layer projections from which they originate (Angelucci et al., 2003). Feedforward projections originate in the superficial layers and project to layer 4. Feedback projections originate in the superficial and deep layers and go on to terminate in the superficial and deep layers of earlier cortical regions. This demonstrates the complex nature of cortical interconnectivity and can make labels such 'early' and 'late' processing somewhat imprecise (Lennie, 1998, Van Essen et al., 1992).

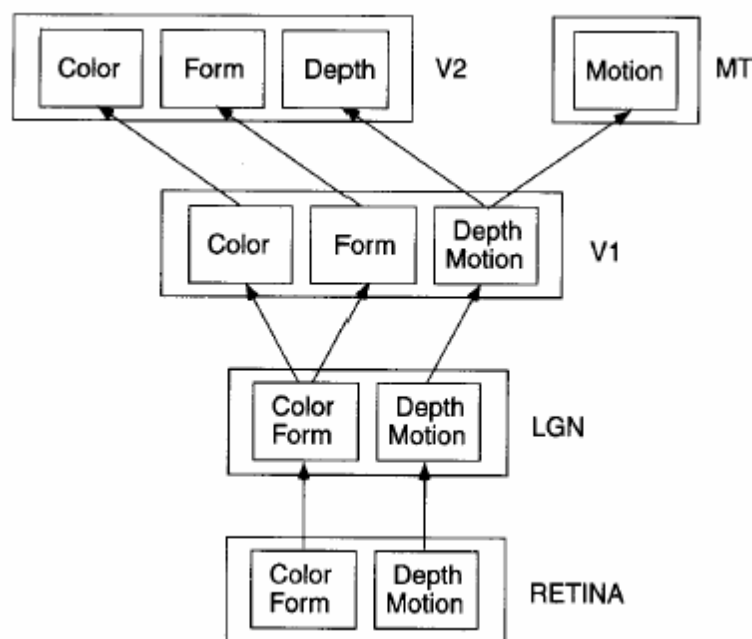


Figure 10 A schematic demonstrating how different forms of visual information may be organised within the visual system (Palmer, 1999)

However, lesion, brain imaging and neurophysiological studies provide some evidence that different forms of visual information are processed by neural pathways that are anatomically distinct (Zeki, 1978). In other words, specialised groups of neurons process particular categories of visual information separately (see Figure 10). This model for visual processing is known as modular processing and is thought to begin as early as the retina. Two different classes of ganglion cells are thought to be involved in the processing of colour/form and depth/motion. Other stimulus features are also separated as the pathway progresses upward through the nervous system (Livingstone et al., 1988). As visual information progresses to visual areas outside the occipital lobe, the segregation of information shown in Figure 10 may be combined with the dorsal/ventral hypothesis of (Mishkin et al., 1983). Specifically colour and form may project from V2 to the ventral 'what' pathway, and depth and motion may project from V5 to the dorsal 'where' system. However, the extent to which this model may be generalised across the nervous system is called into question by human psychophysical evidence of early interaction between differing stimulus attributes (Lennie, 1998).

2.2 Spatial vision

Images which fall on the retina contain information of different spatial scales, ranging from very coarse down to very fine. The visual system has neurons with different sized receptive fields (see previous section). These detect the presence of patterns at different scales. Photoreceptors are connected to bipolar and ganglion cells in different ways. Cones are directly connected to midget bipolar cells as well as indirectly via horizontal cells. The direct pathway may have an excitatory or inhibitory influence on the electrical potential generated within the bipolar cell. The indirect pathway can also be either excitatory/inhibitory but its polarity will always be the converse of the direct pathway. This creates a 'centre-surround' organisation of the bipolar cells receptive field (RF) such that if the centre of the receptive field is of the ON-centre type the surround is of the OFF-centre type (Werblin et al., 1969). This is demonstrated in Figure 11 which shows how the spatial distribution of light across the cells' receptive field governs the firing rate of the ganglion cell. For (a) and (c) the ON and OFF fields of the cell are equally stimulated. If the whole of the receptive field of the cell is uniformly illuminated the cell's firing rate is the same as that produced in the absence of incident light. However when the ON field of the cell only is stimulated the result is an increase in firing rate.

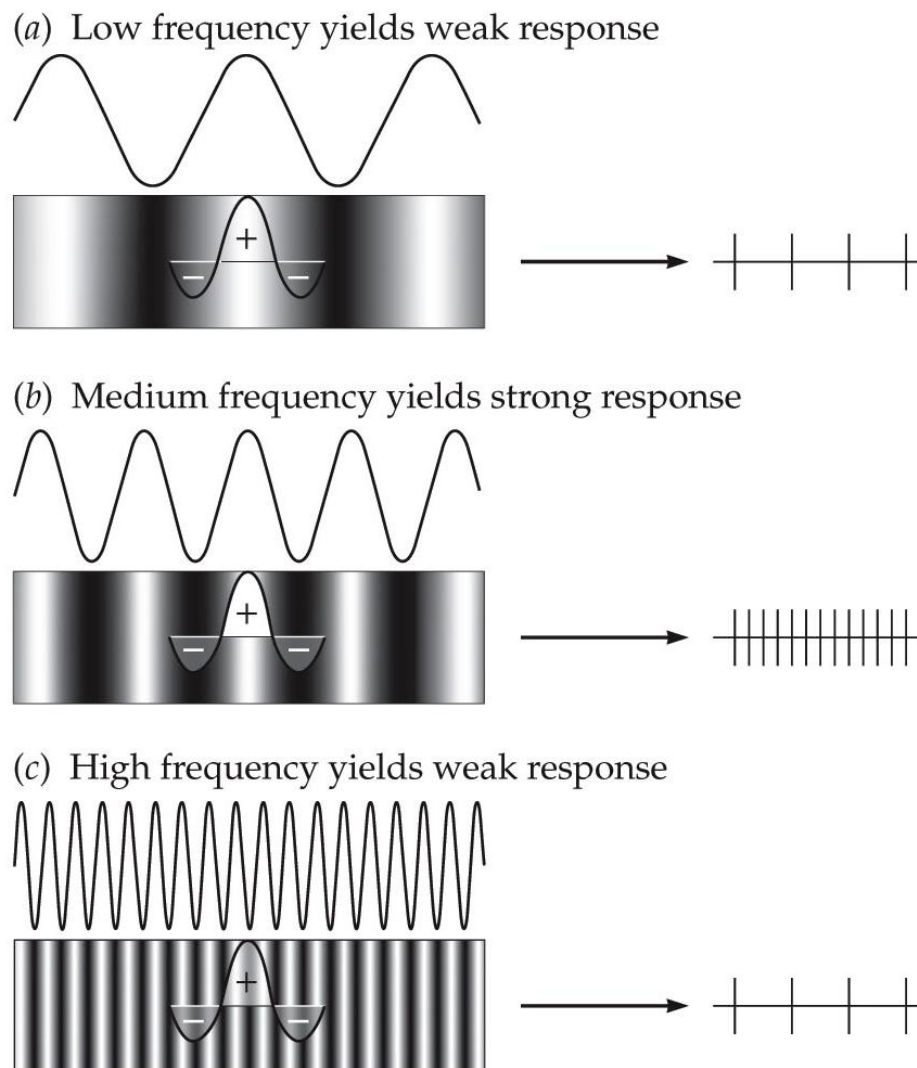


Figure 11 The organisation of excitatory & inhibitory receptive field regions and how this organisation dictates the way in which firing rates are influenced by different spatial distributions of light.

http://homepage.psy.utexas.edu/homepage/faculty/pillow/courses/perception09/slides/Lec05_V1.pdf

The organisation of striate neurons within V1 has been mapped out via animal studies (Levay et al., 1975, Tootell et al., 1982, Hubel et al., 1968a). The hemispheres have distinct regions of neurons dominated by inputs from left and right retinae. These may be seen on the cortical surface as

approximately alternate layers called 'ocular dominance columns' that are oriented at right angles to the cortical surface.

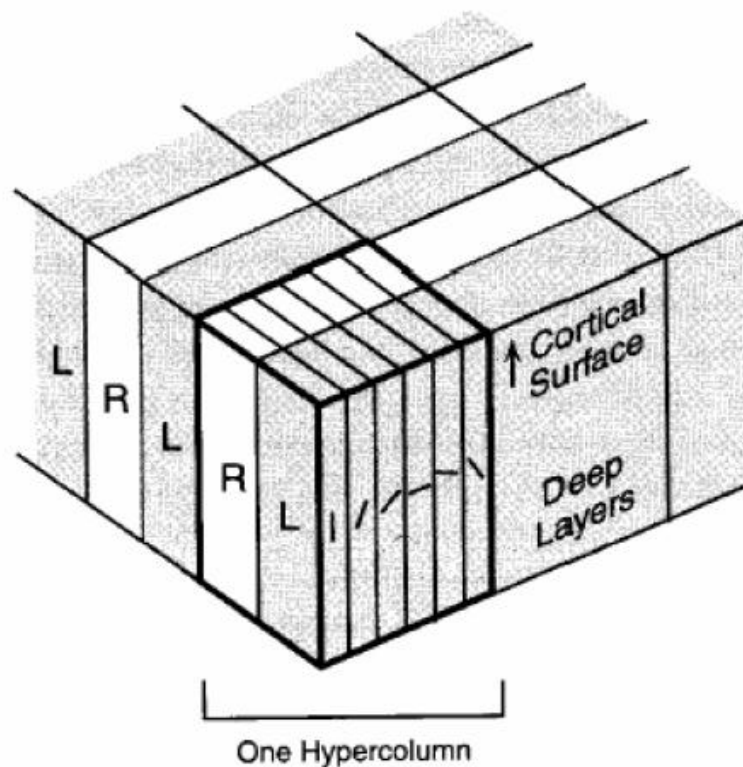


Figure 12 A schematic showing a hypercolumn made up of two layers of right and left dominant cells which run perpendicular to the cortical surface. Within the layers, orientation sensitivity variations run parallel to the cortical surface (Palmer, 1999)

The columns within each hemisphere make up a retinotopic map. On a smaller scale, the retinotopic map of space consists of smaller, approximately 1mm x 1mm sub-units known as 'hypercolumns' which run perpendicular to

the left/right divisions of the ocular dominance slabs (See Figure 12). Each hypercolumn represents a specific point within the retinotopic map and has many orientation columns which are sensitive to a particular stimulus position and orientation. Research involving primates has revealed more about the nature of the striate cortex. It has been suggested that in some areas orientation columns may be arranged around a series of points in a circular fashion such that the different orientations are arranged around each point and meet in the middle (Braitenberg et al., 1979). These arrangements are known as pinwheels and are thought to have an area at their centre (or hub) which has reduced orientation selectivity and is known as a singularity (Horton et al., 1981, Blasdel, 1992). In other areas fractures have been found in the sequence of the orientation selective neurons (Hubel et al., 1974). It has been suggested that fractures and pinwheels could be defects which occur during cortical development (Wolf et al., 1998). Other authors hypothesise that pinwheels and fractures are adaptations which allow for a more efficient use of cortical neurons by minimising wiring length (Swindale, 1996, Chklovskii et al., 2004).

In addition to orientation columns, there are areas of the visual cortex which stain with cytochrome oxidase known as CO blobs. CO blobs tend to sit in the centre of ocular dominance columns in evenly spaced lines. They receive input from M, P, and K layers of the LGN, are usually sensitive to colour and have little or no orientation sensitivity (Wong-Riley, 1979, Livingstone et al., 1984). It has been suggested that pinwheel singularities and CO blobs may coincide in foveal regions, but are separate in parafoveal and peripheral regions (Alexander et al., 2010)

One possible problem for cortical cells could be the ambiguity of two or more combinations of stimulus attribute producing identical firing rates as a result of a cell's RF properties. For instance, a cell could respond identically to a low contrast stimulus of its preferred orientation or a high contrast stimulus at a less than optimal orientation.

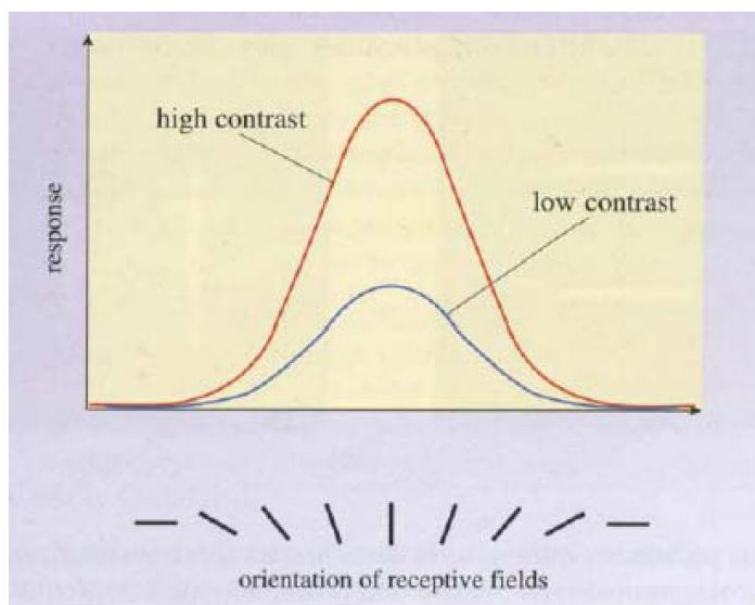


Figure 13 Population coding theory. The peak of the response to a vertical grating (as identified by comparing activation across a population of neurones) signals the same orientation regardless of stimulus contrast (Meese, 2002).

This risk may be reduced by comparing the distribution of activation across a group (or 'population') of neurons that respond to the same region of space. Thus the overall firing rate may differentiate between the two conditions. This process is known as population coding and it is demonstrated in Figure 13.

The way the visual system analyses and constructs a representation of an image is thought made possible by “psychophysical channels”. This theory is known as ‘spatial frequency theory’. It is based upon the concept of ‘Fourier Analysis’. Fourier analysis says that a visual image can be broken down into its constituent parts (Campbell et al., 1968, Blakemore et al., 1969(a), DeValois et al., 1990).

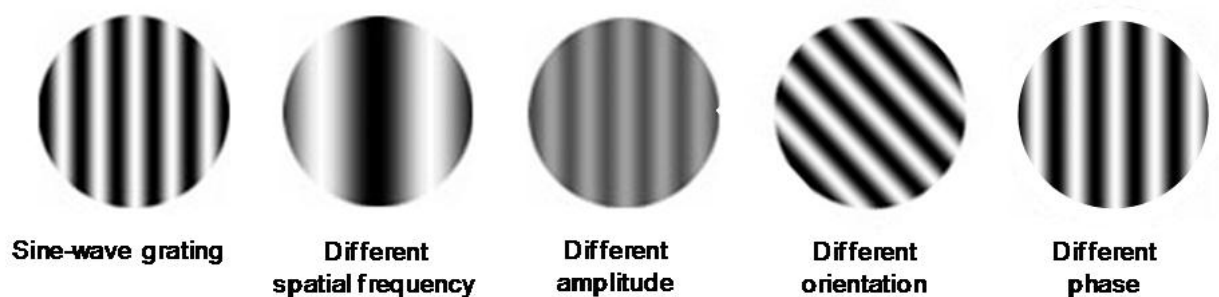


Figure 14 Differences in spatial frequency, phase, amplitude and orientation of a grating (De Valois & De Valois, 1988).

These parts consist of sine wave ‘gratings’ which are shown in Figure 15. Differences in spatial frequency, phase, amplitude and orientation of any given grating reflect differences in the spatial distribution of luminance across the grating (see Figure 14).

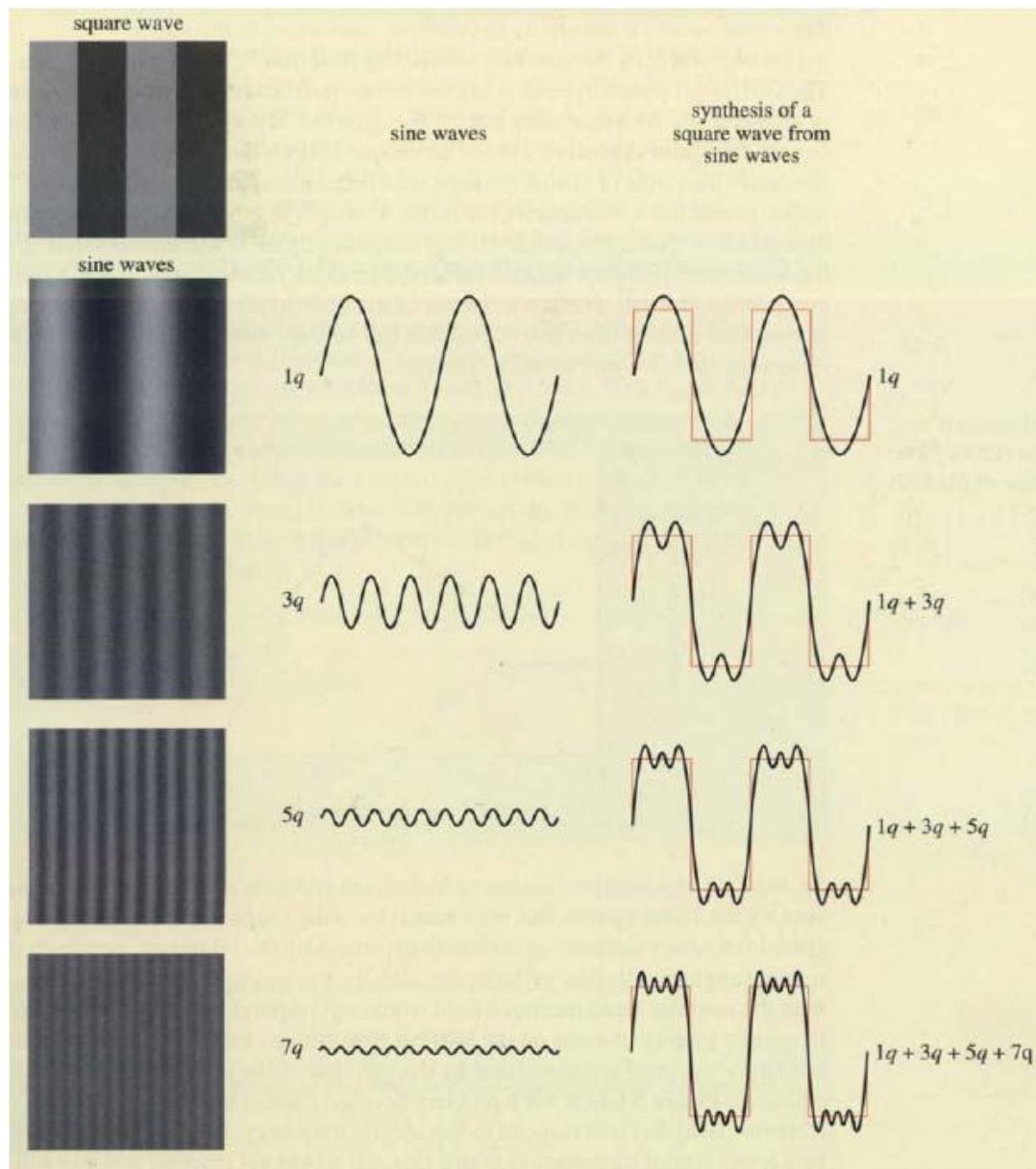


Figure 15 The 'hard-edged' square wave grating can be built up from adding together numerous sine waves of different frequencies and amplitudes (Meese, 2002).

Spatial frequency refers to the number of sine wave cycles (one dark and one light bar) within a grating of given size. This is expressed as cycles per degree (cpd) of visual angle. Put another way, spatial frequency tells us how

coarse (lower spatial frequencies) or fine (higher spatial frequencies) the detail of an image is.

Orientation refers to the angle of the bars within a grating with respect to vertical. Amplitude with respect to the mean luminance of the grating, relates to the contrast of the bars within a grating ($\text{contrast} = \text{amplitude} / 2 * \text{mean luminance}$). The larger the distance between peaks and troughs of the sine wave, the greater the contrast of the bars and the more visible is the grating.

Finally, phase refers to the 'lateral' position of a grating relative to either a fixed spatial position ('absolute' phase) or to relative to other spatial frequencies within a visual image ('relative' phase).

A fundamental frequency of a grating is the sine wave with the highest amplitude and the lowest frequency (1q in Figure 15). If we add together numerous gratings of progressively smaller amplitude and higher spatial frequencies we may produce an image (e.g. the hard-edged square wave grating shown in the top panel of Figure 15) from seemingly very different collection of sine waves (DeValois et al., 1990).

A substantial body of human psychophysical evidence exists which supports the theory of spatial frequency 'channels'. These "channels" may be thought of as filters which only allow the passage of information from spatial frequencies which coincide with the range of frequencies to which the channel is sensitive (Hubel et al., 1968a). This selectivity range is known as the "band width" of the channel. High pass filters are sensitive to high spatial frequencies and low pass to low spatial frequencies.

Our ability to detect a visual stimulus depends upon its spatial frequency and its contrast. Contrast detection thresholds may be measured and used to define the minimum contrast necessary for any particular spatial frequency to be detected by the human visual system. The reciprocal of the contrast detection threshold is known as the contrast sensitivity. The variation of contrast sensitivity across different spatial frequencies for an observer is called the contrast sensitivity function (Figure 16).

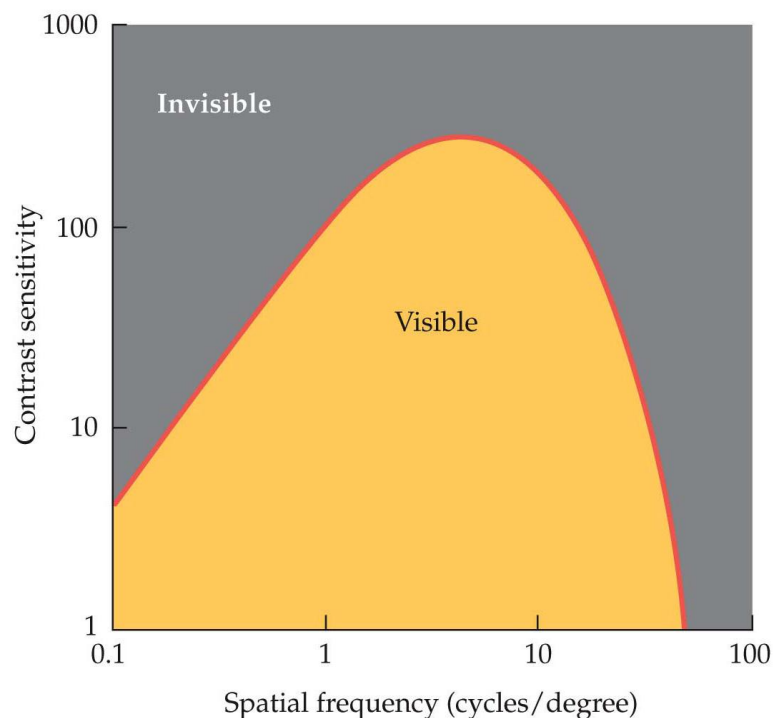


Figure 16 A schematic displaying the shape of a normal human contrast sensitivity function.

http://homepage.psy.utexas.edu/homepage/faculty/pillow/courses/perception09/slides/Lec05_V1.pdf

The contrast sensitivity function shown in Figure 16 illustrates the differential sensitivity to different spatial frequencies with peak sensitivity at around 4

cpd demonstrating the low contrast levels required for stimulus detection at this frequency. Spatial frequencies either side of this peak require greater contrast in order to be detected and it can be seen that there are frequencies at the extremes of the function which cannot be detected whatever the contrast level (Figures 16 and 17). The area under the curve is known as the 'window of visibility'.

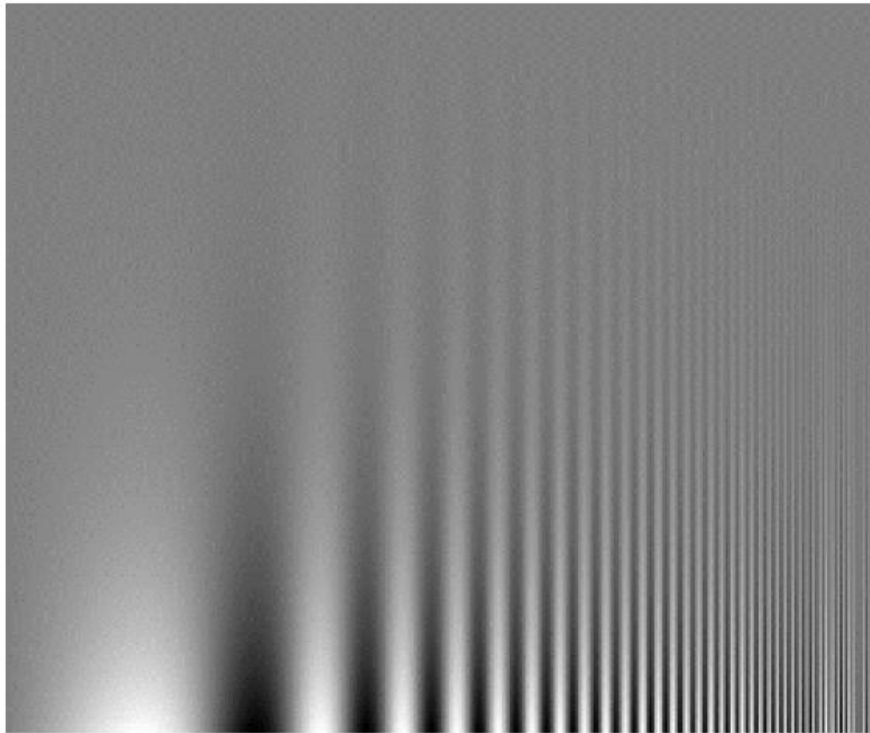


Figure 17 Visibility varies as a function of spatial frequency (bar width) and contrast
(<http://www.brad.ac.uk/acad/lifesci/optometry/resources/modules/stage1/pvp1/CSF.html>).

Contrast sensitivity varies considerably with differing luminance levels and also with differing retinal eccentricity. Extremely high or low spatial frequencies are invisible to humans. The inability of humans to detect very high spatial frequencies may be explained by the optics of the human eye. The eye is not a perfect optical which limits which consequently limits its spatial resolution. For extreme low spatial frequencies receptive field (RF) size is the limiting factor.

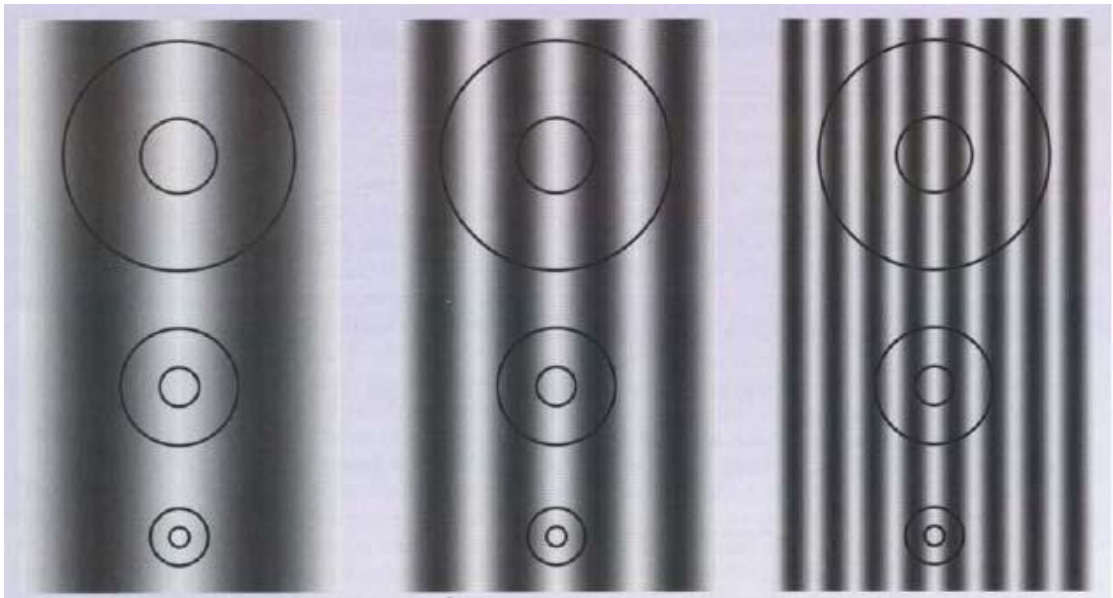


Figure 18 Three different sine wave grating with three different centre-surround RF's sizes superimposed on top of them. A notional diagonal line -joining top left with bottom right RFs - represents the RFs which will respond maximally to each grating (Meese, 2002).

Figure 18 demonstrates how a RF's spatial extent governs its spatial frequency selectivity. For a given spatial frequency, there will be a RF size

which responds maximally to that spatial frequency. The left most grating will initiate similar responses to the medium and smaller sized RF's as they both have a relatively uniform distribution of luminance across the ON/OFF components of their receptive fields. In contrast, the luminance distribution within the area of the largest RF (top left most) produces a maximal response to this grating by virtue of its centre surround (ON/OFF) antagonism. By following a diagonal line in Figure 18 from the top left RF to the bottom right RF it can be seen that each RF size responds maximally to the spatial frequency whereas the other six RFs acts a 'filter': blocking access to spatial frequencies not optimal to their centre-surround (ON/OFF) organisation. Spatial frequencies that are unable to initiate a response from the largest filters available to the visual system will therefore be 'invisible' (DeValois et al., 1990).

With increasing retinal eccentricity, only progressively larger RF sizes are available to process stimuli. Near fixation, a range of small (single or low number of ganglion inputs per V1 neuron) to large (numerous ganglion inputs per V1 neuron) RFs are accessible (Westheimer, 1984).

2.3 The auditory system

The auditory system transduces sound energy into an electrical impulse via a mechanical mechanism. Sound waves travel inward along the auditory canal to the tympanic membrane (ear drum). These waves strike the tympanic membrane and the resulting vibrations are transmitted to the cochlea by three small bones known as the 'ossicles' (see Figure 19). This mechanical energy is transferred across the bones to the stapes (the innermost bone). The stapes is connected to a membrane (the oval window) which covers an aperture in a spiral shaped bone structure known as the 'cochlea.' The oval window lies at the base of the cochlea. Also at the base of the cochlea is another membrane covered hole known as the round window (Bear, 2001). The other end of the spiral cochlea is known as the apex. The middle ear consists of an air filled area between the tympanic membrane and the oval window. Here, atmospheric pressure is maintained via the Eustachian tube which is connected to the throat. The ossicles increase the magnitude of the mechanical energy transmitted to the oval window. This increase in magnitude is due for the most part to the difference in size between the tympanic membrane and the oval window. An increase in energy magnitude transmitted through the middle ear is necessary because the membrane covering the oval window is more resistant to vibration, by virtue of the fluid contained within the cochlea. The mechanical energy transmitted by the ossicles proceeds via the oval window into the fluid filled cochlea.

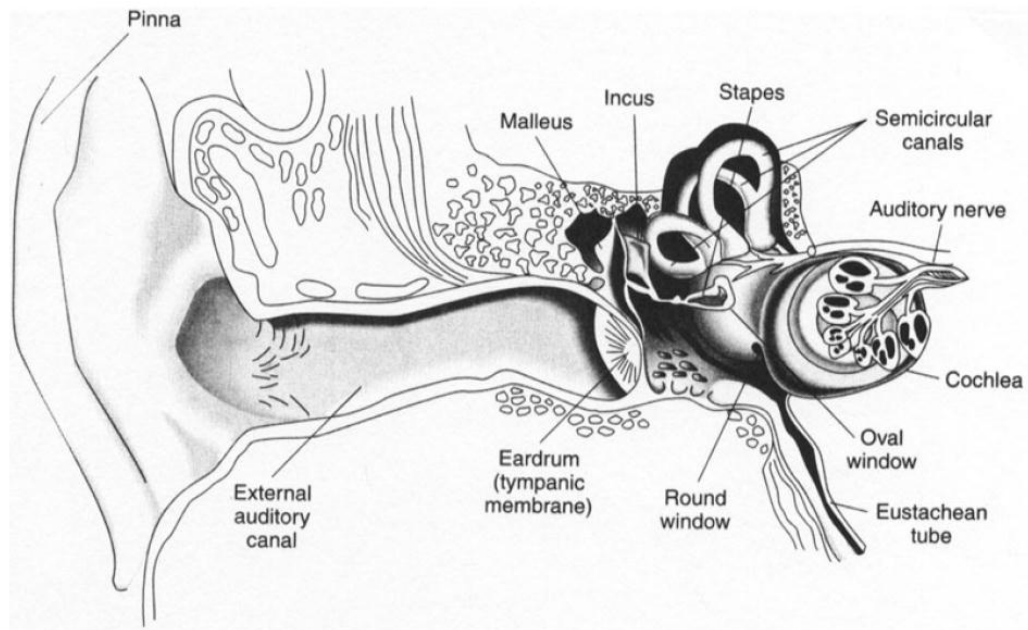


Figure 19 Diagram showing the anatomy of the human auditory system. Sound waves are filtered by the pinna and enter the auditory canal. When the sound waves meet the tympanic membrane, the vibration of the membrane is transmitted to the oval window of the cochlea by the malleus, incus and stapes (collectively known as the ossicles). From the cochlea, neural signals corresponding to the nature of the incident sound are initiated which travel toward the brain.

www.skidmore.edu/~hfoley/Perc9.htm.

It is the cochlea which converts the mechanical energy produced by the sound into neural signals (Figures 19, 20 and 21). The cochlea is made up of three parallel, fluid-filled canals known as the vestibular canal, tympanic canal, and middle canal (scala vestibule, scala, tympani and scala media). The vestibular canal and the middle canal are separated by Reissner's membrane whilst the basilar membrane separates the tympanic canal from the middle canal (Bear, 2001). The organ of corti which, contains auditory receptor neurons, lies on the basilar membrane. Superior to the organ of corti

is the tectorial membrane. The middle canal is closed off at the apex of the cochlea and the tympanic and vestibular canals are connected by a hole known as the helicotrema. At the base of the cochlea the tympanic canal meets the round window and the vestibular canal, the oval window.

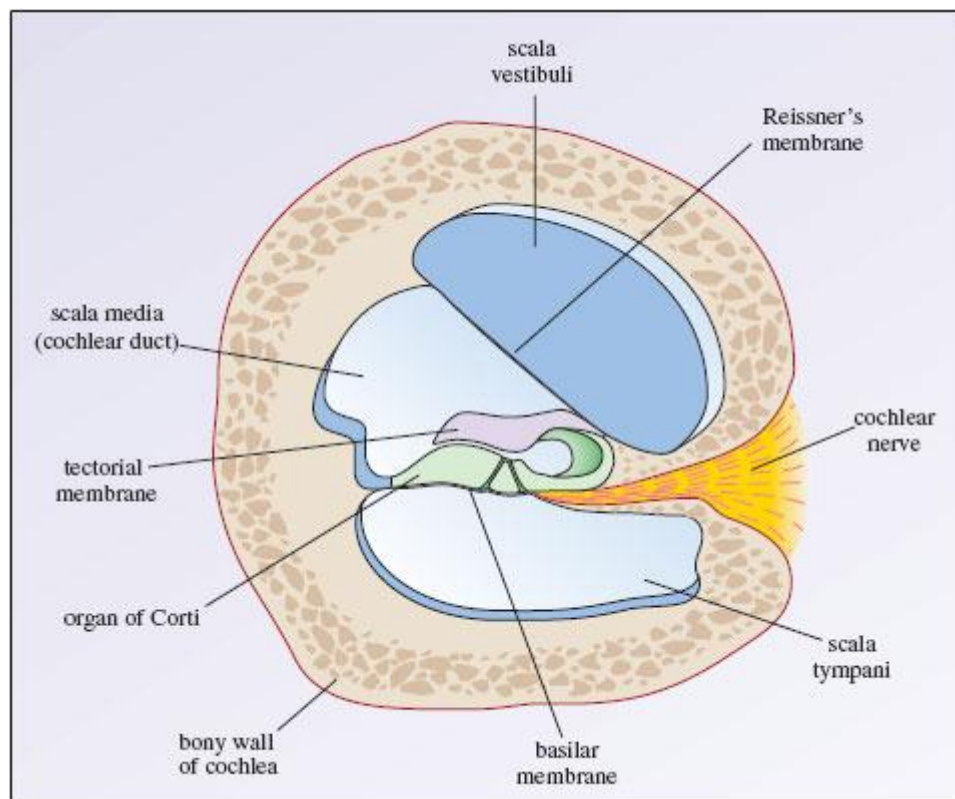


Figure 20 Cross section of the cochlea showing the three fluid filled canals. The inner hair cells of the basilar membrane discharge in response to friction with the tectorial membrane.

<http://labspace.open.ac.uk/mod/resource/view.php?id=432270>

Movement of the oval window membrane causes a pressure wave in the fluid of the canal which travels away from the stapes. Because the two canals are

connected, the round window is distended outwards by the pressure wave, absorbing the remaining energy and stabilising the pressure within the two canals.

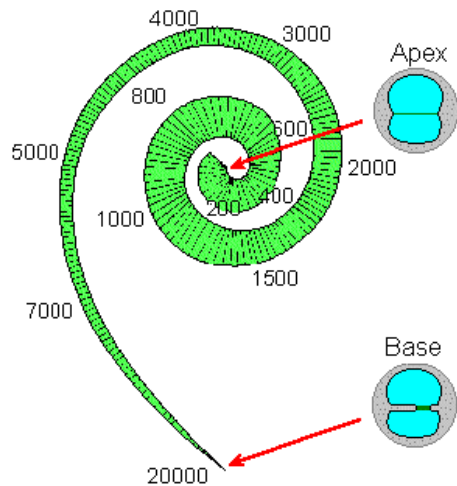


Figure 21 The difference in cochlear cross-sections at the apex and base (right). On the left is shown frequency selectivity (Hz) as a function of basilar membrane location.
(<http://www.sissa.it/multidisc/cochlea/utls/basilar.htm>).

The pressure waves passing through the two canals cause a wave-like movement in the basement membrane. The point at which this wave peaks is known as the location of maximum membrane displacement. The location of maximum membrane displacement depends upon the frequency characteristics of the sound wave which produce the movement of the stapes. High frequency waves peak closer to the base and low frequency will peak nearer the apex (Figure 21). Attached to the basilar membrane within

the cochlear canal, the organ of corti, (Figure 22) contains two banks of receptor cells, known as the inner hair cells and outer hair cells.

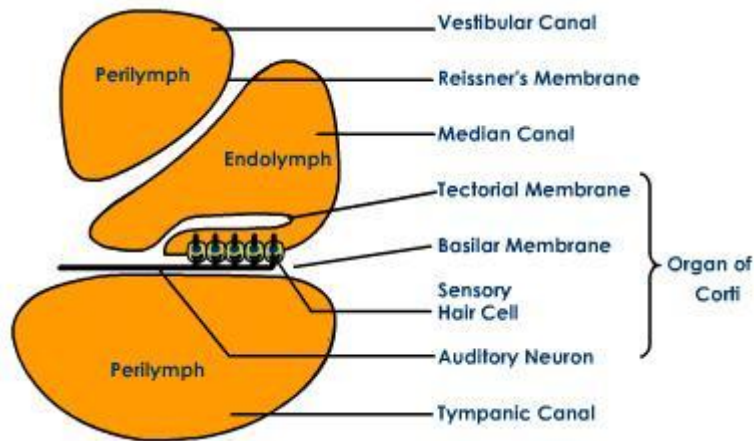


Figure 22 The organ of corti. <http://www.tutorvista.com/content/biology/biology-iv/nervous-coordination/ear-structure.php>

It is the inner hair cells of the basilar membrane which generate nerve impulses. Vibration of the basilar membrane causes friction between the hair cells and tectorial membrane; this friction causes the inner hair cells to produce a neural discharge. The outer hair cells are not active in transduction, but it is believed that they influence sensitivity and frequency tuning via efferent connections from the auditory centres of the brain (Moore, 2005).

Following discharge of the inner hair cells, the neural signals travel along the auditory nerve to the brain.

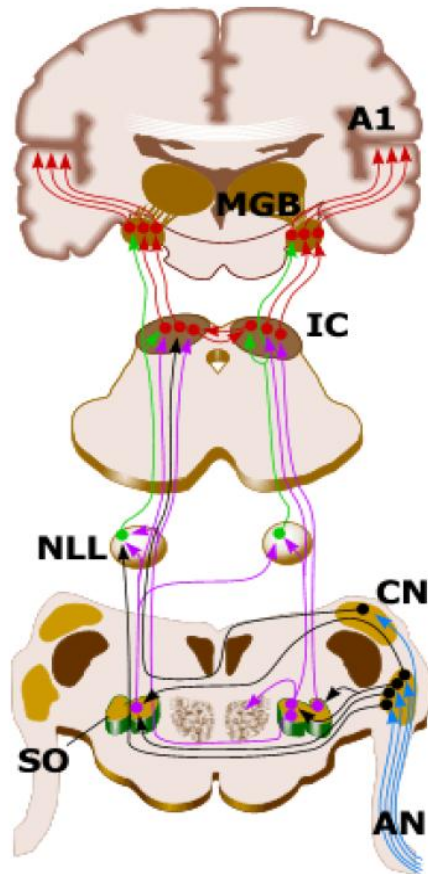


Figure 23 Diagram showing the main structures of the human ascending auditory pathway. Abbreviations are AN (auditory nerve), CN (cochlear nuclei), SO (superior olive), NLL (nuclei of the lateral lemniscus), IC (inferior colliculus), MGB (medial geniculate body), A1 (primary auditory cortex). Taken from <http://www.urmc.rochester.edu/smd/Nanat/faculty-research/lab-pages/KevinDavis/passways.gif>.

Figure 23 shows the main structures and synapses involved in the auditory pathway. The auditory nerve first projects to the cochlear nuclei of the midbrain. It is here that sound frequency processing takes place. Next, the auditory pathway projects from the ipsilateral cochlear nucleus to both

superior olive nuclei. This is the first stage in the auditory pathway where signals from both ears interact, a process which helps enable the auditory system to locate the origin of a sound in space (Boudreau et al., 1968, Hackney, 1987) These binaural interactions convey valuable information about the location of a sound source (Grantham, 1995, Sekuler, 1994). Nerve fibres leaving the superior olive are known as the lateral lemniscus. The nuclei project to the inferior colliculus at the midbrain region of the brainstem, as well as to the contralateral lateral lemniscus. The inferior colliculus receives input from both the ipsilateral superior olive and contralateral cochlear nuclei, although most of the inferior colliculus input is originates from the contralateral ear. The inferior colliculus also receives somatosensory afferent input, and input from the contralateral inferior colliculus. Prior to its termination at the primary auditory cortex the auditory pathway projects to the medial geniculate body of the thalamus. Here, the dorsal division of the medial geniculate body receives visual and somatosensory afferent input as well as auditory information. The pathway progresses to the ipsilateral primary auditory cortex, at the superior temporal gyrus. Cortical processing of auditory events is performed in the auditory cortex and associated secondary auditory cortical areas (Figure 24), although some of the auditory processing has already been carried out before the auditory information reaches the auditory cortex.

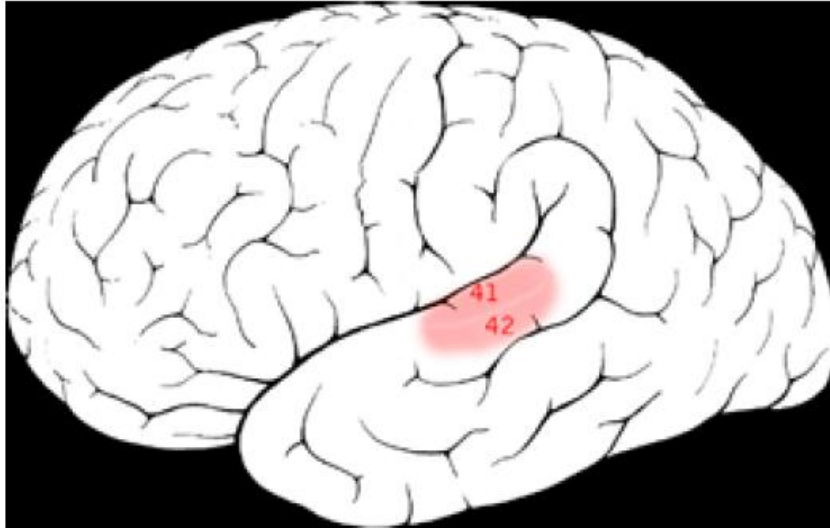


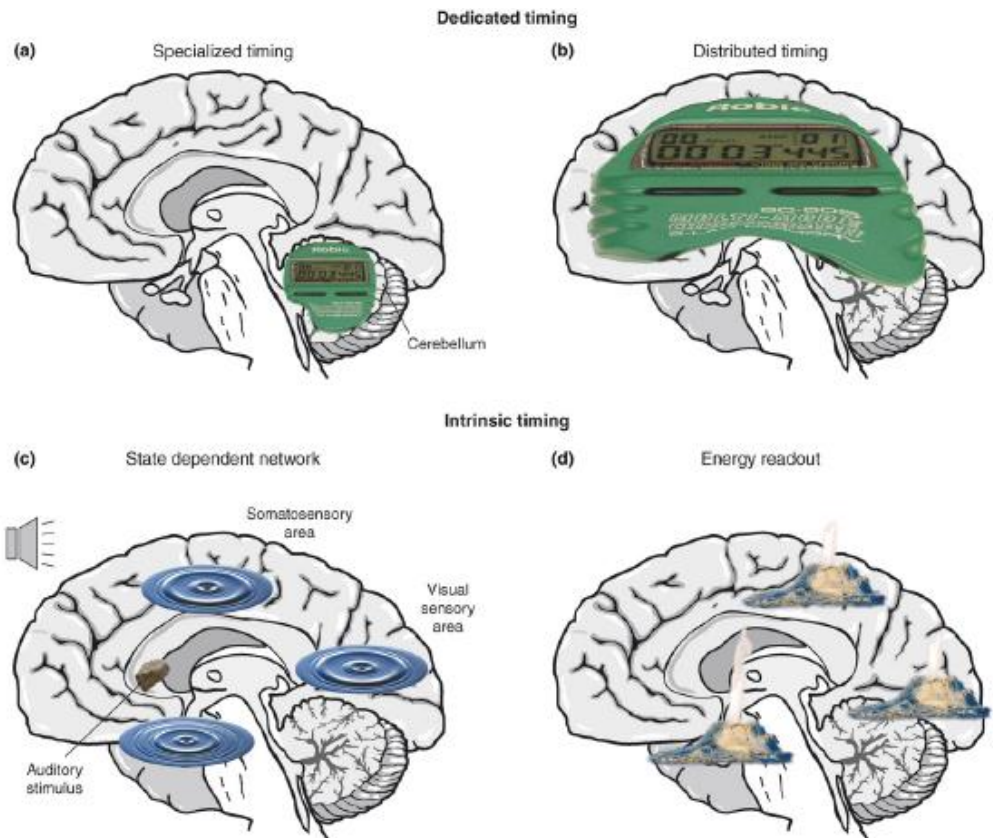
Figure 24 Sagittal view of the human brain illustrating the location of primary auditory cortex (A1; Brodmann areas 41 & 42). The occipital lobe (containing V1) is at the right of the diagram. Taken from [http://en.wikipedia.org/wiki/Image: Brodmann_41_42.png](http://en.wikipedia.org/wiki/Image:Brodmann_41_42.png).

This is in contrast to the visual system where the overwhelming majority of processing does not occur until the primary visual cortex. Consequently, the subcortical auditory pathway is extremely complex in comparison with the visual pathway (Hackett et al., 2003). Auditory input to the multisensory regions of the superior colliculus is from the inferior colliculus, superior olive, and lateral lemniscus (Edwards et al., 1979). Many of the experiments in this thesis involve the judgment of auditory and visual durations. Interestingly, although in the spatial domain vision is more reliable than audition, in timing judgments the reverse is the case.

Chapter 3

Models of time perception

Timing is an essential element for almost all tasks performed by humans and animals alike. Yet our understanding of the way time is processed lags well behind that of auditory or visual space. There is, however, a body of experimental data which may be used to make inferences about the nature of temporal perception. These inferences have given rise to a number of models for the way time is processed. These models may be divided into two groups: Those in which time is processed by a “dedicated” mechanism and those in which the passage of time is revealed via the intrinsic neural activity involved in representing the stimulus to be timed. Dedicated mechanisms may be centralised (in a particular brain area) or distributed across different areas of the brain. Intrinsic timing models use the state of a network of sensory neurons to encode time or time may be deduced from the amount of neural energy expended in response to a stimulus (Figure 25). The most enduring class of the dedicated models is commonly known as pacemaker/accumulator or scalar expectancy theory (Gibbon, 1977).



TRENDS in Cognitive Sciences

Figure 25 Dedicated (top) and intrinsic (bottom) models of timing. (a) A dedicated system for the measurement of temporal extent situated in the cerebellum. Alternative suggested loci are the basal ganglia, supplementary motor area or the right prefrontal cortex. (b) a dedicated system in which timing is distributed across multiple brain areas (see also Figure 29). (c) Localised modality specific timing in which spatial patterns of neural activity represent the passage of time. (d) Expended neural energy used to represent a stimulus is used to measure the passage of time (Ivry et al., 2008).

3.1 Pacemaker/accumulator theories

The nervous system has been found to operate using rhythms over a number of different time scales. For example, circadian rhythms operate over a 24 hour cycle and regulate sleep, hunger and other metabolic and behavioural systems (Rusak et al., 1979). In the spinal column, rhythmic pattern generators have also been shown to produce sequences of motor activity in lamprey eels (Grillner et al., 1991). This rhythmic quality has led scientists to suggest models in which pulses or oscillations are used to measure the passage of time.

Probably the most often quoted group of models for the perception of time are pacemaker/accumulator models. Creelman (1962) and Treisman (1963) are cited as being the first to suggest models in which oscillations or pulses are used to measure the passage of time. Creelman's theory was that a pacemaker would produce pulses and a judgment regarding the total length of any given duration is based upon the total number of pulses produced during the interval. Treisman proposed that a pacemaker produces pulses at a fixed rate, but this rate may be influenced by an observer's level of arousal. The pulses are counted and transformed into their log values, then compared with other similar events (Treisman, 1963). The log transform of pulses to internal duration produces a larger just noticeable difference as duration increases and thus the model may be used to explain Weber's law for time perception which states that the minimum difference between two durations

which may be detected is proportional to the duration itself (Treisman, 1963, Allan, 1979, Matell et al., 2004). A more detailed explanation of Weber's law may be found in section 3.1.1.

Arguably the most widely cited pacemaker model, (Church et al., 1991, Church et al., 1994, Rakitin et al., 1998) is referred to as the Scalar Expectancy Theory (Gibbon, 1977). The theory consists of three stages. The first stage consists of a pacemaker, which emits periodic pulses and is linked to an accumulator which counts the pacemaker's pulses when it is "switched on" by attention to a stimulus. Information from the accumulator feeds into a working memory (short term) of the duration and also into a more long term (reference) memory bank of similar events which is updated after each event. The temporal information from the working and reference memories are then compared to ascertain whether the relative difference between the present duration and the reference memory duration is below a particular ratio. This model describes data obtained by peak interval studies very well. Typically, these studies involve training an animal to produce a response to a signal after a specific period of time by the use of a reward for a correct response. Animals are capable of performing these tasks over periods of several tens of seconds. The responses produced tend to form a Gaussian-like response function, centred on, or very close to the interval which produces the reward. This may be predicted by the model if the animal holds a specific number of pulses representing the desired interval in its reference memory. When the required number of pulses is passed to the accumulator from the pacemaker a response is triggered and the animal collects its reward. It is this comparison, together with predicted fluctuations in pulse production, memory,

and decision making, that produces the scalar property and the Gaussian type responses seen in these studies (Church et al., 1991, Church et al., 1994, Rakitin et al., 1998).

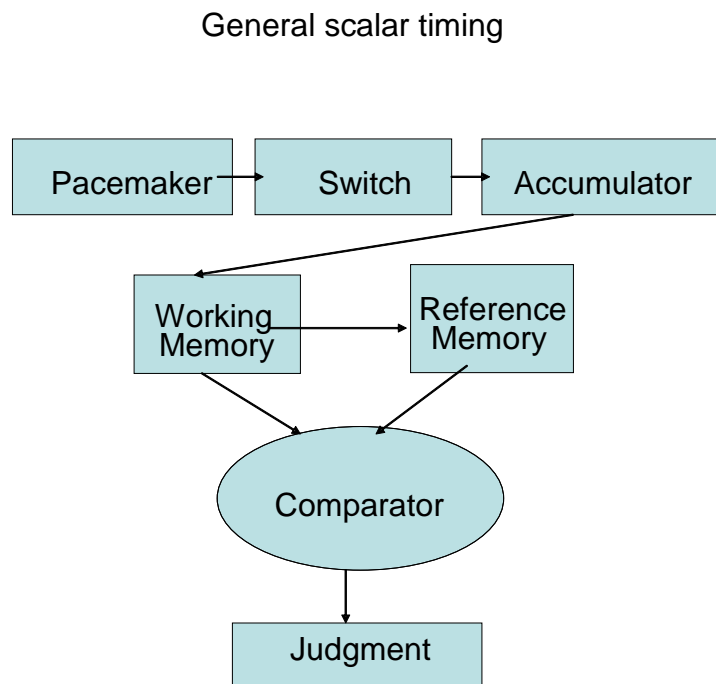


Figure 26 A schematic showing General scalar theory. Pulses are produced by the pacemaker. Attention to a stimulus trips the switch and sends the pulses to be counted by the accumulator. When the stimulus ends the accumulated pulses pass to the work working memory and are compared with memories of a particular event or a range of similar events. Finally a temporal judgment is made on the basis of this comparison (Gibbon et al., 1984).

A schematic of Scalar expectancy theory is reproduced in Figure 26. (Gibbon et al., 1984, Gibbon, 1977).

The Gaussian-like responses produced in studies involving animals and peak interval procedures are well explained by Scalar expectancy theory. Findings such as these and the need to explain the relationship between timing data and Weber's law make the pacemaker accumulator model an attractive description of temporal processing.

A variation on this theme proposes two types of attention processing which run concurrently: temporal information processing and non-temporal information processing. Perceived duration is said to depend upon the way attention is divided between these two processes. This model predicts that as more attention is given over to the non-temporal aspects of a stimulus, less is left for temporal judgment since the two processes share a common attention pool (Thomas et al., 1975).

A number of studies have found that temporal discrimination is adversely affected and subjective duration reduced when subjects are asked to perform a concurrent task (such as count the number of animal names in a random series of words or count the number of wrong notes in a short piece of music) and also by the difficulty of this task. This may be explained by the division of attention and the consequent missing of pulses produced by the pacemaker and cues from the stimulus. Fewer pulses are counted and perceived duration is reduced. (Macar et al., 1994, Brown et al., 2002, Zakay, 1993, Zakay et al., 1997).

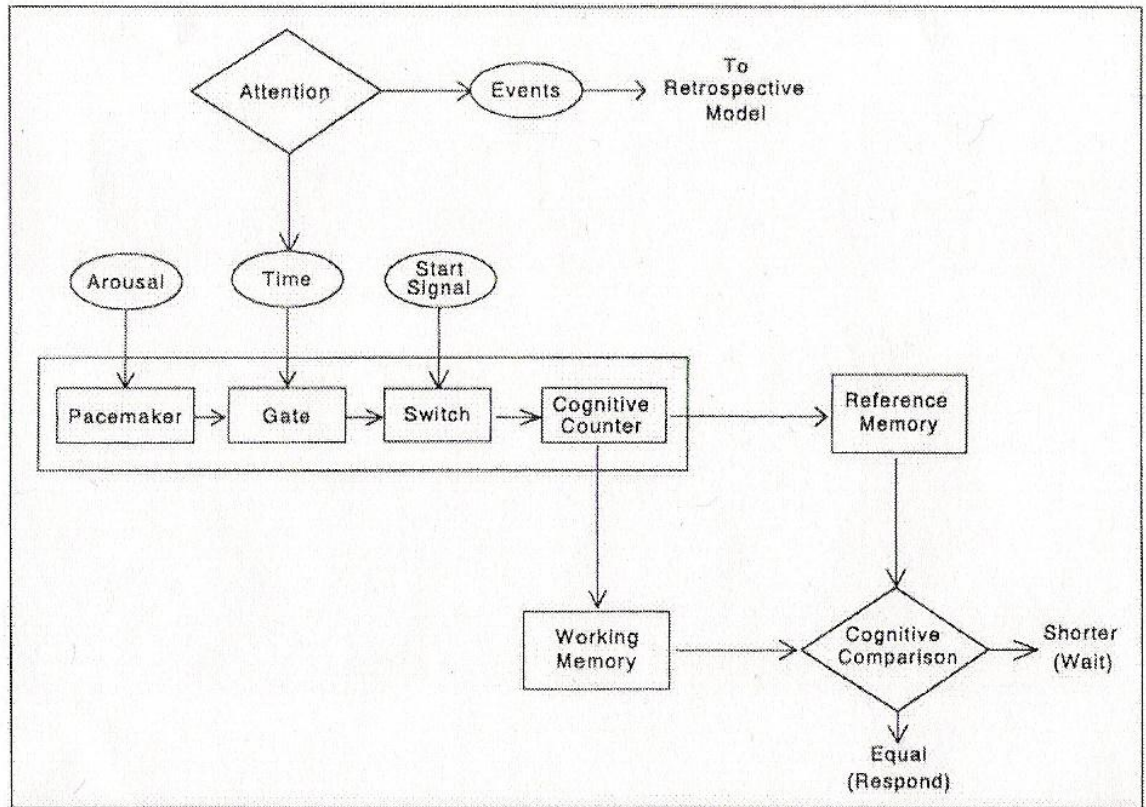


Figure 27 The attentional gate model. The pacemaker produces pulses which are influenced only by arousal. When attention is given to time, the attention gate opens and the pulses are sent downstream. At the start of the duration a switch allows the pulse stream through to the counter from whence it is transferred to the working memory. When the duration has ended the switch closes and the pulse total is sent to the reference memory for comparison. In the example shown above the task is one of reproduction. In the case the subject waits until the correct number of pulses are counted before responding that the duration has passed (Zakay et al., 1997).

A model which builds upon Scalar Expectancy Theory and the attention model of Thomas and Weaver (1975), is known as the attentional gate model (Zakay et al., 1997) and is reproduced in Figure 27. This model is said to explain how temporal estimates are arrived at when different psychophysical methods are used in timing experiments. For example, if a subject is required

to make a verbal estimation, the pulse total is compared to other pulse totals in the reference memory, which have verbal labels for different durations. In the case of a reproduction task the pulse total is held in the reference memory for comparison during reproduction until a similar number of pulses are counted at which point reproduction is complete. A two alternative forced choice task would work in a similar way with the first stimulus pulse total held in reference memory for comparison with the second stimulus.

The attentional gate model describes an important factor in time estimation, that of prospective and retrospective time estimation. Under prospective conditions subjects know that they are required to judge duration prior to its presentation and therefore focus their attention on time. In retrospective conditions the subject may be unaware that a timing judgment is required until the stimulus has passed, so the tendency is to focus on the stimulus alone and the time estimation is produced using contextual information stored in the memory. Prospective time judgments depend on arousal level (greater arousal increases the rate at which pulses are produced), the amount of attention given over to time (greater attention allows more pulses through to the cognitive counter via the gate), and the speed of onset of pulse counting which may also be affected by attention. According to the attentional gate model the most important factor in duration judgments is whether the judgment required is prospective or retrospective since this determines where the locus of attention lies (on the stimulus itself or on its duration) (Zakay et al., 1997).

These models provide a potential explanation for much of the empirical data found in time perception studies. A centralised mechanism for time is

appealing because it explains how subjects are able to compare durations from different modalities (Grondin et al., 1991b, Ulrich et al., 2006(b), Westheimer, 1999). Pacemaker/accumulator models can also be used to predict the relationship between data from a variety of different experiments and Weber's law (Creelman, 1962, Treisman, 1963, Getty, 1975, Getty, 1976, Allan, 1979, Wearden et al., 2007, Wearden et al., 2008). They explain how variations in timing may occur due to differences in the level of attention (Brown et al., 1999, Zakay, 1993, Zakay, 1998), arousal (Treisman, 1963, Ulrich et al., 2006(a), Droit-Volet et al., 2009, Zakay et al., 1997) and how stimulus characteristics have been found to influence perceived duration (see Chapter 4).

However, although examples may be found of neurons with regular rhythms (5-15 Hz) (Meck et al., 2008) in the cortex, there is to date no credible example of a neural mechanism which saves neural pulses over several minutes, although this may be possible for shorter periods (Matell et al., 2004, Gibbon et al., 1997), so presumably pacemaker accumulator mechanisms would only be feasible for shorter durations. Furthermore, the site of this putative central timing mechanism with its pacemaker, counter, gateway etc has yet to be confirmed over the 50 years since pacemaker/accumulator models first appeared. Likewise we have no physical evidence of mechanisms which discriminate between prospective and retrospective timing or mediate timing based on differing levels of attention.

To sum up, whilst pacemaker accumulator models do not as yet describe verifiable mechanisms of the brain, they are able to predict many of the findings reported in the timing literature.

3.1.1 Weber's law

Weber's law states that there is a relationship between a quantity of something (e.g. length, weight, time) and the amount by which it must be changed before a subject can detect the change. Experiments which measure this relationship are known as discrimination threshold experiments. The amount to be added before the change is detected is often referred to as the just noticeable difference (JND). According to Weber, the JND between two weights is approximately proportional to the mass of the weights. If the base value of a weight is increased, the amount to be added before the difference may be detected will increase in a linear fashion. In the case of time perception, if the initial length of a duration is denoted by D and the duration which needs to be added in order for the difference to be detected by ΔD , then $\Delta D/D = K$ where K is a constant value regardless of the value of D . This value K is known as the Weber fraction. An alternative measure often used in timing studies, is the coefficient of variation (cv). The coefficient of variation is proportional to the Weber fraction and is given by $cv = \frac{\sigma}{t}$, where σ is the standard deviation of the duration judgments (reproduced duration or duration estimates) and t is the mean of the judgments. The Weber fraction is often used to compare duration discrimination across differing time scales and modalities. However duration discrimination may be influenced by factors other than the duration magnitude. Getty suggested that duration discrimination variability should be divided into two components, one which is relative to duration length and one which is independent of stimulus length (Getty, 1975, Getty, 1976).

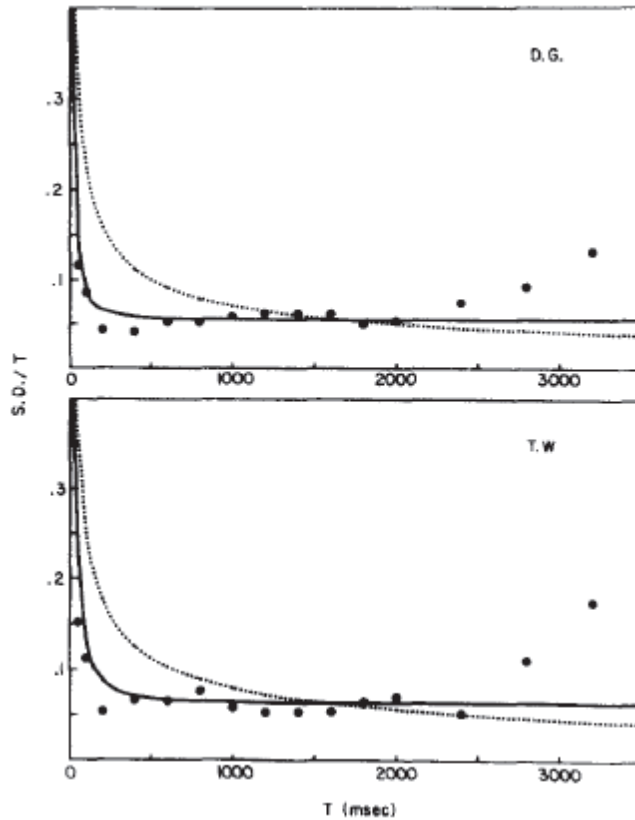


Figure 28 Data obtained from an auditory discrimination experiment run by Getty, along with predictions of Creelman's counter model (dashed line) and the generalised Weber's law (solid line). The X axis denotes Weber fractions and the Y axis, the standard durations tested. It can be seen that Weber fractions cease to follow the predictions beyond 2,000 – 2,500 ms (Getty, 1975).

It is assumed that as duration increases, the relative effect of this duration independent variability decreases, thus Weber's law may not be constant across all values. At short durations, duration independent variability is very high. However its effect rapidly reduces and appears to level off at around 150-200ms. This generalised form of Weber's law successfully predicted data for durations up to 2 seconds (see Figure 28, also (Getty, 1975) for a mathematical explanation). Beyond this level a rise in Weber fraction was found. However since shorter durations are thought to be mediated by

different mechanisms than those of several seconds (Buhusi et al., 2005, Buonomano et al., 2002, Lewis et al., 2003) this is perhaps not surprising.

The method used in psychophysical experiments can have an influence on the results obtained. For example, it has been shown that when measuring the just noticeable difference, if the standard duration is always presented before the comparison (which may be longer, shorter or equal to the standard), and the subject is asked to indicate whether the comparison is longer or shorter, a smaller JND is obtained than if the order of the standard and comparison stimuli is randomised. The authors of one such study (Ulrich et al., 2009) conclude that it is the order in which the standard stimulus is presented which produces this effect and that the comparison of studies which use different methods to obtain the JND should be treated with caution, especially when comparing Weber fractions. This is because the Weber fraction could differ between trials as a direct result of the methods used. It has been suggested that the reason for this difference in discrimination performance is due to subjects' use of an internal representation of the previous standard durations, as well as the actual duration, when the standard is always presented first (Lapid et al., 2008)

3.2 Rival models for time perception

Given that the neural locus of a pacemaker / accumulator model has not yet been identified, it seems reasonable to ask whether other candidate models for neural timing exist.

3.2.1 Distributed timing

It has been suggested, that the different components of a pacemaker/accumulator could be distributed through different brain areas, rather than being a centralised system in a particular location (Hazeltine et al., 1997, Meck, 1996) (see Figure 29). This proposal is based upon the findings of a number of pharmacological and lesion studies which suggest that the basal ganglia system forms a pacemaker/accumulator system driven by dopamine. The temporal reference memory and attention mechanism, it is suggested, lie in the frontal cortex which is dependent upon acetylcholine. The model proposes that the substantia nigra (located in the midbrain; the substantia nigra serves as a input to the basal ganglia and supplies the striatum with dopamine) produces regular pulses via the striatum (a subcortical area of the forebrain and major input station of the basal ganglia system) which acts as a gateway to the Pallidum (a component of the basal ganglia) which fulfils the role of the accumulator. The accumulated information is sent to the Thalamus (situated between the cerebral cortex and midbrain; the thalamus relays sensory and motor signals to the cerebral cortex) which in turn forwards it on to the frontal cortex for comparison with representations of similar events (Meck, 1996, Hazeltine et al., 1997). Whilst this model is feasible, it is worth noting that multiple areas of the brain have been implicated in pharmacological, lesion and imaging studies (see neural correlates section).

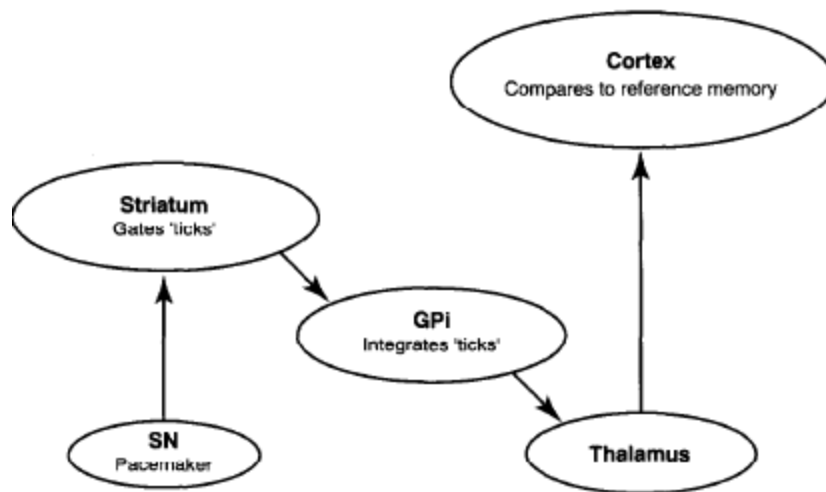


Figure 29 A distributed pacemaker/accumulator model. The Substantia Nigra produces a regular pulse which passes through the striatum to the Pallidum. The internal segment (GPI) acts as an accumulator. Information from the Pallidum passes via the thalamus to the cortex, where it is compared to memories of similar events. This model has been used to account for data produced from animal timing studies which used durations of 20 seconds or more (Hazeltine et al., 1997).

Other models differentiate between automatic (beyond conscious control) and cognitive timing (interval estimation-see introduction) and suggest that different mechanisms (Ivry, 1996), or different areas of the brain (Lewis et al., 2003) are employed depending on the timing task to be undertaken. Automatic timing tasks are said to take place mainly in motor areas of the brain, whereas cognitively controlled timing, it is suggested, takes place in prefrontal and parietal regions (Lewis et al., 2003). However the precise areas and mechanisms remain a mystery.

3.2.2 Labeled lines/Channels

Unlike classic pacemaker models, these models represent different intervals using different channels (Ivry, 1996, Church et al., 1990, Miall, 1996). Labelled line models are based on a mechanism in which neurons respond maximally to a particular duration. Duration tuned neurons have been reported for both audition (Potter, 1965, Leary et al., 2008, Faure et al., 2003) and vision (Duysens et al., 1996). In order for labelled lines to react preferentially to a specific duration they require some kind of time dependent property. Suggestions for this property include delayed spiking, slow biochemical reactions, and oscillators. They would also require a mechanism for relating the onset and offset of a stimulus to the time dependent property (Buonomano et al., 2002). These channels could be dedicated to a particular duration only (Figure 30) or maximally sensitive to its preferred duration with sensitivity dropping off progressively for durations further away from this duration (Figure 31). Time is then encoded using the responses across the different detectors (Walker et al., 1981b, Buonomano et al., 2002, Ivry, 1996).

The models propose that timing neurons work in a similar way to those involved in processing orientation or spatial frequency (Buonomano et al., 2002).

..

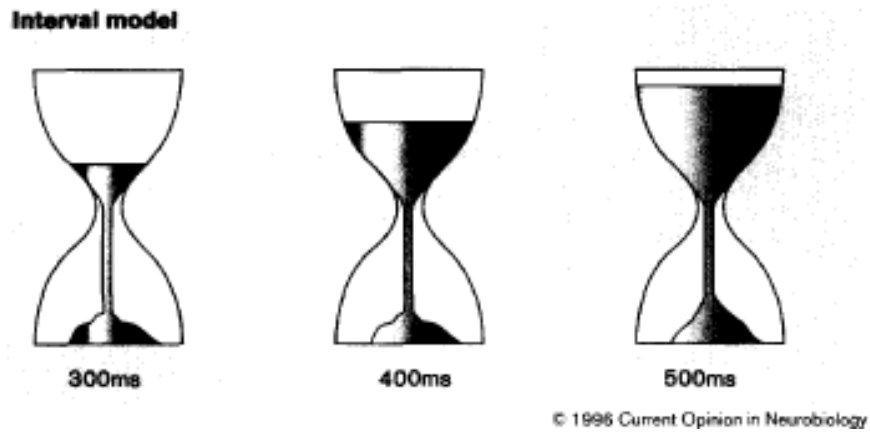
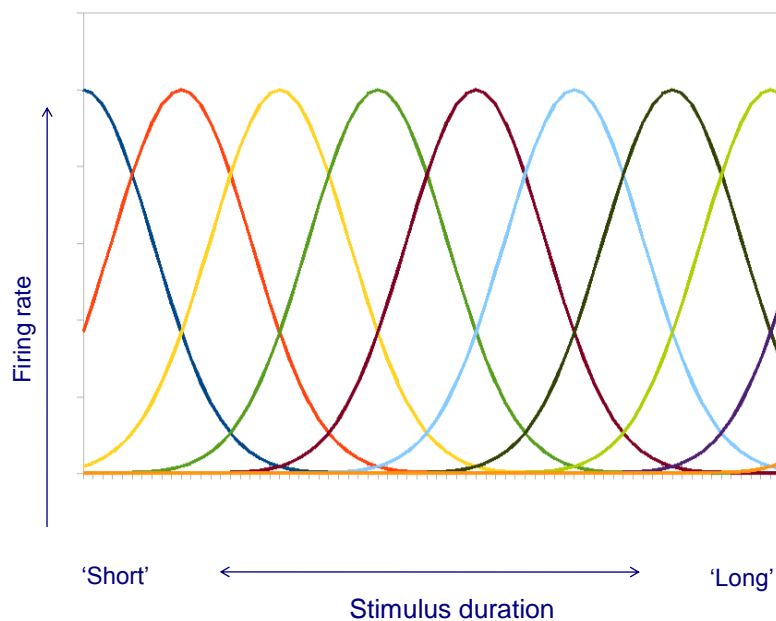


Figure 30 Interval (channel) based models posit that different durations are represented by separate elements each of which represent a particular duration (Ivry, 1996).



12

Figure 31 A range of duration sensitive “channels” each of which is maximally sensitive to its own duration, but which may also react in a limited fashion to neighbouring durations. Duration is estimated by comparing responses across channels.

Another variation of channel based model suggests groups of oscillating neurons with different frequencies which spike at a point in each cycle. Any

two of these spiking oscillators will periodically spike simultaneously. Thus these two oscillators will have their own individual cycles as well as a longer cycle produced when they both spike together.

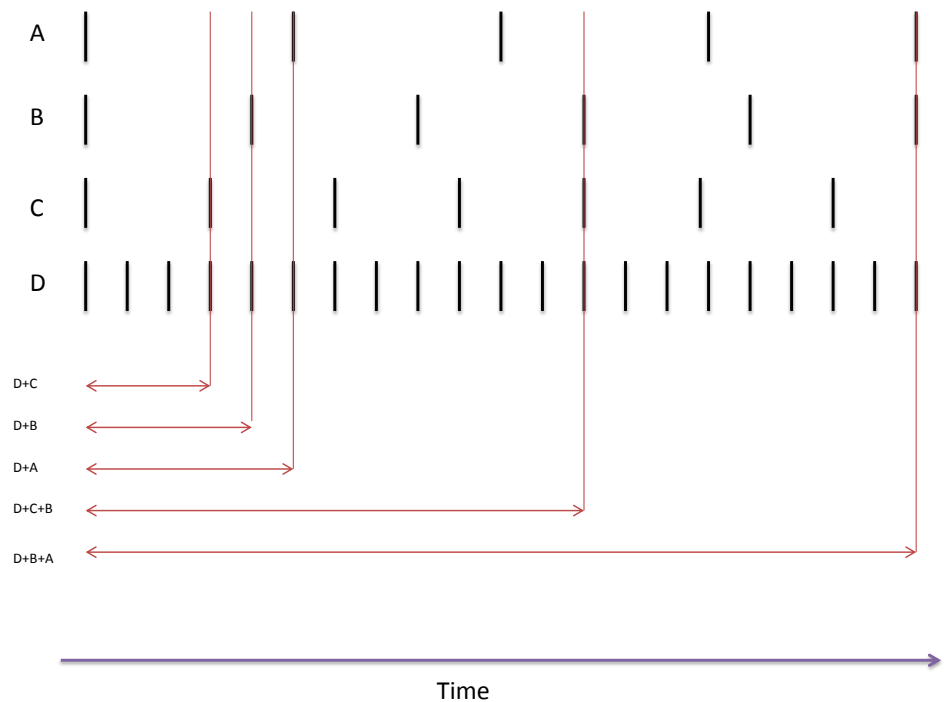


Figure 31 Schematic showing the activity of 4 oscillators. Different durations may be encoded by different combinations of neurons firing simultaneously. The shortest duration which may be encoded equals the rate of the fastest oscillating neuron. The longest duration depends upon the number of oscillators, and the distribution of the oscillator frequencies. Simulations involving 250 or 500 pacemakers have produced an upper limit of 20 seconds (Miall, 1996).

The beat frequency of any group of oscillating neurons will be the lowest common denominator of their individual beat frequencies. In this way a relatively small number of oscillators would be able to represent a wide range of intervals (Figures 31 and 32). The beats which match the onset and offset of the duration in time encode the interval length (Miall, 1996).

A salient feature of these models is that groups of neurons are dedicated to a particular duration. These groups of neurons may be isolated; the model may be tested by using techniques such as duration adaptation.

3.2.3 Intrinsic timing models

A relatively recent group of models propose that rather than time being processed by specialised mechanisms, it is in fact an inherent part of sensory neural dynamics. In other words our sense of time is a natural by-product of ongoing neural responses to physical stimuli. This group of models is often referred to as intrinsic timing. In some intrinsic models timing arises within different modalities as groups of neurons respond to a stimulus. For example, the timing of an auditory stimulus would be encoded in auditory regions of the brain, whereas a visual stimulus would have its duration encoded in visual areas (Buonomano, 2000, Burr et al., 2007a). This theory is viable because sensory stimuli cause spatiotemporal patterns of action potentials which are relayed to the central nervous system.

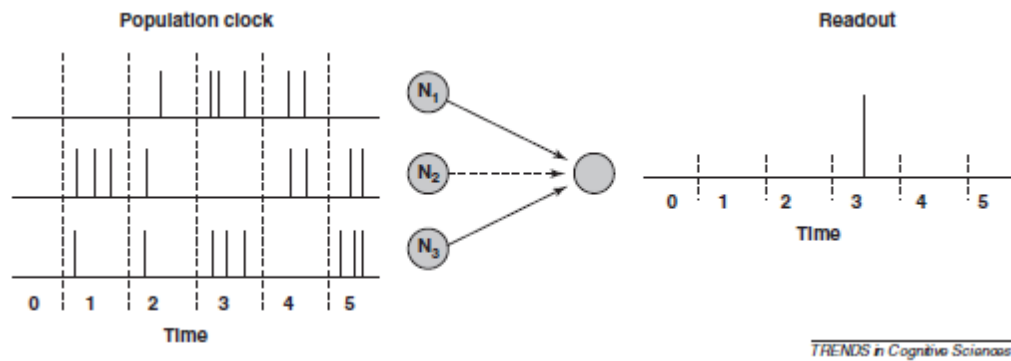


Figure 32 A simple population clock model consisting of 3 neurons. The neurons display a reliable pattern of firing. The activity of each neuron changes over time. Duration may be coded based on the activity across the neurons at any given point in time. For example, the code representing 3 units of time would be 3, 0, 3. After another unit of time has passed the time signature becomes 2, 2, 0 and so on (Buonomano et al., 2010).

These patterns change over time and so the state of a network of neurons will be different when a stimulus has ceased than it was at stimulus onset. Thus the state of the network provides a representation of the duration of the stimulus. This may be achieved by chain reaction to a stimulus within a population of neurons so that any particular point in time could be represented by the activity of a small group of neurons (Eagleman, 2008) or time could be encoded by a larger group of neurons where it is the unique pattern of firing across the group which enables the passage of time to be measured (Buonomano et al., 2010) (Figure 32).

An alternative theory states that perceived duration could be dependent on the amount of energy expended during the neural response to a stimulus. This theory has been suggested as a cause for the oddball effect (in which a

“unexpected stimulus” is perceived as being longer than a repeated “expected stimulus” of equal duration), in that when a stimulus is repeated the corresponding neural response is reduced. This phenomenon is known as repetition suppression and leads to less energy expenditure for repeated stimuli. Thus in a train of repeated stimuli the first stimulus is perceived as being longer than the subsequent ones (Rose et al., 1995).

Much of the supporting evidence for intrinsic timing comes from physiological experiments which demonstrate localised, sensory specific timing (Jantzen et al., 2005, Morrone et al., 2005, Buetti et al., 2008b). Examples of these are a transcranial magnetic stimulation experiment in which TMS was applied over V5. Performance in judging a visual stimulus duration was affected, whereas the judgment of an auditory stimulus was left unimpaired (Buetti et al., 2008a) and an fMRI study in which subjects were exposed to an auditory or visual rhythm which they subsequently tapped out (Jantzen et al., 2005). In the visual condition, activity was found to be high in V5 after the initial primer rhythm ceased in line with state dependent models where continuing sensory specific pattern of activity would be expected in order to provide a template for the tapping task. No such activity in V5 was found when the rhythm was primed with an auditory stimulus.

However, a number of studies have shown that if subjects train on discriminating a particular duration, performance improves and is transferred across modalities (Warm et al., 1975, Nagarajan et al., 1998) and from sensory to motor timing (Meegan et al., 2000). It would seem to be problematic to explain this via intrinsic models. Intrinsic models which suggest an early locus for the encoding of time have a further difficulty

explaining the results of Westheimer (1999). This study had subjects train in temporal discrimination using a stimulus in the left visual field. The gains from this training were found to transfer across to the right visual field, suggesting that time is encoded at higher level visual areas (Westheimer, 1999).

Chapter 4

Factors influencing time perception

4.1 Stimulus nature

Performance in the judgment of durations has been demonstrated to differ between the senses. A widely reported finding is that auditory duration discrimination thresholds are consistently lower than their visual counterparts. This is the case for both filled and empty intervals (Grondin et al., 1991b, Wearden et al., 1998).

In addition, for both filled and empty stimuli, perceived auditory durations are typically longer than the perceived duration of physically identical visual durations (Behar et al., 1961). This has been found for a wide range of durations (Goldstone et al., 1974, Wearden et al., 1998, Wearden et al., 2006). These findings may be explained in terms of the pacemaker accumulator hypothesis if we consider audition as causing an increase in the rate of pulses produced by a pacemaker. If this were the case then more accumulated pulses would equate to a greater perceived duration. Also, because each pulse produced by a faster pulse rate demarks a shorter period of time, finer distinctions may be made, resulting in the smaller JNDs found in the literature. In addition it has also been found that low level visual stimulus characteristics influence perceived duration and that under certain

circumstances perceived visual duration may exceed that of an auditory stimulus (see Chapter 8.10).

Filled intervals have been found to be judged as being longer than empty ones of the same duration (Allan, 1979). Filled intervals have also been shown to be judged with greater sensitivity than empty ones (Rammsayer et al., 1998), although there are some conditions in which this is not the case (Grondin et al., 1998). Also, an “empty” interval punctuated with flashes or beeps has been found to be perceived as being longer than a completely empty one (Goldstone et al., 1976, Allan, 1979).

With regard to inter-sensory bias, it has been shown that if a duration is demarked by transient bimodal (visual and auditory of physically equal duration), the duration is perceived as being the same or very similar to when it is marked by audition alone (Walker et al., 1981b). Welch and Warren (1980) propose that when faced with conflicting information from the senses, we give priority to the most appropriate sense for the task in hand. In this case, audition is more sensitive in temporal perception hence its dominance over vision (Welch et al., 1980a). When judging empty durations which have cross modal markers, performance has been shown to drop significantly (Rousseau et al., 1983, Westheimer, 1999, Rousseau et al., 1973). It was originally proposed that this points towards different timing mechanisms for intra modal and cross modal timing (Rousseau et al., 1973). However, a later study concluded that the differences found in discrimination thresholds were due to noise caused by the attentional shift between visual and auditory modalities required for the cross modal task. Therefore the findings are not in conflict with a supra modal timing hypothesis (Rousseau et al., 1983).

Westheimer (1999) finds that subjects' discrimination thresholds may be lowered with training and that this effect transfers between visual hemispheres and is therefore located beyond areas of retinotopic representation putting the locus of timing at late processing stage. Westheimer also suggests that the increase in thresholds found with cross-modal stimuli is due to difficulties in collating signals from different sensory areas with differing neural firing patterns and different sources of background noise (Westheimer, 1999).

Woodrow (1928) found that in the case of empty auditory intervals, the markers used can influence perceived duration of the silence between the sounds. Making the markers longer resulted in an increase in the judged interval (Woodrow, 1928). This effect was found to be more pronounced as the onset marker was lengthened.

A study using a pattern of eight flashing lights to mark an interval in a reproduction task found that the more ordered and simple the flashing pattern, the shorter the duration reproduced. This effect was found to be more pronounced with shorter durations (Bobko et al., 1977a).

A study using a variable number of box-like stimuli moving along an invisible predetermined pathway found that increasing the speed of the stimuli lengthens the perception of its duration but varying the number of stimuli had little effect (Brown, 1995). A subsequent paper (Kanai et al., 2006) using a variety of stimuli, found that when randomly moving black dots were used as a stimulus, the speed of movement had a direct effect on perceived duration. The same paper describes an experiment involving the use of an expanding

concentric grating in which temporal frequency was manipulated to see whether temporal frequency or speed produced the greater time dilation. The authors concluded that speed had little effect and that temporal frequency was the significant factor. However, a more recent paper used a vertical grating and concluded that the opposite was in fact the case (Kaneko et al., 2009). The authors suggest that the differing results were due to the different stimuli used in each study and that the use of concentric rings produced a variance in luminance over the interval being judged. This variance produced a flicker effect which could have produced the time dilations found rather than them being as a result of temporal frequency. If either speed or temporal frequency can be shown to be totally responsible for the time dilation found by these studies, then this may have implications for the locus of the effect. If speed is totally responsible, this would place the locus for the effect at a later stage than if the effect is due to temporal frequency. Neurons tuned to speed independent of spatial frequency have been found in the macaque MT. In contrast to this, neurons in V1 prefer temporal frequency independent of spatial frequency (Priebe et al., 2006) .

4.2 Stimulus expectancy

Our perception of event duration appears to be modulated by our recent sensory history. For example, the perceived duration of the first stimulus in a stream of identical stimuli is typically overestimated (Rose et al., 1995). A related effect concerns the perception of infrequent or unexpected “oddball”

stimuli (see Figure 33) whose perceived duration is expanded relative to that of expected or frequent “standard” stimuli (Tse et al., 2004). It was initially suggested that this “subjective time dilation” increased the perceived duration of oddballs by approximately 30-50% (Tse et al., 2004). However, subsequent studies have suggested that this figure was grossly overestimated (Seifried et al., 2010), revealing a more modest expansion of around 10% (Ulrich et al., 2006(a), Chen et al., 2009, Pariyadath et al., 2007, van Wassenhove et al., 2008). The effect seems to be most robust for stimuli that are expanding in size, i.e. looming or approaching (Tse et al., 2004, van Wassenhove et al., 2008, New et al., 2009) and can be eliminated (New et al., 2009) with contracting or receding oddballs. The effect is reduced (Tse et al., 2004) or reversed (van Wassenhove et al., 2008), when a static oddball is presented within a stream of expanding standards. This implies an ecological “alerting” function in which an organism may respond to a possible threat more quickly and is consistent with reports of time slowing down in threatening situations (Campbell et al., 2007, Stetson et al., 2007). Inconsistent with this explanation, however, is the fact that similar effects have been reported for stationary stimuli (Chen et al., 2009, Tse et al., 2004, Pariyadath et al., 2007).

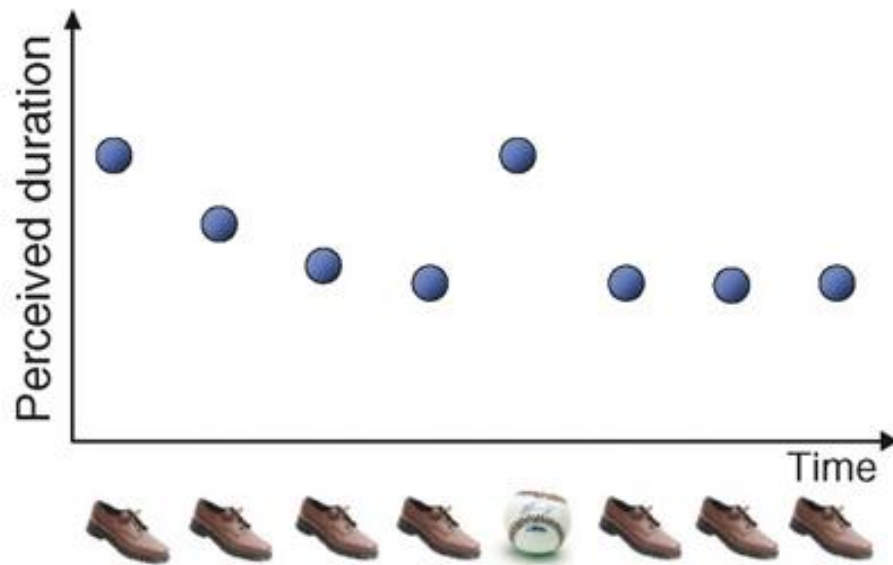


Figure 33 The initial stimulus and “oddball” appear to be expanded in duration compared to other stimuli in the series. This may be due to their unexpected nature or it may be that the other stimuli in the series are showing a contraction perceived duration due to repetition suppression of neural firing rates (Pariyadath et al., 2008).

There are two main competing explanations of subjective time dilation in relation to the oddball effect. The arousal theory claims that the alerting effect of an oddball causes a central internal pacemaker (Creelman, 1962, Treisman, 1963) to speed up, resulting in a subjective prolongation of time (Seifried et al., 2010). Alternatively an increase in attention could lead to more pulses being counted. These explanations receive support from the finding that subjective time dilation is a global phenomenon affecting the whole visual field, not just the oddball or its immediate surround (New et al., 2009). The centralised arousal theory is, however, difficult to reconcile with several experimental results: multisensory versions of the subjective time

dilation show asymmetric transfer between senses (Chen et al., 2009, van Wassenhove et al., 2008); the expansion of perceived duration can be generated with oddballs that are entirely predictable (van Wassenhove et al., 2008) and the fact that supposedly 'emotive' stimuli do not result in a greater expansion of perceived time (Pariyadath et al., 2007).

The information processing theory (Tse et al., 2004), on the other hand, proposes that the rate at which information is processed acts as the pacemaker component of our timekeeping system. In other words, 'bits' of information act as a counter with which we estimate the passage of time. This model suggests that the additional processing resources brought to bear for novel stimuli increase the overall rate at which information is processed, and the greater number of bits processed per unit time leads to an expansion of perceived duration. A related model is the "coding efficiency" model (Eagleman et al., 2009) where perceived event duration is directly related to the neural resources expended during the event's processing by the nervous system. In this model, repeated presentations of the expected or 'standard' stimulus leads to progressively more efficient encoding of this stimulus – a phenomenon termed repetition suppression (Grill-Spector et al., 2006, Henson et al., 2003) - such that, on re-appearance, reduced neural activity levels induce a perceived *contraction* in the duration of the *standard*, relative to the non-suppressed oddball stimulus. This same mechanism could also explain the "novelty" effect of Rose and Summers (1995).

The coding efficiency hypothesis arose following a series of experiments examining the oddball effect and the arousal explanation (Pariyadath et al., 2007). The authors increased the emotional impact of the oddballs in their

experiments by using emotionally charged images, such as growling dogs and spiders. They found that there was no increase in duration expansion with these oddballs when compared to the more neutral oddballs used in earlier experiments. This led them to conclude that if attention is the crucial element in the oddball effect it must either saturate at around 15% or the oddball effect is caused by some other mechanism. The authors propose that the latter is in fact the case. This suggestion is based the findings in the electrophysiology literature that repeated presentations of the same stimulus lead to a reduction of neural responses. Unexpected stimuli with non-suppressed firing rates would therefore appear longer by comparison (Pariyadath et al., 2007). However, since attention has been shown to increase the firing rate of cortical sensory neurons (Chik et al., 2009, McAdams et al., 1999, Moran et al., 1985) and also the coherence of neural firing (Doesburg et al., 2008, Fell et al., 2003) the two ideas need not be mutually exclusive. As attention to familiar stimuli wanes, neural firing rates may reduce, a new stimulus would then produce an increase in attention and an accompanying increase in firing rates. Using a variation of the flicker fusion paradigm, the authors sought to investigate the two hypotheses.

The experiment involved repeatedly presenting letters for very brief periods such that explicit temporal judgments were impossible. In some trials the letters were the same, whilst in others they were different. Because the letters were presented for very brief periods, more than one appeared to be on the screen at a time due to visual persistence.

Pariyadath and Eagleman found that fewer characters were perceived at a time during the repeated condition, due, they suggest, to a contraction of

visual persistence of the repeated stimuli (see Figure 34). They call this the proliferation effect and favour a neural suppression hypothesis because neural activity may be measured whereas as attention is more subjective and difficult to gauge. The authors suggest that a way to distinguish between attention and neural activity would be to conduct a range of experiments with repeated stimuli in which attention is manipulated by increasing the mental load of the experimental tasks (Pariyadath et al., 2008).

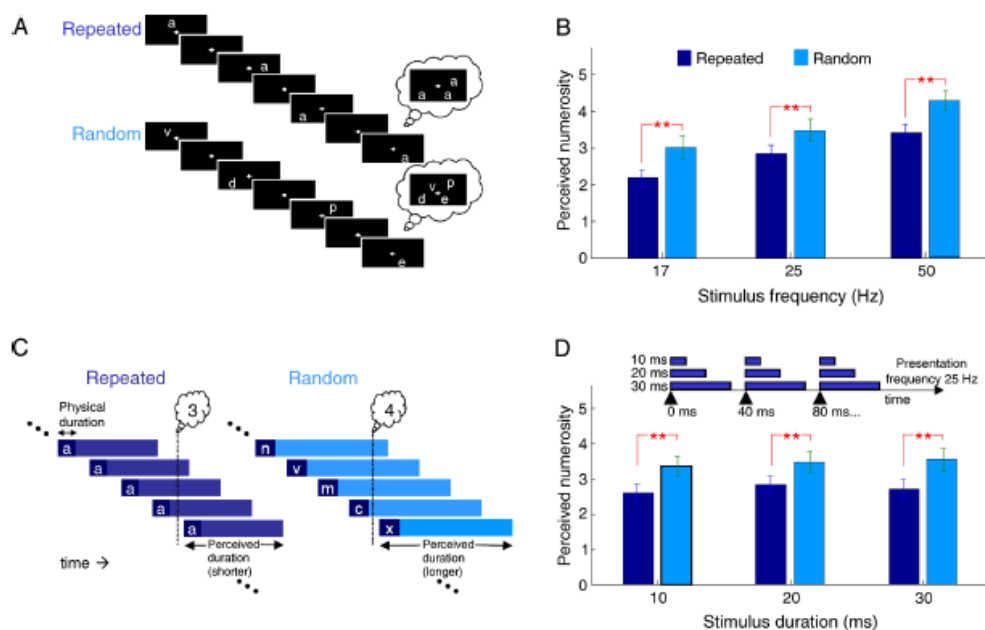


Figure 34 An experiment in which letters are presented to subjects for very brief durations. A) Shows sample sequences of stimulus presentations and the perceived numerosity when the letters are repeated and when they are different. B) Shows the number of letters perceived when they are repeated and when they are random. C) The authors suggest that the visual persistence of the repeated letters is reduced relative to the persistence of the random stimuli. D) When letters were presented for different physical durations whilst keeping the presentation frequency constant, no significant difference in perceived numerosity was found (Pariyadath et al., 2008).

(** = $p < 0.01$, * = $p < 0.05$, paired t -tests) $n=31$. Error bars SEM .

Although arousal and information processing models offer appealing explanations as to why perceived duration is context dependent, a problem common to such models is an inability to explain the criterion adopted by the nervous system when deciding which events should be designated as ‘expected’ or ‘unexpected’. In other words, how ‘odd’ does an oddball stimulus need to be before its perceived duration is deemed to differ from its neighbours? The diverse nature of standard and oddball stimuli deployed makes inferences on this topic somewhat problematic. For example, oddballs have variously been defined by changes in geometric shape (Tse et al., 2004), stimulus size/intensity (Seifried et al., 2010, New et al., 2009, van Wassenhove et al., 2008), alphanumeric character and photographic image properties (Pariyadath et al., 2007), often altering multiple stimulus features simultaneously between oddball and standard trials. Although it has recently been proposed that high level factors play a role (Pariyadath et al., 2007), inferences as to the nature of this role are difficult without precise control over the stimulus parameters in question.

4.3 Attention

Attention may be considered as falling into two different categories. “transient” attention, also known as “exogenous” or “involuntary” which is an automatic response to a stimulus and beyond conscious control, and “sustained” attention, also known as “endogenous” or “voluntary” which is subject to a measure of conscious control (Remington et al., 1992). A number of papers

have suggested that it takes around 120ms-150ms for attention to be allocated to a new stimulus (Hikosaka et al., 1993, Nakayama et al., 1989). This finding was used to further investigate the role of attention in time perception by Tse et al (2004) who found that attention to a stimulus produced an expansion in its perceived duration. This effect does not become apparent until at least 120ms after stimulus onset and appears to be similar for auditory and visual stimuli (Tse et al., 2004). The authors suggest that it is an increase in the level of attention which is responsible for the oddball effect.

It has been suggested that for some low level tasks, attention may be controlled by separate mechanisms for different modalities. Alais, Morrone et al (2006) used a dual task to show that temporal discrimination thresholds were unaffected by a concurrent task in a different modality and that the expected reduction in performance only occurred when the distracting task involved the same sense (vision or audition) (Alais et al., 2006).

One major difficulty with the question of attention and its role in the perception of sub-second durations is the idea that they are processed using different mechanisms than durations of 1 second and above (Fraisse, 1984). The problem appears to be in deciding where one mechanism ends and the other begins. This is thought to be somewhere between 500ms and 1,000ms but may change when a conscious effort is made to judge a brief duration in timing experiments. Transient attention is said to peak at around 120ms-150ms from stimulus onset before declining as sustained attention takes over at around 525ms (Hikosaka et al., 1993). It is tempting to assume that these timings are somewhat fluid depending upon the context in which the stimulus

is being judged and that there may well be a certain amount of overlap between the two different timing mechanisms (automatic and cognitive).

4.4 Adaptation

The ability to adapt to changing conditions is a fundamental requirement for living organisms. Our senses are subject to constantly changing input and have developed strategies for adjusting to these changes in order that we maintain a useful perception of the world around us. In the case of the visual system a variety of methods are employed for adapting to different types of stimulus change including brightness, motion, orientation, spatial frequency and contrast. For example, one of the major challenges for the visual system is the wide range of light intensity it is required to process. Adaptation to different light levels is accomplished by changes in pupil size, duplicity (rods and cones sharing the task with intermediate light levels) and also by neural adaptation. This occurs because the visual system is capable of varying its response to light depending upon the average level of illumination across the retina, the important criteria being the relative light levels of different areas of the visual field not the absolute light level. Put another way, the same response in the visual system may be produced to different stimulus intensities if the relationship of the stimulus to the general level of light is the same. In this way the visual system is able to be more sensitive and cover a wider range of light levels than if responses were fixed according to absolute

light levels. This is known as lightness constancy (Tovee, 1996, Blake et al., 2006).

The use of adaptation has been a fundamental method in psychophysics and has produced many valuable insights into human perception. One famous example of this is described by Blakemore and Campbell (1969b). Adapting to low spatial frequency stimuli causes subsequently viewed medium spatial frequencies to appear higher than they actually are (a repulsive after-effect). The same is true if we adapt to a high spatial frequency – medium spatial frequencies appear lower (see Figure 35) (Blakemore et al., 1969b), Blakemore et al., 1970). Other examples of rebound effects following adaptation include the motion aftereffect and the tilt aftereffect. The motion aftereffect is sometimes referred to as the “waterfall effect” because it can be produced by looking at a waterfall. After staring at the waterfall for a period of time, if we shift our gaze to the bank it appears to drift upwards (Goldstein, 1958). In a similar fashion if we stare at a set of parallel lines which are slightly tilted away from vertical and then shift our gaze to a set of vertical lines they appear to be tilted in the opposite direction.

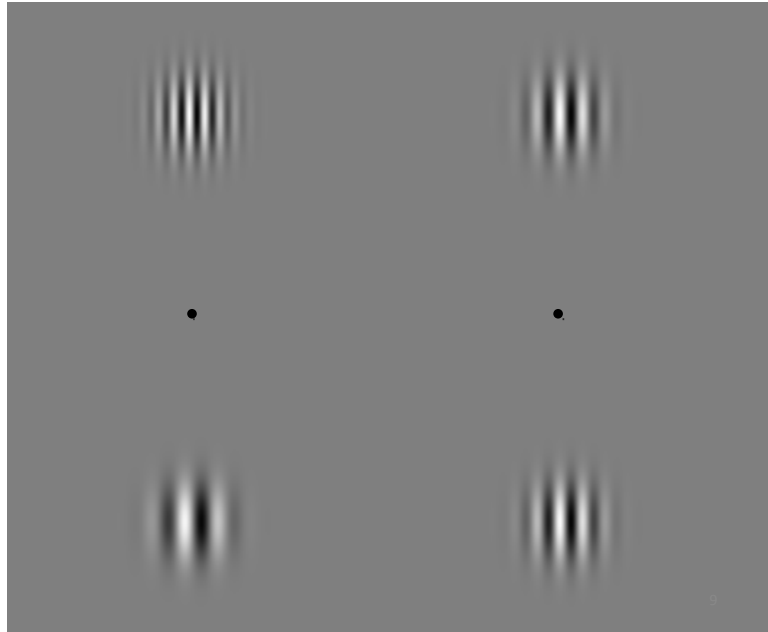


Figure 35 If we gaze at the dot between the 2 different spatial frequencies on the left for around 30 seconds and then transfer our gaze to the equal intermediate spatial frequencies on the right, they appear to be different from each other. The top Gabor appears to be a lower spatial frequency than the lower one. Put another way, the perceived spatial frequency of these stimuli has shifted away from the adaptor spatial frequencies. This is known as the rebound effect (Blakemore et al., 1970).

Neurophysiological studies have revealed neurons which are sensitive to particular stimulus features (spatial frequency, motion or tilt) (Hubel et al., 1962, Hammond, 1978, Tolhurst et al., 1981). These features are perceived by comparing the responses of neurons sensitive to different variations within the features (different spatial frequencies, direction of motion or angles of tilt). After adapting to a particular direction of motion (say), the response of the neurons sensitive to that direction is reduced relative to the neurons sensitive

to other directions. Hence when we shift our gaze away from the waterfall the bank appears to drift upwards.

Adaptation has also been used to investigate time perception. It has been shown that exposure to an auditory or visual “anchor” stimulus duration presented a number of times within a series of random durations from the same modality, can influence the apparent duration of subsequent stimuli in a categorisation task (Behar et al., 1961). This study describes a series of 5 interrelated experiments, the first of which involved presenting a light of 1,2,3,4 or 5 seconds duration in random order, with no anchor, an anchor of short duration (0.2 seconds), or an anchor of long duration (9 seconds). The anchors were presented interspersed with each series of random durations as every fourth stimulus. The term “anchor” signifies the most frequently occurring member of the series. In addition to this increased frequency, the anchor is also presented at regular intervals, allowing it to become a reference point against which non-anchor stimuli are judged. The observers were asked to judge the durations (including the anchors) using an eleven point scale which ran from very, very, very short to very, very, very long. The results showed that observers judged the series of durations with the short anchor as being longer than the no anchor series, and the series with the long anchor as being shorter. The effect was four times as pronounced for the longer anchor than for the short one. The second experiment went on to look at the relative effect of anchors in the modalities of audition and vision and found that they are similar. Experiment four used series of random durations interspersed with anchors of the alternate modality to see whether a series of durations from one modality is affected by an anchor from the

other. The results showed that the crossmodal anchors also produced shifts in perceived durations.

The final experiment looked at the effect on an audio-visual series, of anchors of a single modality. Comparison with data from their 3rd experiment showed that the cross modal anchor effects were smaller than those from the same modality. To sum up, the main findings of this study were, 1) Cross modal anchor effects are similar for visual and auditory durations and are easily produced, 2) Cross modal anchor effects may be produced but differ in scale from intra-modal effects. The study concluded that the cross modal effects found point to a central temporal judgment process.

An adaptation effect has also been found for sub-second visual and auditory intervals using a reproduction task in a series of experiments involving adaptation to visual or auditory stimuli (Walker et al., 1981a). The study demonstrates what the authors call "simple" and "contingent" after-effects. A simple after-effect experiment consists of a single visual or auditory stimulus (either of long duration 0.80-1.0 seconds or short duration 0.20-0.40 seconds) which is repeated over a 60 second period before the subject is asked to reproduce the duration of an intermediate stimulus (0.60 seconds).

The results showed that adaptation to the longer stimuli made the intermediate one seem shorter than in the pre inspection phase and adaptation to the shorter stimulus made the intermediate stimulus seem longer. These effects were similar for visual and auditory stimuli, although the visual stimuli appear to be perceived as longer across durations during the test phase. In contrast to Behar and Bevan (1961), no cross modal adaptation effect was found. The authors suggest that their aftereffects may

be accounted for by a neural adaptation model which supposes a range of duration detectors in the auditory and visual systems each of which has maximum sensitivity to a particular duration and is less sensitive to durations as they become further from its peak sensitivity. Perceived duration would be determined by the response across a range of these detectors. If these detectors were adapted to say, a comparatively long auditory stimulus this would displace the perceived duration of an intermediate auditory stimulus so that is perceived as shorter. This idea is not as unlikely as it first appears. Dedicated channels have been found for the detection of several stimulus features. For example, the auditory system appears to have dedicated channels for the estimation of location (Kashino et al., 1998) and pitch (Fletcher, 1940, Regan et al., 1979). In the visual domain, clusters of neurons in visual area V1 respond vigorously when presented with horizontally oriented stimuli (i.e. their output is tuned) (Hubel et al., 1968b). A critical component of these channel based systems is the presence of individual neural units that respond selectively to a relatively narrow range of afferent sensory information.

The presence of spatial channels (neurons which respond preferentially to particular spatial properties such as spatial frequency or orientation) is a well-established phenomenon (Blakemore et al., 1969(a)). The possibility that something similar may occur in the timing of sub-second intervals with particular neurons responding to durations at or around a particular interval has also been mooted, as discussed earlier in section 3.2.2.

Specifically, a putative channel-based system for duration might contain neural units that respond selectively to a narrow range of stimulus durations

centred on their preferred duration (Ivry, 1996, Becker et al., 2007). By comparing relative activation states across banks of these duration-tuned neurons, a 'population response' would emerge, which would signal the most likely perceived duration. Although behavioural evidence for human temporal judgements sub-served by duration channels remain sparse, it is noteworthy that several neurophysiological studies provide examples of visual (Duysens et al., 1996, Yumoto et al., 2011) and auditory (Casseday et al., 1994b, Faure et al., 2003) neurons displaying band-pass duration tuning. Such an arrangement would confer several advantages to the nervous system. First, population-based estimates tend to be relatively free of the potential ambiguity associated with absolute activity levels within individual channels (e.g. events with similar durations but differing levels of salience/intensity). Second, a system capable of extracting features from a population response is able to interpolate across individual channels, thus facilitating accurate estimates of duration over a range far greater than predicted by its total number of constituent channels. However, while this framework appears theoretically feasible (Ivry, 1996), it awaits experimental validation.

The two major studies investigating duration adaptation fail to agree on whether the effects are sensory specific (Walker et al., 1981b) or transfer to non-adapted modalities (Behar et al., 1961). One possible explanation for this discrepancy could be that the Behar and Bevan paper investigated durations of several seconds, whereas the Walker et al experiments involved sub-second durations.

4.4.1 Spatially specific adaptation

Adaptation to a high temporal frequency visual stimulus has been found to produce a distorting effect on perceived duration which is specific to the area of visual field occupied by the adapting stimulus (Johnston et al., 2006). Specifically, adaptation to a drifting grating which is spatially localised and has alternating direction of motion has an effect on subsequently presented stimuli in that area of the visual field and nowhere else: when the test stimulus has a temporal frequency of 10Hz, adaptation to a 20Hz grating produces a temporal compression. This effect was found using both forced choice and reproduction paradigms (see Figure 36).

These visual adaptation stimuli do not influence the perceived duration of tones. The results of this study suggest the involvement of spatially localised temporal mechanisms in the perception of brief visual stimuli. They also have implications with regard to the locus of timing mechanisms involved in sub-second time perception: This is because if a moving grating were to affect the pacemaker of a central timing mechanism then all spatial positions would be affected and the position of the adaptor stimulus would be irrelevant. This would seem to undermine the model of a central, dedicated timer for humans, at least for durations under 1 second.

Although perceived temporal frequency was affected by both high (reduced) and low (Increased) frequency adaptation, the duration distortion effect was found to be specific to high temporal frequency stimuli only. In other words the duration compression found was independent of perceived temporal

frequency. Adaptation to a stimulus of 5 Hz produced little effect on perceived duration.

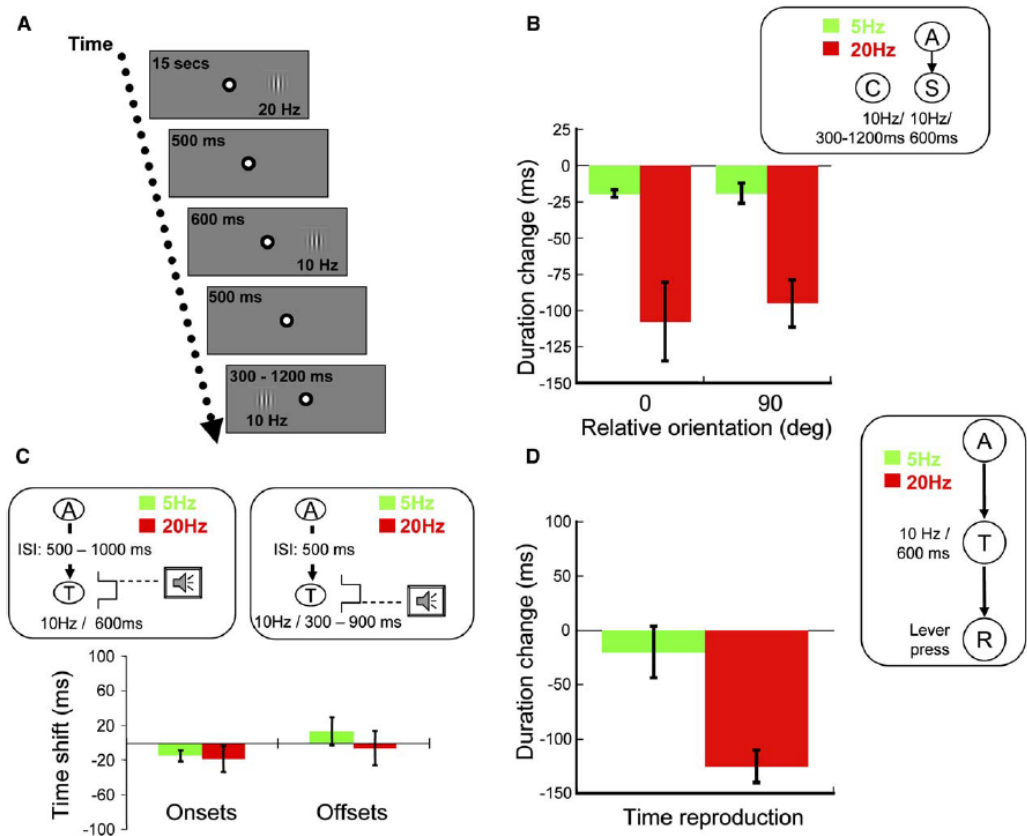


Figure 36 Spatially specific duration adaptation. A) Subjects adapted to a oscillating grating. Subsequently they were required to judge the relative durations of two moving gratings, one of which was in the same position as the adaptor stimulus and one of which was in a different position in visual space. B) Denotes the perceived duration of a 600ms grating after adaptation to gratings drifting at 5Hz and 20Hz. Test gratings could be at the same orientation as the adapting stimulus or at 90 degrees. C) Perceived onsets and offsets of post adaptation gratings relative to a burst of white noise. Perceived duration of post adaptation grating measured by reproduction. A= adaptation, S= standard, C= comparison, T= test, R= response (Johnston et al., 2006).

The authors suggest that post adaptation duration compression may arise via a shift in temporal impulse responses in the magnocellular pathway (Johnston et al., 2006).

A related finding has also been reported by Burr et al. The authors state that these spatially specific adaptation effects occur in real world coordinates. In other words they occur in a particular area of external space which is independent of head or eye position and not in a particular area of the retina (retinotopically) (Burr et al., 2007b). The authors postulate that the locus for this compression of perceived duration lies at a very late stage of the visual system. This hypothesis is based on the spatiotopic nature of the adaptation effect found (external space is represented retinotopically throughout nearly all of visual system) and the results of another experiment which found that adaptation to a drifting grating in one eye produced a spatiotopic effect in the other.

It should be pointed out however, that a later study found no such effect (Bruno et al., 2010). A spatiotopic adaptation effect would imply that the locus of the timing mechanism involved is at a higher level than that mooted by Johnston et al (2006) since external space is not thought to be represented spatiotopically until some areas of the parietal cortex (Galletti et al., 1995, Duhamel et al., 1997). Burr et al (2007) suggest the lateral intraparietal area as a possible candidate. This area has receptive fields which are affected by saccades, after which a remapping of the external world is necessary in order

to maintain continuity of visual perception (Duhamel et al., 1992). Saccades have also been shown to distort duration perception (Burr et al., 2007b).

However, spatiotopic duration adaption to a drifting grating is not a universal finding. A number of experiments have claimed that this phenomenon is in fact, retinotopic (Bruno et al., 2010).

In addition, spatial adaptation has been found with adaptors as narrow as 1 degree of visual angle (Ayhan et al., 2009b). The fact that adaptation has been found for such a narrow visual stimulus points to an early part of the visual pathway being responsible for this effect since the receptive fields of cells in the visual cortex beyond V1 are too large to show sensitivity to changes in the position of such small adapting stimuli. Wherever this effect is located, the idea of a central pacemaker would seem to be incompatible with these findings as it would be expected that adapting to a stimulus in a particular area of the visual field should produce an effect equally across retinotopic or spatiotopic space.

The magnocellular-LGN pathway has been suggested as a likely locus for perceived duration compression following adaptation to a high temporal frequency stimulus (Johnston et al., 2006, Ayhan et al., 2009b, Bruno et al., 2010). This hypothesis is supported by the observation that magnocellular neurons are particularly responsive to high temporal frequency stimuli since significant duration compression was not found when adapting stimuli were lower in temporal frequency (5Hz).

The model which may be used to explain the perceived duration compression outlined above is known as a “content dependent clock model”. This model

suggests that time is measured by predicting how a visual stimulus will appear after a fixed period of time so that when the stimulus matches the prediction the aforementioned period has lapsed and the clock is reset. This model links timing to the visual perception of stimuli with changing characteristics. Perceived duration compression is said to be produced due to differential adaptation effects in magnocellular and parvocellular neurons. To date there does not appear to be a hypothesis which explains how spatiotopic duration compression following adaptation to a drifting grating could occur in the absence of a retinotopic component. To sum up, that a spatially specific compression following adaptation to a high temporal frequency visual stimulus is well established. However, whether this effect is spatiotopic or retinotopic has yet to be proven one way or the other.

Chapter 5

Multisensory integration

5.1 Multisensory neurons

Often, in everyday life, events processed in either the auditory or visual systems will also be detected by the opposing modality. Information from both senses will therefore need to be integrated to produce a coherent representation of the outside world. The interactions between visual and auditory information is managed by multisensory neurons. These neurons respond to events involving more than one sense. Multisensory neurons are thought to be present in a number of cortical areas including the superior temporal sulcus, the temporo-parietal association cortex, ventral and lateral intraparietal areas of the parietal lobe, premotor and prefrontal cortex and insular cortex (Calvert et al., 2004). Multisensory neurons are also thought to be present in subcortical regions (see Figure 37). These include the superior colliculus, inferior colliculus, claustrum, supragenicolate nuclei, medial pulvinar nuclei, and amygdale (Calvert et al., 2004). These areas are also thought to contain some unisensory neurons (Meredith, 2002).

A multisensory response can be said to have occurred when there is a significant difference between the neuronal response to a multisensory

stimulus and the response produced by the most potent unimodal component of the stimulus when it is presented on its own.

Multisensory neurons are especially common in the cat superior colliculus and this area has provided a number of insights into the mechanisms mediating multimodal integration.

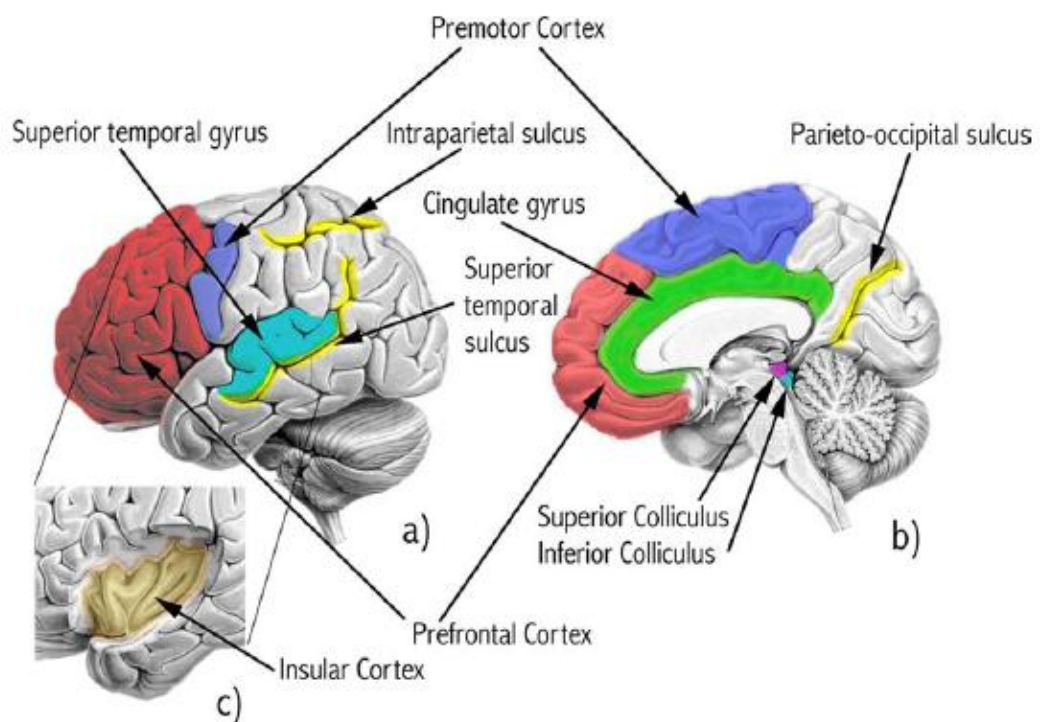


Figure 37 Multisensory brain areas from (a) lateral and (b) mid-sagittal viewpoints. (c) Shows insular cortex after temporal lobe dissection. Different areas of multimodal cortex are shown in distinct colours for both lateral and mid sagittal viewpoints. Yellow defines the boundaries of multisensory regions implicated in cortical sulci. (Calvert et al., 2004).

The superior colliculus is situated in the midbrain, inferior to the thalamus. It has seven layers of alternating grey and white matter. Grey matter consists of neural cell bodies, dendrites, unmyelinated axons and glial cells. It is involved in muscle control and sensory perception. White matter consists

mainly of glial cells and myelinated axons and relays information between different areas of the brain. The visual field is mapped out in the superficial layers, whilst the deeper layers have spatial maps from vision, audition and somatosensory input. The receptive fields from the different modalities converge in the deep layers of the superior colliculus. The multisensory neurons have an excitatory receptive field for each modality, such that the receptive fields of an audio-visual neuron overlap in space. Therefore, the location of an event as opposed to its modality is the dominant factor in determining whether a neuron is activated (Stein et al., 2008). The stimuli from the different modalities will be defined as coming from a common source providing they lie within the spatial region corresponding to the overlapping receptive fields of the multisensory neuron, even if they do not come from the same precise point in space. Within the overlapping receptive fields there is an area known as the “best point” which produces the maximum unisensory response to unisensory stimuli and maximum multisensory enhancement in response to multimodal stimuli (Kadunce et al., 2001). The excitatory receptive fields of some multisensory neurons are surrounded by an inhibitory area. Stimuli which lie within the receptive field of one modality but not the other will not increase the activity of a multisensory neuron; in fact if one of the stimuli is located in an inhibitory area bordering a receptive field, the result may be enough to cancel out the excitation produced by the other stimulus producing an inhibition of activity. A multimodal stimulus will produce a response across a great many receptive fields both excitatory and inhibitory. When stimuli from different modalities coincide in space there will be a maximum enhancement of the multisensory

response. With increasing disparity the number of excitatory receptive fields stimulated is decreased and the number of inhibitory receptive fields increased. It is thought that the balance between the two makes possible the discrimination of small amounts of spatial disparity in spite of the comparatively large area of receptive fields in the superior colliculus (Kadunce et al., 2001). The receptive areas for the different modalities are linked to eye movement, so that auditory and somatosensory fields shift with eye movement. In other words, multisensory neurons in the superior colliculus perceive the world in retinotopic terms (Stein et al., 2008).

In order for multisensory stimuli to be integrated they must also be located in similar time frames, but these can be relatively long, sometimes as long as hundreds of milliseconds (Wallace et al., 1997). It has been suggested that this may be made possible by discharge chains produced by sensory stimuli which have been found to be of relatively long duration.

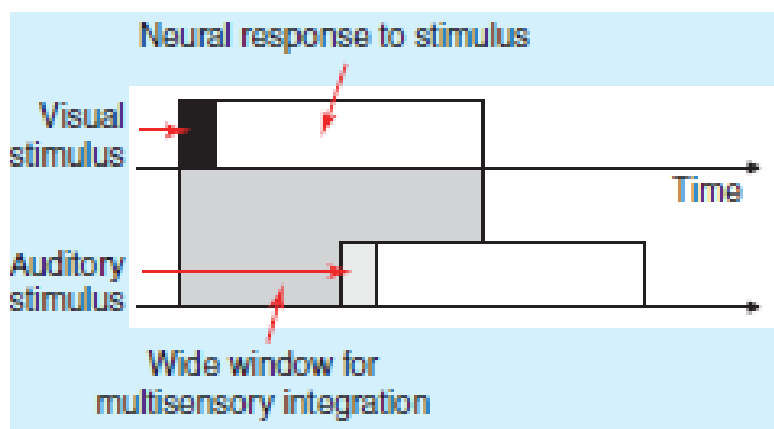


Figure 38 Visual and auditory stimuli occurring at slightly different moments in time, may still be perceived as simultaneous providing the neural response to the first stimulus (in this case visual) overlaps with the neural response to the second stimulus (auditory) (Spence et al., 2003) .

Multisensory enhancement, it is proposed, occurs when these discharge trains overlap, (see Figure 38) although this theory is not universally accepted (Spence et al., 2003) This allows integration of stimuli from the different modalities even though visual, auditory and somatosensory modalities have differing conduction speeds and response latencies. However, an integrated response will be greatest when the peak periods of response from the individual modalities coincide (Stein et al., 2008).

There are three categories of multimodal response. These are known as additive, superadditive and subadditive. An additive response is equal to the sum of the unimodal responses to a stimulus, whereas a subadditive response is less than this total (Stein et al., 2008) and a superadditive response will be greater (Meredith et al., 1986). Often superadditive responses are greatest when unimodal responses are very low. This is known as inverse effectiveness (Meredith et al., 1986). Inverse effectiveness allows an increase in the detection rate of weak stimuli as well as the speed of their detection - an obvious ecological advantage (see Figure 39). Multisensory integration in the superior colliculus is dependent upon input from an area of the association cortex known as the anterior ectosylvian sulcus. The anterior ectosylvian sulcus consists of three unimodal regions which are divided by multimodal neurons where they meet (Jiang et al., 2003).

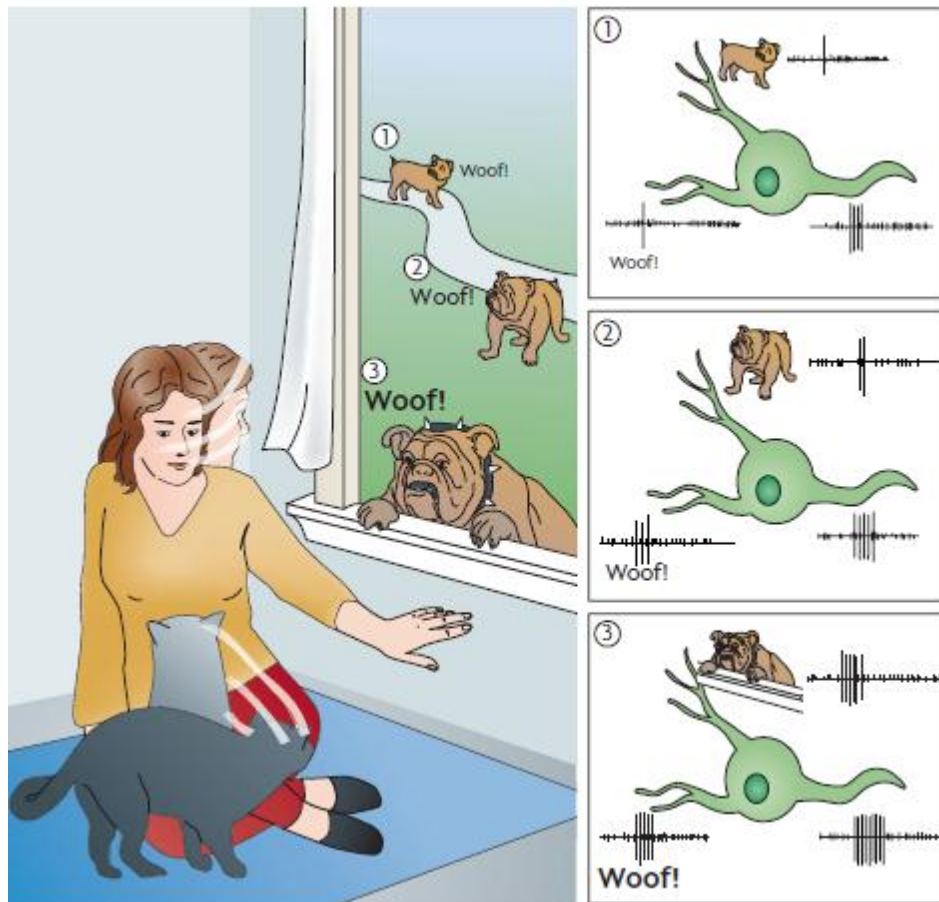


Figure 39 The ecological advantage of multisensory integration. A dog approaches a woman and a cat. When the dog is relatively far away unimodal cues are weak, but the multimodal response is superadditive. The multimodal response exceeds the combined unimodal responses. This superadditive response allows the dogs' presence to be detected at an early stage, thus avoidance is still possible (top). As the dog moves closer, unimodal responses become greater, and the multimodal response is proportionally weaker. At this stage the multimodal response is additive (middle). Finally as the unimodal responses increase further, the response multimodal response becomes subadditive (bottom) (Stein et al., 2008).

Input from these unimodal regions is received in the deep layers of the superior colliculus and facilitates the multimodal integration effect. If input from the anterior ectosylvian sulcus is disrupted, the responses of multisensory neurons in the superior colliculus are no more effective for

multisensory stimuli than for unimodal stimuli and the ecological advantages of multisensory integration are lost (Jiang et al., 2003).

5.2 Audio-visual multisensory integration in humans

Areas of the brain involved in human multisensory integration may be deduced using a variety of brain imaging techniques. Functional magnetic resonance imaging (fMRI) is a very useful tool in this regard owing to its high spatial resolution. The superior temporal sulcus (STS) has been shown to produce high levels of activation relative to other areas of the cortex during audiovisual speech processing (see Figure 40). The effect is at its greatest when audio visual information is congruent (Macaluso et al., 2004, Wright et al., 2003, Marchant et al., 2012, Noesselt et al., 2007). This increase in activation has been found to correspond with an improvement in performance of target detection and object categorization tasks (Werner et al., 2010, Marchant et al., 2012). Areas traditionally thought of as sensory specific may be activated by some surprising stimuli. For example, the auditory cortex produces a response when subjects view images of lip movement with no accompanying sound. When 'none-speech' based visual stimuli were used (random facial gurning) the response effect was absent (Calvert et al., 1997). More recent studies also suggest that multisensory interactions take place in the auditory cortex (Foxe et al., 2005, Ghazanfar et al., 2006). These findings demonstrate that the model of a human brain with sensory specific areas is

probably an over simplification and the true picture is likely to be more complex.

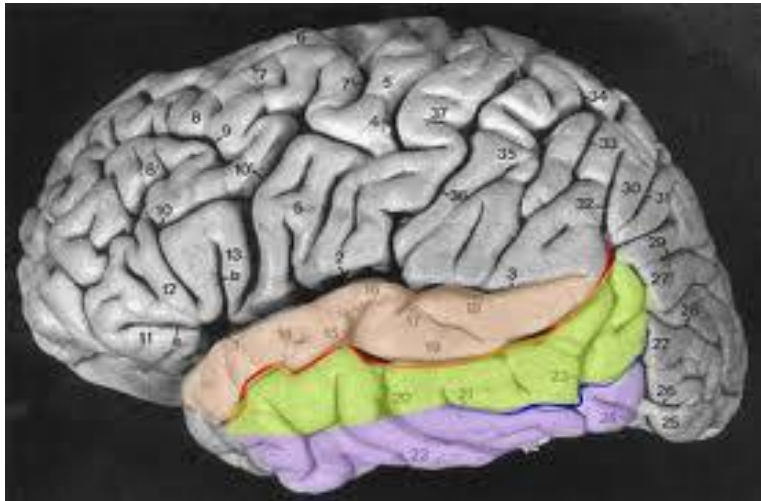


Figure 40 the superior temporal sulcus (red), lies between the superior temporal gyrus (peach) and the middle temporal gyrus (green).

http://culhamlab.ssc.uwo.ca/fmri4newbies/Images/temporal_lobe_sulci.jpg

Animal (primate) studies have shown that direct connections between auditory and visual cortical areas exist (Falchier et al., 2002, Rockland et al., 2003). The implication of this is that audio-visual integration is possible at early stages of neural processing as well as at higher level association areas. It has been suggested that convergence of multisensory information occurs in low level sensory specific areas (Driver et al., 2008) and that in 'real world'

situations the senses may never actually operate independently of one another (Ghazanfar et al., 2006).

A study comparing activation of the visual and auditory cortices during a variety of speech conditions (heard speech, seen speech and congruent audio-visual speech) found that cortical activation during the audio-visual conditions was greater than the sum of that produced by audition and vision individually (Calvert et al., 1999). These response enhancements are absent when bimodal stimuli are unrelated or 'none-speech' based. This points to these multiplicative responses being confined to the processing of speech based information which is congruent in nature (Calvert et al., 1999).

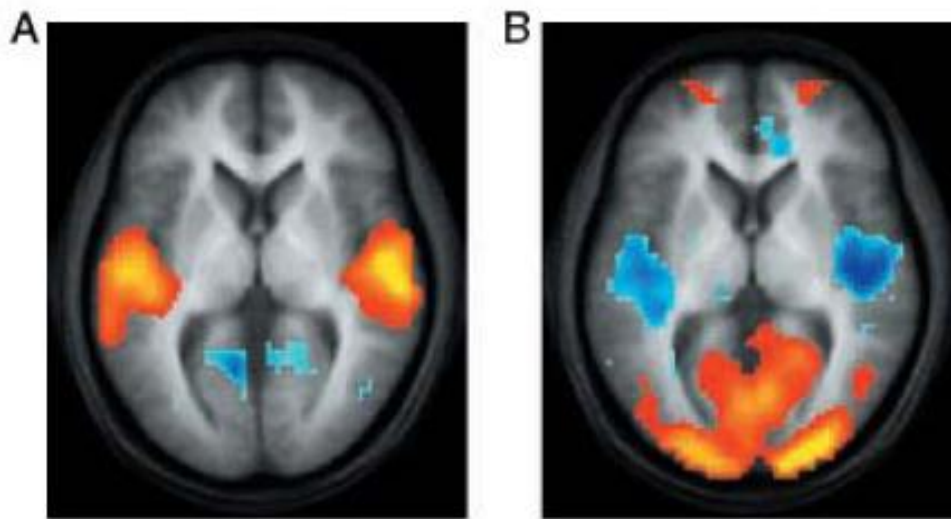


Figure 41 Activation (yellow/orange) and deactivation (blue) of cortical areas following: A) visual stimulation and B) auditory stimulation. Visual stimulation results in visual cortex activation and auditory cortex deactivation. Auditory stimulation results in auditory cortex activation and a deactivation of the visual cortex (Laurienti et al., 2002).

An audio-visual fMRI study using low level stimuli found that activation across the visual and auditory cortices may be reduced as well as increased. When presented with visual stimuli, auditory cortex activation decreased relative to background activity. When auditory stimuli were presented, visual cortex activation decreased to below background levels (Figure 41). Audiovisual stimuli however produced a multiplicative response in both cortices (Laurienti et al., 2002). Support for reciprocal and competitive interaction between auditory and visual brain areas comes from a study which measured responses to identical auditory and visual stimuli some of which were perceived as being bound together whilst others were perceived as separate events. The conditions in which audio-visual stimuli were bound together resulted in a higher response in multimodal areas together with a reduced response in unimodal areas (Bushara et al., 2003).

Although fMRI has the advantage of its high spatial resolution, evoked potential (EP) studies are most useful if we wish to consider the temporal dynamics of multisensory integration. It is thought that EPs produced in less than 200ms from stimulus onset are likely to originate in the visual pathway (Regan, 1989).

However, audio-visual stimuli have been shown to produce modulation of occipital evoked potentials just 40ms from stimulus onset (Giard et al., 1999). The authors conclude that multisensory integration can take place at early stages of sensory processing and operates in both sensory specific and non-specific areas. Further evidence of low level multisensory integration comes from a study produced by Shams et al (2001). The authors presented subjects with a flash of light in their peripheral visual field together with two

concurrent auditory beeps. This resulted in the flash being perceived as two flashes in 81% of trials. They measured subject EPs during the experiment. The results showed that the EPs produced by the illusory flashes were the same as those expected from a genuine double flash of light (Shams et al., 2001). The way in which the activities of early and late stage multimodal processing interact remains unclear. In some EP studies, EP activity is found to increase in multimodal areas before unimodal areas (McDonald et al., 2003, Teder-Salejarvi et al., 2002). Some authors suggest that there exists some sort of top down feedback loop between unimodal and multimodal area (Bushara et al., 2003, Macaluso et al., 2000).

To conclude, it is becoming apparent that sensory information from different modalities interacts at multiple levels. Multisensory integration in the superior colliculus depends on input from the neocortex. In the primate amygdala neurons may be found which respond to visual, auditory and somatosensory stimuli as well as neurons which respond to information from multiple senses. These neurons produce maximal responses to novel, oddball type stimuli, suggesting that some multisensory areas may be adapted to cope with particular external factors. The thalamus has also been shown to be involved in multisensory integration. Multisensory cortical areas include the temporal superior temporal sulcus, the intraparietal areas of the parietal cortex and the frontal cortex (Ghazanfar et al., 2006). Multisensory interaction has been demonstrated at early stages of sensory processing, in particular within the auditory cortex. These areas appear to be connected in a complex network involving cortical and thalamo-cortical connections, the connections within thalamus itself as well as cortical sensory and associative areas. These

findings are leading some researchers to conclude that the neocortex is essentially multisensory in nature (Ghazanfar et al., 2006) and that multisensory processing occurs over a widely distributed system which is refined to deal with a range of specific circumstances (Cappe et al., 2009).

5.3 Cue combination

The process of multisensory integration is often considered as a form of “cue combination” (the combination of multiple sources of behaviourally relevant information). Cue combination produces a fuller perception of an event than may be produced by one modality alone (Ernst et al., 2004, Burr et al., 2006). It is important, for obvious reasons, that events which are combined in this way do actually belong together. Since signals from different events in different sensory modalities may arrive at slightly different times and are processed at different speeds, a certain amount of slippage is necessary in order to allow the discordant spatial or temporal signals from the different modalities to be integrated.

The inclination for humans to bind stimuli together from different modalities when they are perceived as having a common source is known as the “unity assumption” (Vatakis et al., 2007). In this way, information is deemed to have arisen from a single source. However, when the difference in time and space between them is too great they are treated as separate events and multisensory binding does not take place (Klink et al., 2011, Vatakis et al., 2006, Zampini et al., 2003, Spence et al., 2003). One example of the

integration of misaligned sensory information is the Ventriloquist Effect in which vision is said to “capture” perceived auditory position. The illusion works because the perceived location of an auditory stimulus is relatively ambiguous compared to that of a visual stimulus. Other examples include the “McGurk” effect and the “parchment skin illusion”. The McGurk effect relates to speech perception in which subjects listen to a repeated word whilst watching the speakers’ lips (McGurk et al., 1976). If the visual and auditory cues are slightly mismatched (for example the lips make the shape for “goes” and the word spoken is “bows”), subjects may perceive the spoken word as “those” or “doze”. The parchment skin illusion occurs when subjects rub their fingers together accompanied by a grating sound played through headphones. The result is a change in the perceived texture of participants’ finger tips (Jousmaki et al., 1998).

There are several explanations as to how cue combination occurs. One school of thought assumes that information from the most appropriate modality for the task dominates perception, to the exclusion of the less accurate senses (Welch et al., 1980b). For spatial perception this is vision because it is the modality with the highest spatial acuity (Welch et al., 1986). Another approach suggests that no one modality is innately superior; rather the senses are given different perceptual weighting depending on their relative reliability in a given situation, with the most reliable sense being allocated the greatest weight. This approach is known as Maximum Likelihood Estimation (MLE) or optimal combination (Alais et al., 2004, Ernst et al., 2002). The MLE model takes sensory estimates from each modality and assigns perceptual weighting according to their relative reliabilities. It

then combines them to produce the most accurate estimate possible from the available information (the optimal estimate). This model predicts that if information from the more reliable or least variant modality (in this case vision) is degraded in some way, its weighting in the combined perception of a location will be reduced and thus the weighting of the other modalities will increase in a proportional manner. In other words, if simultaneous visual and auditory signals have differing spatial locations, the perceived audio-visual location will be pulled towards the position of the auditory stimulus if the reliability of the visual stimulus is reduced to a sufficient extent. By giving the appropriate weighting to each independent stimulus estimate from the different modalities, and combining them, the MLE model also predicts that bimodal discrimination will be better than that afforded by the most reliable modality alone (Ernst et al., 2002, Jacobs, 1999).

An alternative model for spatial integration proposes that a combination of the MLE and “winner take all” models is more appropriate. This model proposes that the relative reliability of the different modalities is taken into account, but subjects nevertheless have a bias in favour of the visual modality because of previous experience in which vision has been, in general, the most reliable modality. The consequence of this is that weightings are skewed towards vision to a greater degree than would be predicted by relative reliability in any particular situation alone (Battaglia et al., 2003b). Until recently it has been thought that a reduction in attention to a sensory modality would lead to a reduction in its reliability and therefore in its weighting (Prinzmetal et al., 1998). Research has been published which finds that in some situations selective (i.e. top-down) attention influences

multisensory integration for particular stimulus combinations. Specifically, if attention is diverted away from, say visual input, its contribution to the multisensory percept is reduced (Talsma et al., 2005, van Ee et al., 2009). However, a study investigating visual-haptic cue weighting in size discrimination has found that the withdrawal of attention from the visual modality had no effect on the relative weighting of the senses (Helbig et al., 2008a), suggesting that perceptual weighting occurs automatically and at a stage in sensory processing which is prior to the implementation of attention mechanisms (Helbig et al., 2008a, Bertelson et al., 2000, Vroomen et al., 2001). It has been proposed that both top down and bottom up processes influence multisensory integration, depending upon the complexity of the stimuli. Specifically when multisensory stimulus complexity is low, pre-attentive, bottom-up processes are employed in the cue combination process. When there are a large number of stimuli within each modality which require processing, top down attention is necessary in order that the appropriate stimulus events from the different modalities are integrated (Talsma et al., 2010).

For temporal discrimination tasks, it has been demonstrated that audition is more reliable than vision (Walker et al., 1981b, Burr et al., 2009, Wearden et al., 1998). Therefore, audition would be expected to “capture” vision in temporal judgements in the same way that vision captures sound in spatial judgements. Fendrich and Corballis (2001) asked subjects to judge the time at which a visual flash occurred. When the flash was preceded by a sound it was judged as occurring earlier in time. In a similar way when the flash was followed by a sound its perceived timing was shifted later in time towards.

When the modalities were reversed, vision had an influence on the perceived timing of audition but to a lesser degree (Fendrich et al., 2001). Another temporal example of the influence of audition on vision is known as the “double flash” illusion. If a single flash is accompanied by two clicks it is perceived as two flashes (Shams et al., 2000).

Since the MLE model of modality integration predicts many of the findings produced by studies of modality integration in the spatial domain, it seems reasonable to ask whether this model might be applied with equal success to modality integration for timing. Results from temporal experiments so far are not entirely as predicted by optimal cue combination models. In a series of experiments involving an audio-visual temporal bisection task, Burr et al (2009) found that more weight was given to audition than would be predicted by the relative reliabilities of the auditory and visual modalities. Furthermore, bimodal thresholds were found to be no better than those of audition alone (Burr et al., 2009).

Another experiment, involving asynchronous timing of audio visual stimuli found that the perceptual timing of the less reliable stimulus (vision) was pulled to towards the more reliable stimulus (audition). This is in line with MLE predictions, but as with the Burr study bimodal thresholds were no better than those of audition alone (Hartcher-O'Brien, 2011).

A study investigating visual flicker and auditory flutter rate perception found that their data was best modelled in a Bayesian fashion in which it is supposed that subjects combine weightings based on the relative reliability of the visual and auditory information with prior knowledge regarding the

probable relationship between different audio/visual rates to produce an estimate of the most likely bimodal rate (Roach et al., 2006a).

Studies looking at multimodal integration of temporal extent as opposed to those concerned with temporal localisation are relatively sparse. However, a study which looked at duration discrimination judgements concluded that when auditory and visual durations are perceptually bound together, visual durations are influenced by auditory stimuli but perceived auditory durations are unaffected by visual stimuli (Klink et al., 2011). The authors explain their findings in terms of the pacemaker/accumulator model and conclude that we have separate pacemakers for vision and audition, but when an event is perceived as bimodal the audio pacemaker is employed by virtue of its superior sensitivity.

To summarize, models for the way multisensory temporal information is combined may be placed into three main categories.

- The Auditory dominance pacemaker accumulator theory in which we have separate “clocks” for the different senses and the most appropriate/sensitive clock (audition) is used to the exclusion of the others.
- The maximum likelihood model which has been reasonably successful in predicting data produced from spatial integration tasks though less so for temporal integration, in which information from the different modalities is weighted according to relative reliability and then combined to give a multimodal timing estimate.

- The Bayesian model in which prior experience is combined with weighted multimodal information to produce an optimal estimate of temporal extent.

Only one of these models originates from experiments involving filled durations – that of Klink et al who propose a pacemaker/accumulator explanation for their findings (Klink et al., 2011). Hartcher-O'Brien et al (2011) do argue that their temporal localisation task could be construed as a type of duration discrimination task. However although they find temporal ventriloquism in line with the MLE model, they report that discrimination was no better in bimodal than unimodal conditions, so this may hardly be considered as total support for the MLE model. As for the Bayesian model, this has yet to be tested using a filled duration discrimination task. We investigate multisensory integration in Chapter 11.

Chapter 6

Neural correlates of duration perception

Studies looking at prospective timing in humans and animals have produced a general consensus that an important area in the timing of intervals in the seconds range, consists of fronto-striatal circuits which consist of recurrent loops between the frontal cortex, the supplementary motor area, the caudate-putamen, pallidum and thalamus. These circuits are driven by the dopamine system (Harrington et al., 2004). Motor functions are controlled via cells in the substantia nigra which produce dopamine and release it into the striatum (a subcortical part of the forebrain made up of the caudate and putamen) which lies close to the thalamus. The striatum sends input to the basal ganglia system and receives input from the cortex (Figure 44).

Haloperidol, which is a dopaminergic antagonist, has been shown to impair duration discrimination in healthy subjects (Rammsayer, 1999). Other studies have found that dopaminergic agonists and antagonists increase and decrease perceived time respectively (Buhusi et al., 2002, Cevik, 2003). It is suggested that this may be due to a speeding or slowing down of the internal clock (Wittmann, 2009). Reduced dopamine in the basal ganglia of patients with Parkinsons is linked with a decrease in performance in duration discrimination and reproduction tasks (Hellstrom et al., 1997). Taken together these studies appear to point to the importance of dopamine neurotransmission within frontal and striatal brain areas in the accurate judgment of duration.

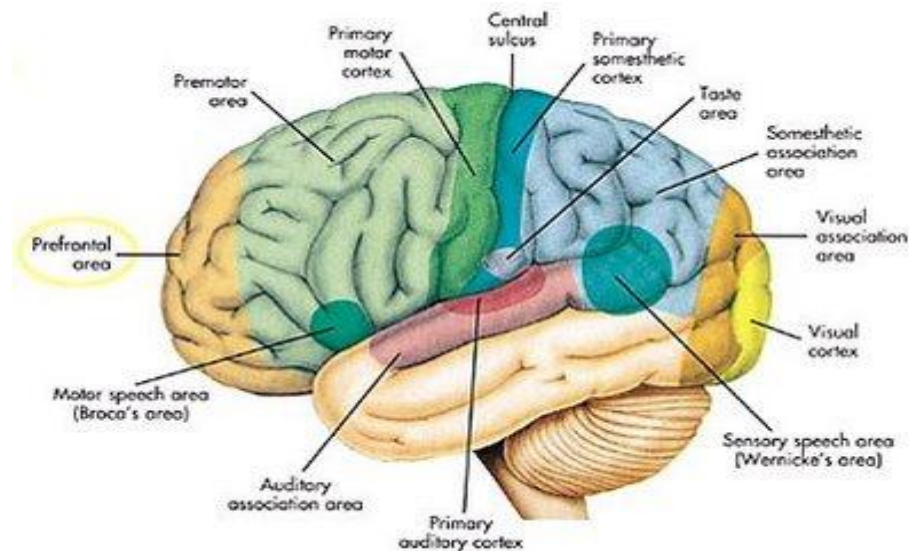


Figure 42 Diagram of the brain showing the prefrontal cortex which is situated at approximately the left third of the frontal lobe (far left).

http://4.bp.blogspot.com/_6Zb1XQameA/say/8pil7pn...prefrontal.jpg.

The right prefrontal (Figure 42) and striatal areas (Figure 44) have been shown to be active during duration judgment in a number of neuroimaging experiments (Coull et al., 2004, Ferrandez et al., 2003)

Patients with traumatic brain injury to the frontal lobe have been shown to have attention and memory deficits which cause them to have a greater variability in performance of production and reproduction tasks. However, their mean productions/reproductions do not deviate from those of normals (Pouthas et al., 2004). The authors suggests that although this area of the brain is important in timing, its significance lies not in timing as such, but rather in attention and working memory and the effect they have on temporal processing.

The suggestion that different processes are involved for the timing of sub second and supra second intervals has support from experiments involving

the use of pharmacological drugs in timing tasks. Agents which affect working memory (such as midazolam), have been shown to impair duration discrimination of intervals of around 1 second, but not those of a few hundred milliseconds (Rammsayer, 1999). This has led to the hypothesis that very brief intervals of a few hundred milliseconds are judged without the use of working memory or influence of conscious attention, whereas in the judgment of longer intervals cognitive strategies such as counting come into play and work alongside timing processes common to sub-second and supra-second timing which are controlled by dopamine (Rammsayer et al., 2001). A relatively recent study shows that the right posterior-inferior parietal cortex (see figure 43) may be involved in integrating timing information from different senses and in the perception and timing of movement (Buetti et al., 2008c). This study used fMRI imaging to investigate brain activity when subjects reproduced and estimated durations. The study found the right posterior-inferior parietal cortex was involved during the reproduction tasks but not the estimation tasks. Many areas shown to be active in neuroimaging experiments may not necessarily be directly involved in timing since there are a number of different cognitive processes involved in timing tasks such as memory, decision making and attention (Rubia et al., 2004). This may help to explain why so many different pathologies affecting different brain areas have an adverse affect on timing performance. It is also worth noting that as yet, no study has reported the abolition of temporal processing due any disease or lesion (Buonomano et al., 2010).

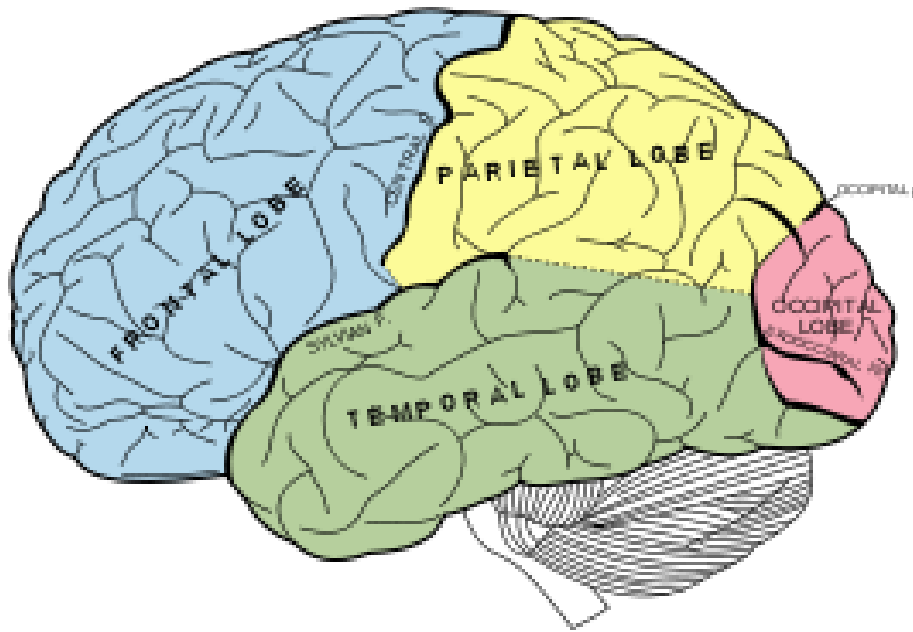


Figure 43 The parietal left cortex is shown here in yellow and for this diagram the right parietal cortex would be on the far (out of sight) side of the brain.

<http://upload.wikimedia.org/wikipedia/commons/thumb/1/1a/gray728.sug.png>

It is this difficulty in finding a particular brain area for timing tasks that has led some researchers to suggest that rather than a central clock-type mechanism being responsible for time perception, distributed neural networks may be responsible (Mauk et al., 2004). In this case, different neural networks would intrinsically encode duration by time dependent neural changes or by virtue of the amount of neural energy expended (Karmarkar et al., 2007). These mechanisms would be limited to automatic sub-second durations which would not use the cognitive processes needed for the estimation of longer intervals (Ivry et al., 2008). This suggestion is supported by a study which analysed neuroimaging data from a number of studies and suggests different mechanisms for second and millisecond timing. It is thought that sub-second timing is served by an automatic system involving the motor and pre-motor

systems, along with the cerebellum (Figure 44) and auditory cortex, whereas for durations of 0.5-1.0 seconds and above, a system involving cognition situated mainly in the right prefrontal and parietal cortex areas (Figures 42 and 43) is used (Lewis et al., 2003).

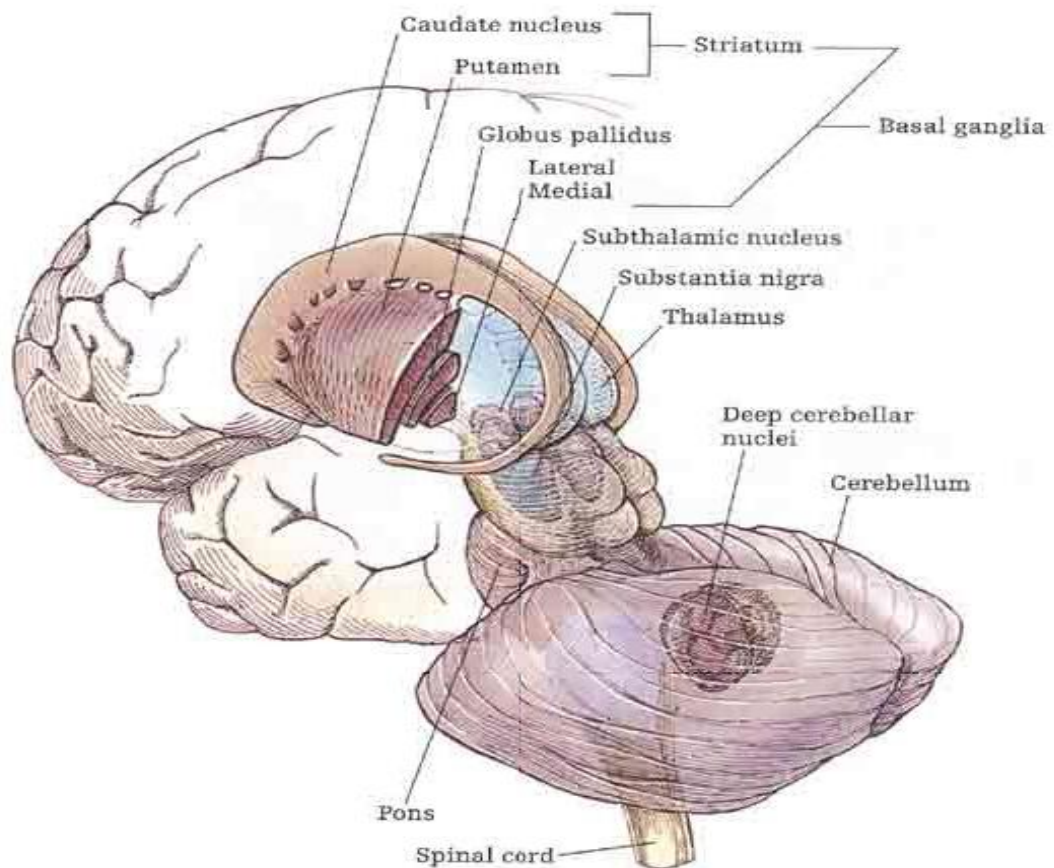


Figure 44 Brain areas thought most likely to be involved in sub second automatic timing. These are the basal ganglia and the cerebellum, - structures which are involved in motor tasks. They receive input from the frontal cortex. www.epistemic-forms.com/limbic-system.

In addition a number of studies have reported that some neurons produce changes in firing rate over time. This phenomenon has been observed in different brain areas including the prefrontal, parietal and motor areas (Roux

et al., 2003, Leon et al., 2003, Brody et al., 2003). These increasing or decreasing firing rates are mostly seen to peak at the time of an anticipated response. However it has been suggested that this activity may not necessarily be involved in timing and that it may be associated with preparation for, or in expectation of, a motor response. Thus this activity could be in response to temporal information from elsewhere in the brain (Buonomano et al., 2010). To sum up, it appears that many different areas of the brain are involved in temporal processing. The involvement of cognitive factors such as attention, memory and decision making all serve to make the “pinning down” of a precise area for duration judgments more difficult, assuming such an area exists. Separating out cognitive influences is no easy task when the act of asking someone to make a duration judgment seems likely to activate the very factors we are trying to eliminate.

6.1 Motor and sensory timing

An unresolved question regarding motor and sensory timing is whether they are achieved by separate temporal processes or by a mechanism which is common to both. For the purposes of clarity, here motor timing is defined as timing involving movement, either external movement of the target or movement of the organism, and sensory timing as more passive duration judgments which do not involve movement. Until relatively recently, studies pointed to a common mechanism (Ivry et al., 1995, Keele et al., 1985). Experiments comparing performance in perceptual and production tasks found that when the performance of a tapping task was compared with that of a discrimination of empty intervals task across various durations, the functions produced (variance plotted against duration) had almost identical slopes, leading the authors to conclude that both types of timing shared a common mechanism (Ivry et al., 1995). An earlier paper came to the same conclusion following a series of experiments which demonstrated that people who were “good” at judging brief perceptual events performed equally well during production tapping tasks with both fingers and feet. In other words, a subjects’ performance in sensory timing tasks correlated with their performance of a motor timing task (Keele et al., 1985). The same study found that on average skilled piano players were better at both tasks (perceptual and motor) than a control group of non-piano players. However, in a study involving a greater range of tasks and conditions, Merchant et al found that even though there was a linear increase in the variability of performance with increased duration across all task types, variability was

greater for motor than perceptual tasks and also greater for visual than auditory tasks (Merchant et al., 2008).

The authors seek to explain these results by proposing the existence of two networks, an automatic mechanism which times predictable sub-second durations defined by movement, and a more cognitive mechanism for longer durations not defined by movement. It is suggested that the two systems overlap. They cite a study which reviewed fMRI literature for different timing tasks and came to a similar conclusion. It was found that sub-second and supra-second tasks produce differing patterns of brain activity. However the authors of this paper suggest that the motor or “automatic” system may also be used to measure brief durations even when no movement is present (Lewis et al., 2003). A proposal that motor and sensory timing may be achieved by non-overlapping systems comes from a review paper by (Buonomano et al., 2010). The paper looks at a particular theory of intrinsic timing known as “population clocks”. A population clock is said to be a group of neurons which respond to a stimulus and also allow the timing of the stimulus by producing a spatial pattern of activity which changes over time. At the termination of the stimulus, its duration is judged based upon the pattern of activity of the neurons at the end point of the stimulus. For motor timing these groups of neurons may be recurrent. In other words, they form a loop so that the activity of a neuron may indirectly feed back into itself. (Buonomano et al., 2010) suggest that sensory events are timed using non-recurrent networks. It is argued that the crucial factor with regard to motor and sensory timing is the strength of the connections between the different neurons. If the strength of the connections between neurons is weak then

activity produced during the timing of an event will cease relatively quickly when the stimulus ends. This, it is proposed, is what happens during “sensory timing”. “Motor” timing on the other hand would involve neurons with strong connections. These groups of neurons would be capable of producing self maintaining activity even when the initial stimulus has ceased. Thus a recurrent network loop could occur. In this way a group of neurons would be able to predict the timing of the next beat in a tapping task for instance or the time taken for a moving object to reach a particular position in space based on previous activity in the neural network (Buonomano et al., 2010). An early population clock theory proposed that interactions between granule (excitatory) cells and Golgi (inhibitory) cells in the cerebellum could encode time and the activity of these cells could be detected by Purkinje cells in order to read out the state of the network and hence time the event they represent (Buonomano et al., 1994, Mauk et al., 1997). However, circuitry in the cerebellum is known to be incapable of sustaining recurrent excitatory activity (Buonomano et al., 2010). This would seem to suggest that although the cerebellum may well be a strong candidate for sensory timing, the population clock theory would rule out its involvement in motor timing. Since cortical networks have the strong excitatory connections necessary for motor timing, the population clock theory would support the idea of motor timing in these areas. The pre and supplementary motor areas have been suggested as likely candidates and are known to be involved in sequence generation (Buonomano et al., 2010). However, since damage to the cerebellum has been shown to produce impaired performance of the timing of movement (Keele et al., 1985, Ivry et al., 2004, Spencer et al., 2005), it is difficult to

believe that it is uninvolved in motor timing if only because of its role in learning these tasks. One possible explanation which seems to circumvent this problem is proposed in a paper by Zelaznik et al (2005). The authors conducted a series of experiments which supported their proposal that motor timing requires an initial event based representation of a duration in which a temporal goal is externally defined (sensory timing), but that after a few “movement cycles” control processes take over, which allows the timing to be emergent (internally driven) and that it is these processes that are observed at work in motor or predictive timing (Zelaznik et al., 2005). If this is the case then it would be possible for disruption of the cerebellum to produce a reduction in performance for motor tasks due to its involvement in the initial stages of the task.

Taken together, it seems likely that sensory and motor timing may have a common source, but when motor timing is required additional process are brought into play. Intrinsic timing models also suggest that in both cases readout neurons may be necessary in order that an ultimate temporal value may be produced.

Chapter 7

Psychophysical methods

Historically, psychophysical methods for the investigation of perceived temporal extent may be divided into two main groups: duration scaling, and duration discrimination (Allan, 1979). These are set out in the table below.

Duration Scaling	Duration Discrimination
Verbal estimation	Method of comparison
Magnitude estimation	Forced choice-fixed standard
Category rating	Forced choice-roving standard
Production/Reproduction/Ratio Setting	Method of single stimulus

7.1 Duration scaling

Verbal estimation involves presenting the subject with a duration and asking them to estimate its length e.g. “about 3 seconds”. Magnitude estimation is similar to verbal estimation but instead of using, say seconds, the observer allocates a number from their own internal scale to represent the length of

the duration. Sometimes a “standard” duration is presented at each trial or at the beginning of a series of trials which the subject judges before the test durations in order to allow the subject to place subject durations in context.

Verbal estimation and Magnitude estimation are both used in a study by (Bobko et al., 1977b) and the results from each contrasted. The authors found that subjects using the methods of estimation had a tendency to slightly underestimate the presented durations. This effect was similar for both methods.

In the case of category rating (a variation on magnitude estimation) the subject is presented with a duration and asked to place it in order with a series of other presented durations (for example, ranging from very, very long to very, very short). An example of this method is used by Behar and Bevan (1961). The main difference between category rating and magnitude estimation is that with category rating the scale of categories is dictated by the experimenter, whereas with magnitude estimation the subject decides the categories used.

“Production” involves the subject being asked to produce a particular period of time, sometimes by means of a button press arrangement or sometimes by performing a task for the required interval. An example of this would be when subjects are asked to tap their finger for a particular period of time, say 2 or 5 seconds. This method is used by Pouthas and Perbal (2004)

With reproduction, the subject is presented with a stimulus before being asked to reproduce it (again, typically by means of a button press). This method is used in a variety of studies, one example among several (Walker

et al., 1981c, Schiffman et al., 1977, Brown et al., 1999, Brown et al., 2002) being a paper by Walker et al (1981), in which subjects are asked to reproduce visual and auditory stimuli by pressing a stimulus button for the required duration. With ratio setting, the subject is asked to reproduce a fixed proportion of the duration presented. For example Warm et al (1975) use this method in which they present subjects with an auditory or visual stimulus and the subject is asked to press a button, which cuts off the stimulus, when they believe that half of the duration has passed. The object of this experiment was to see whether training for accuracy in temporal discrimination is transferable between visual and auditory modalities. The authors concluded that it was and that this suggested a central timing mechanism for temporal discrimination in different modalities (Warm et al., 1975).

7.2 Duration discrimination

The Method of Comparison is a technique in which an observer is asked to judge the relative durations of two intervals of time. This method can be used in different ways. The forced choice fixed standard method or method of constant stimuli involves a standard duration which is presented before or after the test duration. The subject is then asked to specify the relative duration (which was longer/shorter). This method is used by Goldstone and Lhamon (1974), and a variation of this is used in a paper by (Grondin, 2005), in which observers are asked to indicate which duration is longer, the first

(standard) or second (one of a series of other test durations centred around the standard). An important factor here is that subjects are not allowed to say they “don’t know” or “can’t decide” and therefore must make a guess if the stimuli appear to be of equal length (Goldstone et al., 1974, Grondin, 2005). This method is often known as the method of constant stimuli. A roving standard method is the same except that the position of the standard duration (first or second) varies from trial to trial. This technique has also been used by Goldstone and Lhamon (1974). It is more commonly known today as a forced choice method.

The method of single stimuli involves a stimulus being presented and the subject deciding whether its duration is longer or shorter than an established “internal mean” duration built up from repeated exposure to all previous stimuli. This is an efficient method since only a single stimulus is presented on any trial (Brown, 1998).

7.3 Methods used in this thesis

The majority of methods used in the experiments described in this thesis fall into the duration discrimination category and are of the forced choice type in which subjects are asked a particular question and given a number of alternative possible answers from which they have to choose, none of which may be “yes” or “no”. In these experiments the number of choices was two. The majority of questions related to the relative durations of different visual and auditory stimuli (e.g. “which lasted longer – the flash of light or the burst of white noise?”). The crucial point with the forced choice paradigm is the subjects have to make a decision. When they are unable to distinguish between different stimuli they have to guess. A big advantage is that it removes the element of criterion from the experiment. For example, if we ask subjects whether they detected the presence of a visual stimulus, some observers will only respond “yes” if they are “absolutely” sure about what they saw. Others may respond yes if they think they may have seen even a “hint of a flash”. Because of these different criteria, responses may vary widely between observers purely because of the different criteria they use.

Alternatively, asking subjects to tell us whether the light was presented on the left or the right of fixation forces observers to extract as much information as possible, regardless of how sure they are that their decision is the correct one.

Thus, this method can be said to be less subjective than other, more criterion dependent measures such as duration scaling. Using forced choice methods

has been shown to reveal thresholds for the detection of lights and sounds which are lower than those found with the method of adjustment. In other words, subjects have been found to be able to detect these stimuli even when they claim not to be able to do so. It seems that subjects require less stimulus information to confirm a decision when using a forced choice method (Blake et al., 2006).

After using a forced choice method it is necessary to calculate the value of interest (in our example, the point at which a stimulus is detected). This is achieved by plotting the range of responses to our question (e.g. "was the light on the left or the right of the screen?"), from chance performance to 100% correct, against the intensity of stimulus presentations. The resulting data is then fitted with a curve to produce a psychometric function, from which values of interest can be extracted. In the case of a two-alternative, forced choice detection task, subjects could obtain the correct answer 50% of the time purely by guessing. The function has y values ranging between 50% and 100% correct. Conventionally the detection threshold is taken to be the stimulus intensity at which a subject obtains 75% of correct responses, in other words, the point which is half way between chance performance and 100% correct responses. If the task had four alternatives, this point would be 62.5% correct responses, this being the point half way between chance (25%) and 100% correct. However, it is of course, possible for other points to be chosen as representing threshold although the experimenter would need to justify his or her choice.

When a reference stimulus is always presented first in a forced choice paradigm it is known as the method of constant stimuli. This method is used

extensively in the experiments described in the thesis. In this experimental design, one stimulus is designated as the “standard” or “reference” stimulus and the other as the “test” stimulus. In the case of a duration discrimination experiment, the “standard/reference” stimuli would be of a fixed duration, say, 300 ms and the test stimuli would be a range of fixed durations centred around 300ms, each of which is presented for a equal number of times. Subjects are asked “which of the two stimulus durations is the longest?” The “test longer than reference responses” are then plotted against the test duration.

Normally there are five, seven or nine different values of the test stimulus which are equidistant in value and centred on the standard value. In the case of the duration discrimination experiments, subjects were asked to differentiate between the durations of a auditory reference stimulus and a visual test stimulus or vice versa. The test durations are presented in random order. Judgments will be easier when the difference in duration between the standard and test stimulus is significantly greater than the subject’s duration discrimination threshold and more difficult when they are closer to threshold. This introduces some variety into the task and also provides some confidence boosting “easy” judgments for naive subjects so they do not become discouraged by too many “impossible” judgments. Also, this variety helps ease possible fatigue and boredom.

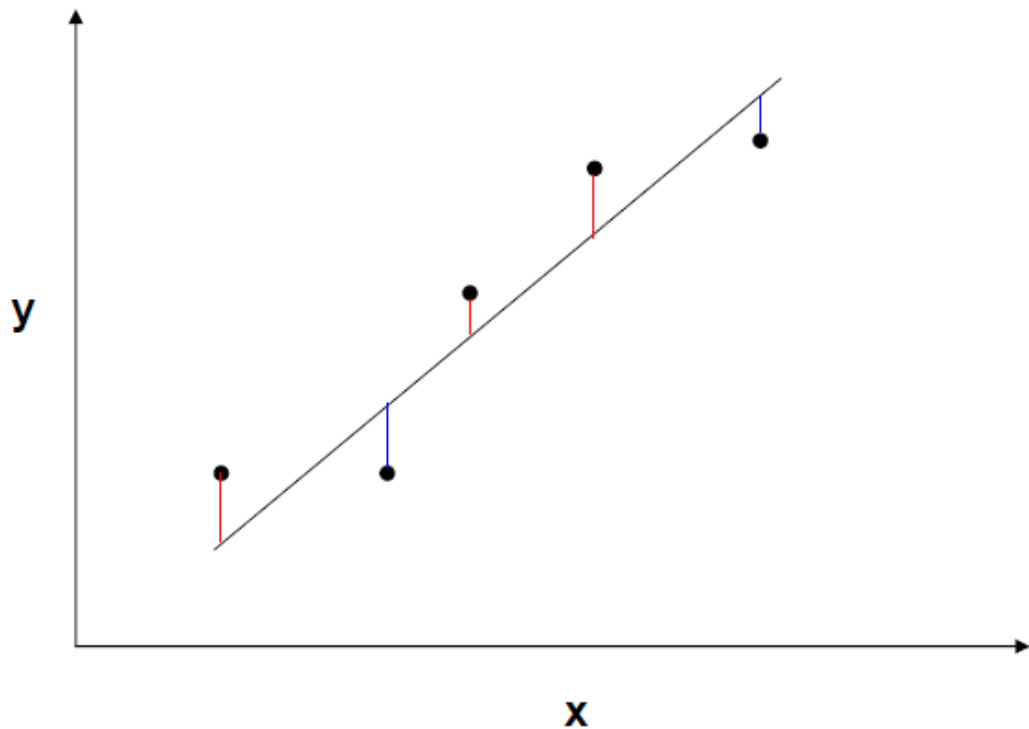


Figure 45 A linear regression through 5 data points. The line gives the minimum vertical offset for each of the data points. The values are squared in order that positive (red) and negative (blue) values should not cancel each other out.

The data points produced may then be fitted with a line. When we fit data points with a line it is known as a regression (Figure 45). The regression must represent all the data points. The method of least squares is used in the fitting of all the psychometric functions shown in this thesis. This involves producing the curve in which the sum of the squared vertical distances between the individual data points and the curve is the minimum possible value.

Software (e.g. Kaleidagraph) calculates a potentially infinite number of curve locations and selects the one which most closely matches the data points i.e.

the one with the lowest total squared vertical distance from the data points. The vertical distances are squared in order to produce positive values for each data point, so that positive and negative values are not offset against each other. This method is considered to be well suited to psychophysical experiments where the number of data points is relatively small. It is possible, however, for an outlying data point to have a detrimental effect because all data points are treated equally. The likelihood of this occurring may be greatly reduced by ensuring that each data point represents a large number of responses and equal numbers of responses are counted across data points. The bulk of the psychometric functions in this thesis were fitted with a logistic curve (Figure 46) whose equation is shown below.

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

Where μ is the test duration corresponding to the point of subjective equality between the reference and test durations, and θ provides an estimate of the discrimination threshold.

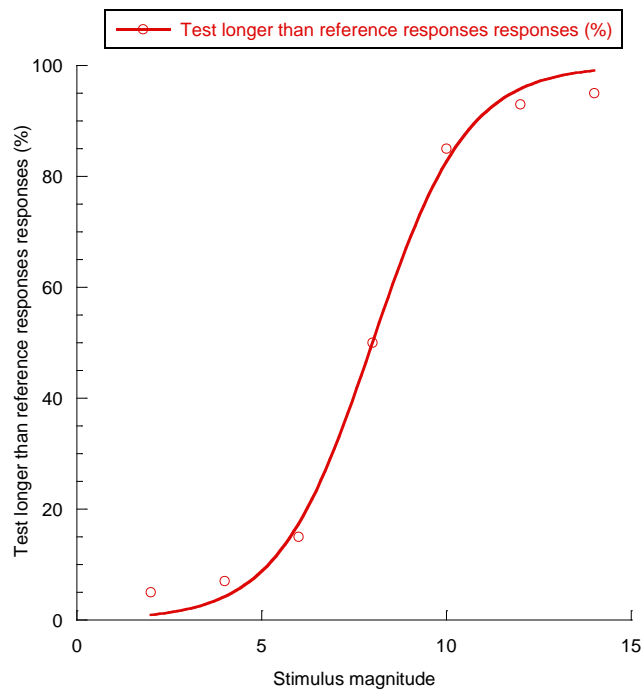


Figure 46 The point of subjective equality is the test duration where test longer than reference responses = 50% on the Y axis. In other words, physical test duration is perceived as being matched to the standard/reference duration.

The logistic function is popular and relatively straight forward method of curve fitting and was thought to be suitable for the experiments described in this thesis, although other methods such as Weibull or cumulative Gaussian would be likely to produce similar results (Klein, 2001, Strasburger, 2001).

The majority of experiments asked subjects to judge the relative durations of auditory and visual stimuli, and the psychometric functions produce two values of interest. Firstly, we are interested in the point at which the visual

and auditory stimuli are perceived as being the same duration. This is known as the point of subjective equality or PSE and corresponds to the physical test stimulus duration that produced perceptual equivalence between standard and test durations

This is the 50% point of the function where subjects are unable to judge which stimulus is longer and have to guess (Figure 46).

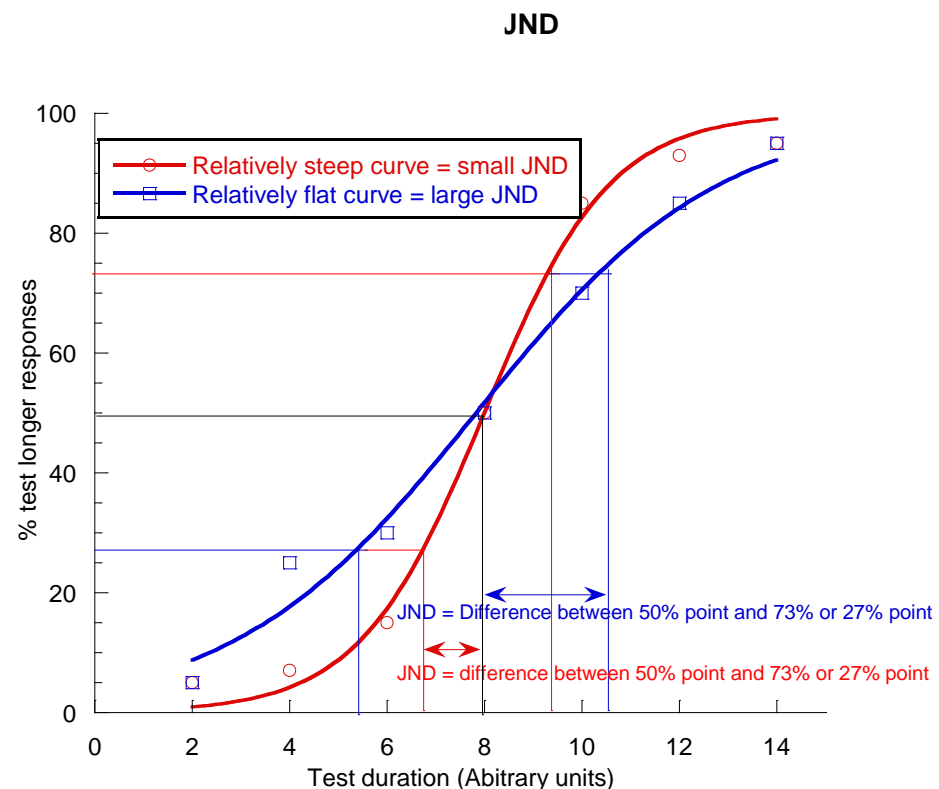


Figure 47 JND is defined as the difference between the point on the X axis corresponding to 50% and the points corresponding to the 73% or 27%. Thus a steep curve gradient corresponds to a small JND and a shallow gradient corresponds to a large JND.

Secondly, we can obtain the difference threshold. This is the minimum difference in physical duration between stimuli which subjects can reliably report that one is duration is longer than the other, often referred to as the “just noticeable difference” or JND. This value is dependent on the slope of the function as can be seen from Figure 44. Different criteria may be used to calculate this. For the experiments described in this thesis JND is defined as the difference between the test stimulus duration which produces a 50% test duration longer than reference duration response from the subject, and the test stimulus duration which produces a 73% (or 27%) test duration longer than reference duration response from the subject. Thus, when the slope of the function is shallow, the JND is relatively large and when the slope is steep, JND becomes relatively small (Figure 47).

The range of test duration values chosen for the experiments is important as this will influence the quality of the data obtained. If the spread of durations used is too wide, the extremes will make little contribution to JND or PSE estimates. This would result in useful data being extracted from only a very small number of the points and much data being wasted (Figure 48 left).

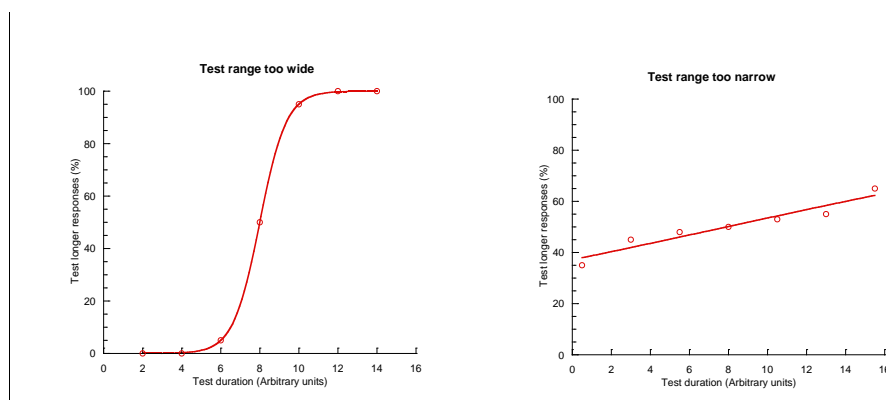


Figure 48 the resulting functions when the range of data points is too wide (left) and too narrow (right).

Also each of the 3 points could have an undue influence on the data extracted from the function. On the other hand, if the range used is too small (Figure 48 right), subjects may find the judgments too difficult and the influence of internal noise would be much greater over a large range of data points and the result would be a shallow function with many of the points close to the 50% “chance” zone of the function. If the responses do not extend beyond the 23% and 73% levels, then estimates of JND will be uncertain. For this reason, the choice of the test durations used is of vital importance as are subject instructions. It is important that when subjects are unable to discriminate between durations they do not guess in the same direction all the time and therefore “skew” the results in one direction or the other. In addition, a large number of repetitions per data point are required, so that the variability of individual responses due to internal noise may largely cancel each other out. In this way reliable data may be obtained. This method is arguably the most thorough and precise psychophysical tool available for this type of experiment (Norton et al., 2002).

A possible problem with the method of constant stimuli is that subjects may choose to ignore the reference stimulus and instead use an “internal mean” of test durations in order to make their judgments. In order to avoid this, in some of the early experiments, the reference stimulus duration was jittered by up to plus or minus 20%. However it became apparent that subjects were not employing this strategy and for the later experiments this method was dropped.

Another possible criticism of the method of constant stimuli is that it is possible to introduce a “time error”. It has been shown that when the test stimulus is always presented second, more “test longer” responses are produced than when they are presented the other way around. Some studies get around this by presenting the standard stimulus first in half the trials and second in the other half (Gescheider, 1985). This was not considered necessary for the experiments described in Chapter 9, since any time order error should apply equally to the all conditions. The majority of the experiments described in this thesis are of the method of constant stimuli variety. The exceptions to this are the contrast matching used in the spatial frequency duration experiments described in Chapter 8 which used a two alternative forced choice paradigm together with a staircase program known as QUEST (Watson et al., 1983), and the reproduction experiments described in Chapters 10 and 11 which fall into the duration scaling category of methods.

The staircase method may be thought of as a refinement of the method of limits which involves presenting a stimulus with whichever property we are interested in set at a very low level in order that it may not be detected by the subject. The property is then gradually increased until the subject is able to detect it. This is known as the ascending method of limits. The descending method of limits is the same except that the property in question is set well above threshold and gradually reduced until the subject can no longer detect it. The ascending and descending methods are run alternatively and the thresholds obtained are averaged. This method has the disadvantage that subjects may get used to responding that they detect the property and so

continue to give a positive response even when it is in fact below threshold. Alternatively, in the ascending condition they may anticipate that the property is about to become suprathreshold and respond prematurely.

The staircase method was developed to avoid these drawbacks. If, for instance, we ask the subject whether the visual stimulus was presented on the left or the right, a descending staircase will usually begin with an easily detected presentation randomly presented left or right. It then reduces stimulus intensity on each successive presentation until the subject makes an incorrect response. At this point the intensity is increased. This is known as a reversal.

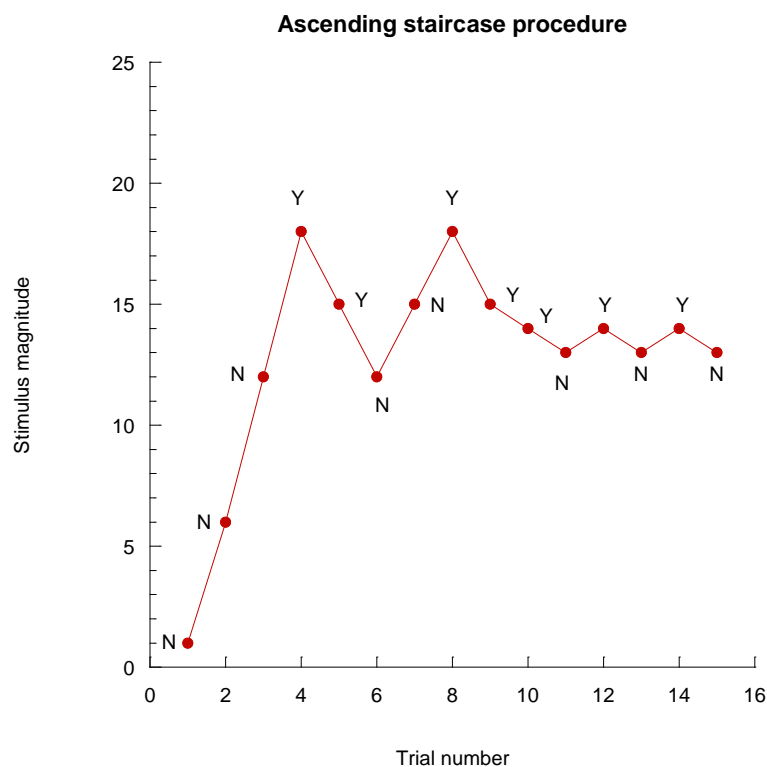


Figure 49 a single staircase procedure. Y denotes correct responses. N denotes incorrect responses.

The stimulus is then increased until the subject makes a correct response whereupon the intensity is reduced and so on until a predetermined number of reversals are reached. An ascending staircase begins with a very low magnitude stimulus and increases it for subsequent presentations until the subject makes a correct response. Then the stimulus magnitude is decreased and so on until a fixed number of reversals are obtained. The points at which the stimulus intensity is reversed are averaged and a threshold obtained (Figure 49).

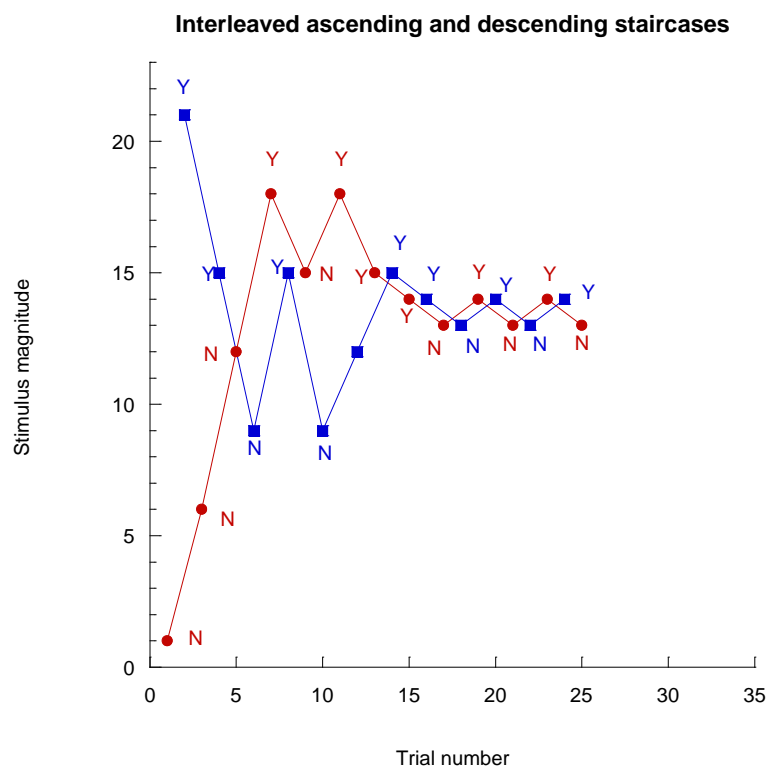


Figure 50 a double staircase procedure. N denotes incorrect responses. Y denotes correct responses. Blue denotes the descending staircase, red the ascending staircase.

This method is known as a standard staircase (Cornsweet, 1962). In its standard form the staircase has many of the disadvantages of the method of limits, in that subjects may make errors of habituation or expectation because the step sizes and direction of magnitude are predictable (Cornsweet, 1962, Levitt, 1971)

One way of overcoming this problem is to use two or more interleaved staircases as shown in Figure 50. As the trials progress, the different staircases converge on threshold, but because they are interleaved in a random fashion the subject is unable to predict which staircase is presented from trial to trial.

Another problem with the traditional staircase is that a poor choice of initial stimulus intensity may lead to a large number of wasted trials (Levitt, 1971). This problem may be overcome by pre-testing in order to discover the range of stimuli which are relevant for each subject or this may be established during the course of the experiment.

A number of variations of the staircase method have been developed in order to overcome the problem of subject stimulus prediction and establishing the optimum range. These include the transformed staircase (Levitt, 1971), PEST (Taylor et al., 1983), QUEST (Watson et al., 1983) and ML-PEST (Harvey, 1986), all of which involve differing rules of stimulus presentation relating to step sizes, rules for when step directions change, when the procedure ends and how the threshold is estimated. All these modifications are intended to improve the accuracy of the results obtained.

QUEST adapts the stimulus intensity whilst the experiment is progressing. The stimulus magnitude is estimated sequentially based on the experimenters' prior knowledge of the appropriate stimulus range and the subjects' response over the previous trials. The point at which different stimulus values appear equal is treated as a normally distributed random variable. This is a function that describes the relative likelihood that perceived equality will occur at any given value, based on previous subject responses. Following each response the Gaussian probability density function is updated. The intensity of each trial is set at the current maximum likelihood estimate of stimulus equality or the mode of the estimates, as is the final threshold estimate (Watson et al., 1983).

Finally, the method of stimulus reproduction was employed in experiments described in Chapter 10 and in Chapter 11. This involves subjects being presented with a visual or auditory stimulus for a particular duration and reproducing the duration by pressing a computer key for the required length of time. This method suffers from another potential source of noise due to the introduction of motor timing into the experiment. It also suffers from its criterion dependent nature (subjects may develop differing strategies to complete the task). It has been reported that subjects tend to overestimate short intervals and underestimate long intervals with this method (Tse et al., 2004). However the advantage of this method was that it allowed us to obtain an absolute measure of perceived duration rather than the perceived duration of a stimulus in one modality relative to the perceived duration of a stimulus in a different modality. It also allowed us to establish that duration adaptation did not transfer across modalities and provided a robust

confirmation of the results found using the method of constant stimuli which are detailed in Chapter 9.

7.4 Apparatus

All the experiments described in this thesis were carried out on one of two monitors. The majority of spatial frequency experiments as well as the cue combination experiments used a Compaq P1220 CRT with a resolution of 1280 x 1024 and a refresh rate of 100 Hz, which was driven by a dual-quad-core Apple Mac Pro desktop computer running Mac OS 10.4. The maximum luminance on this setup was 94cd/m² and the mean luminance was 47cd/m².

The adaptation experiments were carried out on a Sony Trinitron GDM FW 900 with a screen resolution of 1920 x 1200 and a refresh rate of 75 Hz, which was driven by an Apple Mac Pro desk top computer running Mac OS 10.5. The max luminance on this setup was 86cd/m² and the mean luminance was 43 cd/m².

The auditory and visual stimuli were generated using MATLAB 7 (Mathworks, USA) and Psychophysics Toolbox 3 (<http://www.psychtoolbox.org>). The auditory stimulus consisted of a burst of white noise of approximately 75dB presented via Sennheiser HD 280 headphones. The delivery of visual and auditory stimuli and the collection of subject's responses were controlled from within MATLAB using custom software. The physical duration of visual and auditory stimuli were given rectangular onset-offset profiles. All timings were

verified via simultaneous capture on a dual-channel oscilloscope: auditory voltage signals were obtained by accessing the input to the headphones whilst a visual voltage signal was produced using a photodiode attached to the monitor.

7.5 Gamma correction

Gamma correction allows us to control the luminance of our visual stimuli accurately. We can do this by controlling the relationship between the voltage input to the monitor and the resulting luminance of the image. Variations in this relationship affect the brightness of and the ratios of the red, green and blue “guns”. Since the stimuli used in this thesis are white, we require the input of the red, green and blue guns to produce equal intensities of light. Computer monitors do not normally have a linear relationship between voltage input and intensity, so in order to accurately produce the required visual stimuli for our experiments we need to correct for this before the image reaches the monitor. We can achieve this using software. First, we need to establish the present intensity to voltage relationship. We do this by taking a range of luminance readings from the monitor using a photometer. These readings correspond to a range of voltage values between 0 (minimum for the monitor) and 1 (maximum for the monitor). These values may then be plotted (Figure 53 solid line). We can use our software to produce an equal and opposite plot (Figure 53 dashed line) which can then be used to correct

the original input/output relationship in order to produce a linear input/output ratio from 0 to 1 (Figure 53 dotted line).

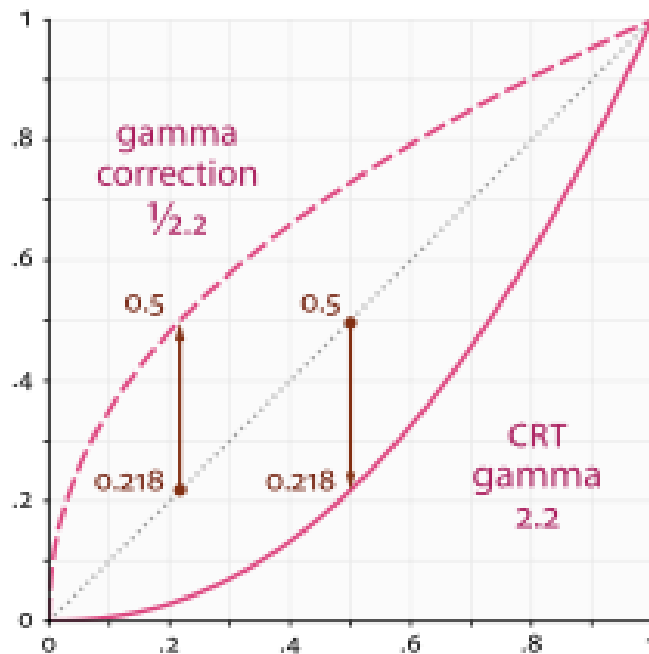


Figure 51 An example of gamma correction showing the original input/output plot for the monitor (solid line), the correction plot (dashed line) and the resulting linear input/output relationship (dotted line).

<http://en.wikipedia.org/wiki/File:GammaFunctionGraph.svg>

Chapter 8

Spatial frequency and perceived duration

8.1 Introduction

The first stimulus presented in a series of identical visual stimuli has been found to be perceived as longer than subsequent stimuli of equal length. It has been proposed that this is due to the increased attention given to the first stimulus (Rose et al., 1995). If an unexpected (different) stimulus is presented within a series of identical stimuli, it too is perceived as having a longer duration than temporally adjacent stimuli (Ulrich et al., 2006(a)). This is sometimes known as the “oddball effect” (Chapter 4.2).

In this Chapter a series of experiments are described in which observers were presented with standard and oddball stimuli that, in phenomenological terms, were obviously different from one another (Figure 52). However, in a departure from previous oddball studies, these differences were limited to a tightly controlled low-level parameter: spatial frequency. This approach had two advantages. Firstly, it minimises the higher-level cognitive factors that cloud the existing oddball literature. Secondly, it allows the introduction of carefully circumscribed levels of difference in standard and oddball appearance (for example, the difference shown in Figure 52 reflects a 2 octave change in spatial frequency). Surprisingly, data from experiment 1

shows that these low-level stimulus differences fail to induce the classic oddball effect, suggesting that stimulus complexity is perhaps a necessary component of the effect and that other cognitive factors must also be involved. In experiment 2 the effect of spatial frequency itself on perceived duration is investigated.

8.2 General methods

8.2.1 Observers

Six observers participated in Experiment 1 (3 authors (Aaen-Stockdale et al., 2011) plus 3 naive subjects) and 6 participated in Experiment 2 (4 authors plus 2 naive subjects).

8.2.2 Apparatus

All stimuli were presented on a Compaq P1220 CRT monitor or via Sennheiser HD 280 headphones. (See Chapter 7.3 for details). Stimuli were viewed from 57 cm and a headrest was used to ensure this.

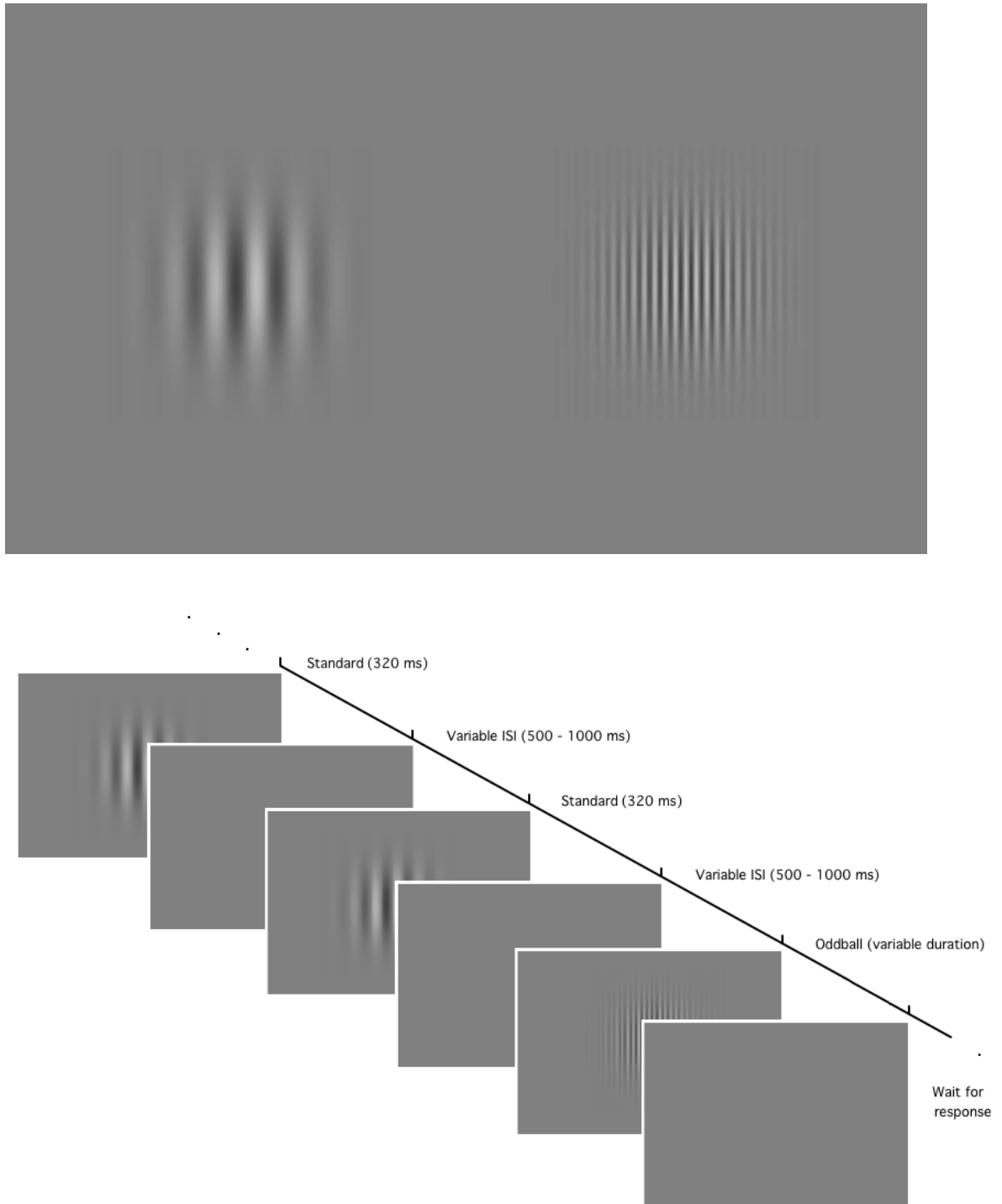


Figure 52 An example of the visual stimuli utilised in all three experiments. The top row of this figure shows two Gabor patches of 2 cpd (left) and 8 cpd (right), the values used for the oddball and standard in experiment 1. The stimuli shown here are of equal contrast, whereas in the actual experiments we presented Gabors of equal perceived contrast. The bottom row shows a schematic of a single trial. In this example the lower (2cpd) spatial frequency stimulus formed the standard stimulus and the higher (8cpd) spatial frequency stimulus formed the oddball stimulus.

8.2.1 Stimuli

The visual stimuli in all experiments were Gabor patches composed of a sinusoidal grating carrier multiplied by a spatial Gaussian envelope presented on a mean luminance background (Figure 52). The Gabor patches were presented in the centre of the screen. The grating component of the Gabor was of varying spatial frequency, but the envelope had a constant standard deviation of 2.7° . Therefore the size of the stimulus was the same for all conditions. The auditory stimulus was a burst of white noise presented binaurally via headphones. The intensity of the auditory stimulus was 65dB. On and offset profiles of both the visual and auditory stimuli were rectangular.

8.3 Experiment 1

8.4 Procedures

This experiment comprised of a typical oddball task based on previous experiments in the literature. The “standard” stimulus was a 2 or 8 c/deg Gabor patch a series of which were presented for 320 ms to the observer separated by a blank screen for a variable inter-stimulus interval. Initially observers were presented with 10 standard stimuli in order to build up an internal representation of the standard duration. Following this initial phase, each trial consisted of a minimum of 3 and a maximum of 10 presentations of

the standard stimulus followed by an “oddball” which was the opposite frequency (2 or 8 c/deg) to the standard. The duration of the oddball was varied symmetrically (Seifried et al., 2010) around the standard duration in seven steps of 20ms (oddball durations were 260, 280, 300, 320, 340, 360 and 380 ms). The observer then had to report (via a keypress) whether the oddball appeared longer or shorter than the standards. In between presentations of the standard or oddball stimulus, the screen was mean luminance grey for a variable inter-stimulus interval (isi) of between 500 and 1000 ms. All stimuli (standards and oddballs) were presented at the same spatial location, the centre of the screen and the phase of the sinusoidal component of the Gabor patch was varied randomly on every presentation. Each of the oddball durations was presented 5 times within each experimental block. Blocks were repeated 4 times to give 20 observations per point. Based upon previous literature (see Chapter 4.2), it was hypothesised that we would obtain an expansion of subjective time in both conditions (standard 2 cpd, oddball 8cpd and vice-versa). It was also expected that this expansion of perceived duration would be roughly equal, on the basis that the spatial frequency of the oddball in both conditions differed from that of the standard Gabor by the same number of octaves.

8.5 Equating for visibility

It has been suggested that perceived duration is systematically biased by the contrast or luminance of a stimulus with the most common finding being that

brighter (or higher contrast) stimuli are perceived as longer in duration (e.g. Terao et al (2008)). Previous oddball-based studies using images or complex geometric stimuli have not equated for such low-level image characteristics. In order to investigate the effect of stimulus novelty under conditions of matched stimulus visibility, an initial experiment was conducted where we controlled for differences in perceived contrast at different spatial frequencies. The lower spatial frequency (2 c/deg) Gabor was set to 50% contrast and asked observers were asked to match the perceived contrast of the high spatial frequency (8 c/deg) Gabor to this value. A temporal 2AFC task was employed in which the observer was presented with either the 2 c/deg Gabor followed by the 8 c/deg Gabor (or vice versa) at the centre of the screen and had to report which interval contained the higher contrast grating. The contrast of the 8 c/deg Gabor was determined by a QUEST staircase (Watson et al., 1983) to match its perceived contrast to the 2c/deg stimulus. Three separate Quest staircases were run and the mean taken. The appropriate contrast value for each observer was then used in Experiment 1. This contrast matching paradigm was adopted because the stimuli used in this experiment were highly suprathreshold. Due to contrast constancy in the visual system, it would not have been appropriate to present stimuli at very large multiples of detection threshold since that may not result in equally “visible” stimuli (Georgeson et al., 1975).

8.6 Results

A logistic function of the form $y = \frac{100}{1 + \exp\left(-\frac{(x-\mu)}{\theta}\right)}$ was fitted to the raw data for each observer, from which the position of subjective equality (PSE – μ in the above equation) - the physical oddball duration that was perceived to match the standard duration - was extracted. Samples of the resultant psychometric functions from one observer are shown in Figure 53.

Surprisingly, approximately equal and opposite shifts from veridical are found depending upon the spatial frequency relationship of the standard to the oddball, rather than the expected subjective expansion of all oddballs. With a standard of 8 c/deg and an oddball of 2 c/deg (green curve), we see a decrease in the PSE value, signifying an expansion in the perceived duration of the oddball, consistent with previous reports. However, when the stimuli are reversed (red curve) we see an increase in PSE values, signifying a *contraction* of the perceived duration of the oddball. PSE values for each observer and the mean PSE for the two conditions are summarised in Figure 54, with bars colour-coded as in Figure 53.

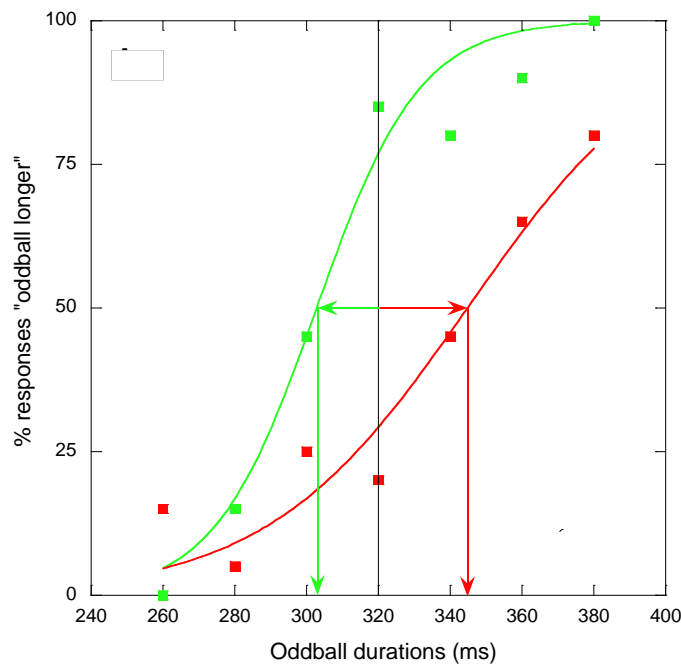


Figure 53 Example psychometric functions for one observer. The green curve shows data from the condition in which the standard Gabor patch is 8 cpd and the oddball Gabor patch is 2 cpd. The red curve shows data from the condition in which the standard is 2 cpd and the oddball is 8 cpd. Arrows show the shift in the psychometric functions from veridical and the corresponding PSEs.

For all observers, when the oddball is lower in spatial frequency than the standard (green bars), there is a decrease in the PSE relative to veridical. Similarly, for all observers, when the oddball is *higher* in spatial frequency than the standard (red bars), there is an *increase* in the PSE from veridical. In order to show this differential effect more clearly, a modified version of the *temporal expansion factor* (TEF) used in previous studies (Tse et al., 2004) was calculated. The TEF is the standard duration divided by the oddball PSE.

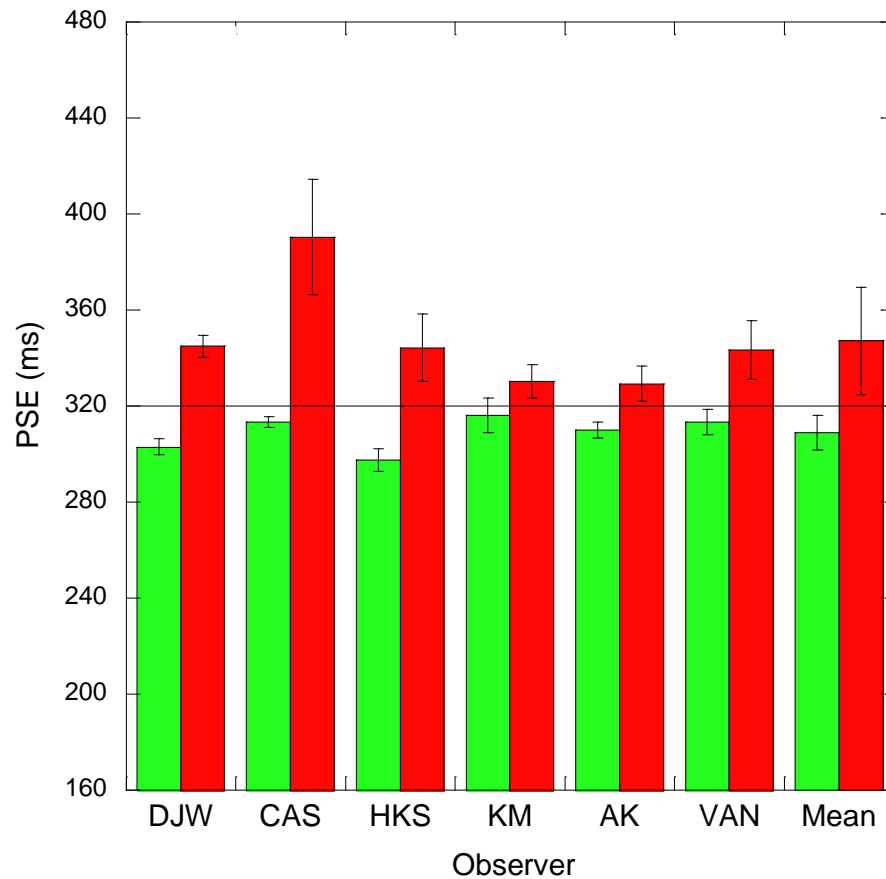


Figure 54 shows PSE data for all observers (colour coding identical to Figure 53). Error bars for individuals show the error of the PSE extracted from the logistic fit to the data. Error bars for the group show the standard deviation.

This value was calculated and subtracted from 1, which gave a positive value for temporal expansions and a negative value for temporal contractions. This analysis can be seen in Figure 55. Notice that all green bars are positive and all red bars are negative.

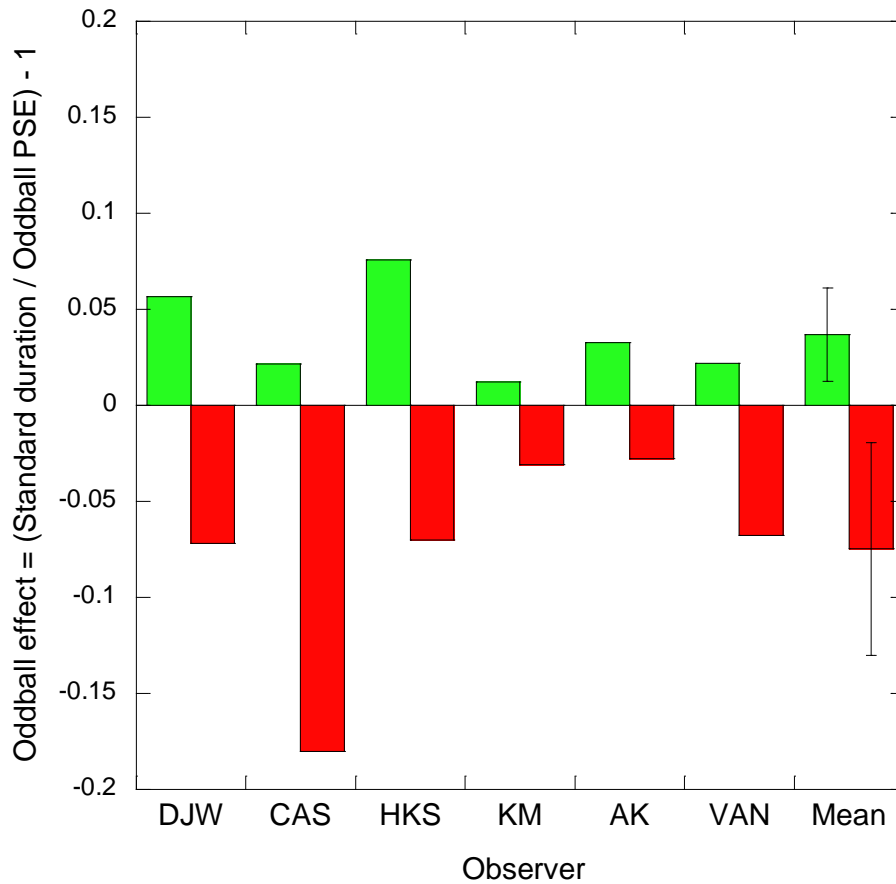


Figure 55 shows that expansion occurs when the oddball is lower in spatial frequency than the standard (2 cpd green), while temporal contraction occurs when the oddball a higher spatial frequency than the standard (8 cpd red). The direction of the effect is modulated according to the spatial frequency relationship of the stimuli, not their “differentness”.

A paired-samples *t*-test showed that the difference between the perceived duration of the oddball in the different conditions was significant ($t(5) 4.1$,

$p < 0.05$, $r = 0.1$), but the small sample size resulted in a significantly non-normal distribution of scores for the condition in which the oddball was of a higher spatial frequency (Kolmogorov-Smirnov, $D(6) = 0.372$, $p = < 0.05$). In light of this non-normality a Wilcoxon rank-sum test was also carried out which also resulted in a significant difference between the conditions ($z = 2.201$, $p < 0.05$). There was no systematic difference in slope between the two conditions ($t(5) = 1.393$, $p > 0.05$, $r = -0.236$).

Given that both oddballs shared matched differences in spatial frequency - relative to their respective standard stimuli - it seems unlikely that the expansion shown for the 2c/deg oddball condition and the contraction shown in the 8 c/deg oddball condition result from separate mechanisms. Rather, the data in Figure 55 suggest that the spatial frequency *per se* may be the dominant factor governing the perceived duration of the stimuli employed in Experiment 1. In Experiment 2, we removed issues surrounding stimulus expectancy and investigated perceived temporal extent as a function of spatial frequency.

8.7 Experiment 2

8.8 Methods

The differences in perceived duration that we obtained in Experiment 1 could be the result of either differential visual persistence of the stimuli or caused

by an inherent biases in the temporal processing of different spatial frequencies. These two different explanations result in two very different predictions. A persistence effect would manifest itself as a constant *additive* difference in perceived duration, which would be proportionally smaller as stimulus duration increased (equivalent to a perceptual increment added to the perceived duration of a particular spatial frequency, which would be constant across different physical durations). On the other hand, a ‘faster’ or ‘slower’ clock for different spatial frequencies would result in an effect that was a constant *proportion* of stimulus duration. In order to test for this, the auditory standard had average baseline durations of 160, 320 or 640 ms, +/- 20%. The trial-to-trial duration of the auditory standard was jittered by +/- 20% around each of these average baseline values so that the observer was forced to pay attention to the auditory standard and could not simply opt to ignore the duration of the noise and compare the duration of the visual stimulus to an internally generated standard. Thus, the exact duration of the standard therefore varied within a block and the average duration was varied between blocks. The duration of the Gabor patch stimulus was chosen from among seven durations that were equally spaced around, and centred on, the standard duration for *that trial*. Logistic functions were then fitted to this raw data for each observer, spatial frequency and baseline duration. From these psychometric functions, the position of subjective equality (PSE) was extracted in the same fashion as previously described for experiment 1. Since the raw data were expressed as percentages of the standard duration the PSEs were multiplied by the relevant three baseline durations (160, 320, and 640ms), so that we could express all the data in millisecond terms.

On each trial, an auditory burst of white noise was presented as a “standard”, after which a Gabor patch with a spatial frequency of 0.5, 1, 2, 4 or 8 c/deg was presented. The spatial frequency of the Gabor patch was randomly interleaved within a block according to the method of constant stimuli. Observers were then asked to make a two alternative forced choice duration discrimination judgment as to ‘which was longer, the visual or auditory stimulus?’ and responded via a keypress. The sound was always presented first.

8.9 Equating for visibility

As with experiment 1 the visual stimuli used were highly superthreshold and therefore may not be equally visible to observers. For this reason visual stimuli were equated for visibility. For each baseline duration (160, 320 and 640 ms), the perceived contrast of the 0.5, 1, 4 and 8 c/deg Gabor patches was equated to that of a 50% contrast, 2 c/deg Gabor of the appropriate duration using a 2AFC task and interleaved QUEST staircases (Watson et al., 1983). Three separate QUEST staircases were run and the mean taken. The appropriate values for each observer were then used in Experiment 2 (Figure 56).

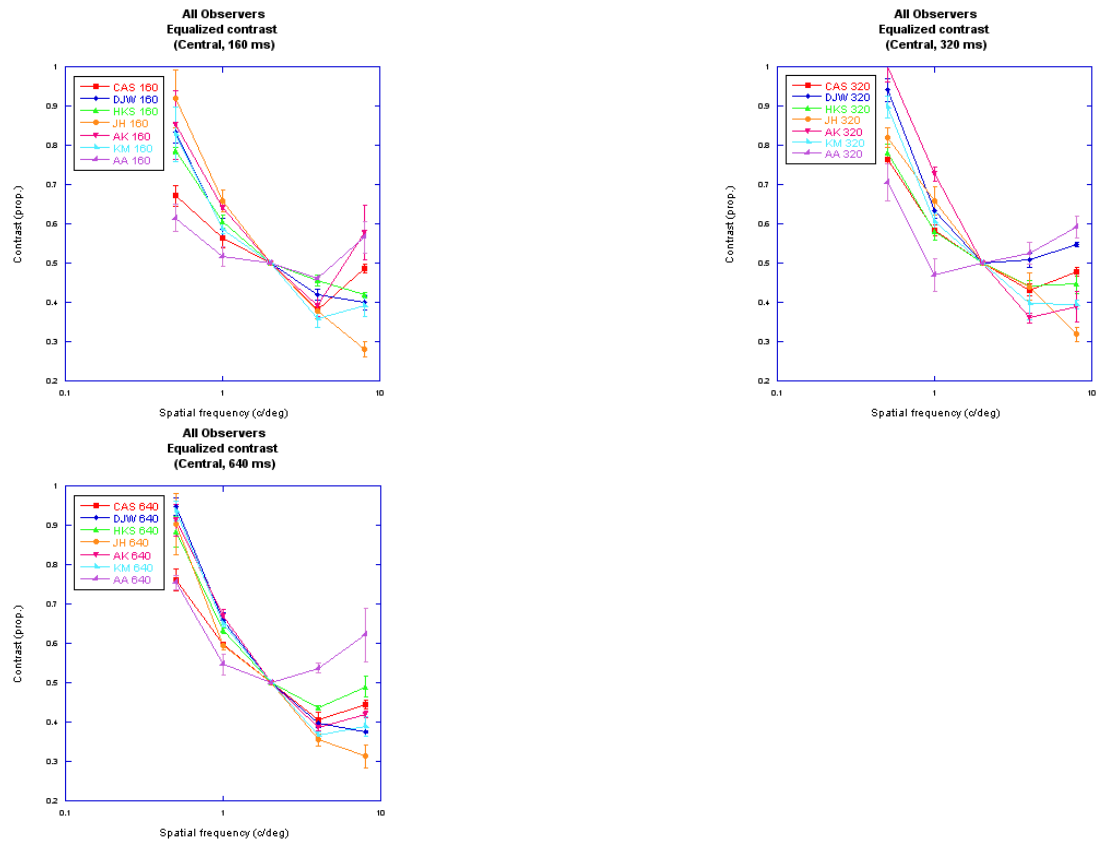


Figure 56 Equivisibility curves for all observers at stimulus durations 160 (left), 320 (right) and 640 (bottom left) milliseconds. Data points show the contrast at which a particular SF was perceived as the same contrast as a 50% contrast 2 c/deg Gabor.

8.10 Results

The average PSEs for seven observers are shown in Figures 57 and 58.

The data are fitted with a Gaussian function of the form

$$y = h/e^{-(\log(x/f))^2/2\sigma^2}$$

Where f is the spatial frequency corresponding to the minimum of the curve, h is the duration in milliseconds at the minimum point of the function and σ is the standard deviation.

Figure 57 shows a clear effect of spatial frequency on perceived duration. The 'u-shaped' distribution of these data shows that, relative to the higher and lower ends of the spatial frequency range, middle spatial frequencies are perceived as having a longer duration. This effect appears maximal at around 2 c/deg, and, the effect appears to be constant, in millisecond terms, as baseline duration increases. This can be seen if we change the y-axis from a linear scale (Figure 57) to a logarithmic scale (Figure 58). Plotted like this, the functions are progressively shallower at longer durations, reflecting a proportionally smaller effect of spatial frequency as baseline duration increases. These effects were confirmed by a two-way repeated measures analysis of variance, which revealed that the effect of both baseline duration ($F_{2,10}=936.9$, $p<0.001$) and spatial frequency ($F_{4,20}=16.17$, $p<0.001$) were highly significant. Importantly, however, there was no significant interaction between these two parameters ($F_{8,40}=0.578$, $p>0.1$) indicating that, in

absolute terms, the spatial frequency effect was similar across baseline durations.

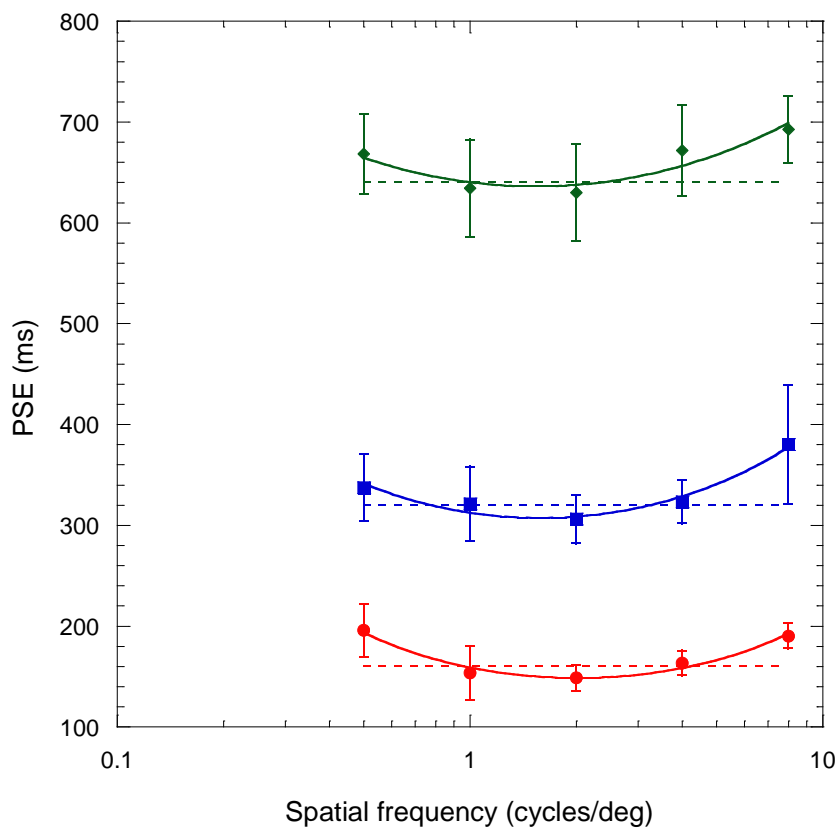


Figure 57 Individual PSEs for five different spatial frequencies and three different baseline durations of 160ms (red), 320ms (blue) and 640ms (green) averaged across observers ($n=7$). The PSE data are presented on a linear scale. Dotted lines represent the veridical duration of the auditory standard. Error bars show the standard deviation from the group mean.

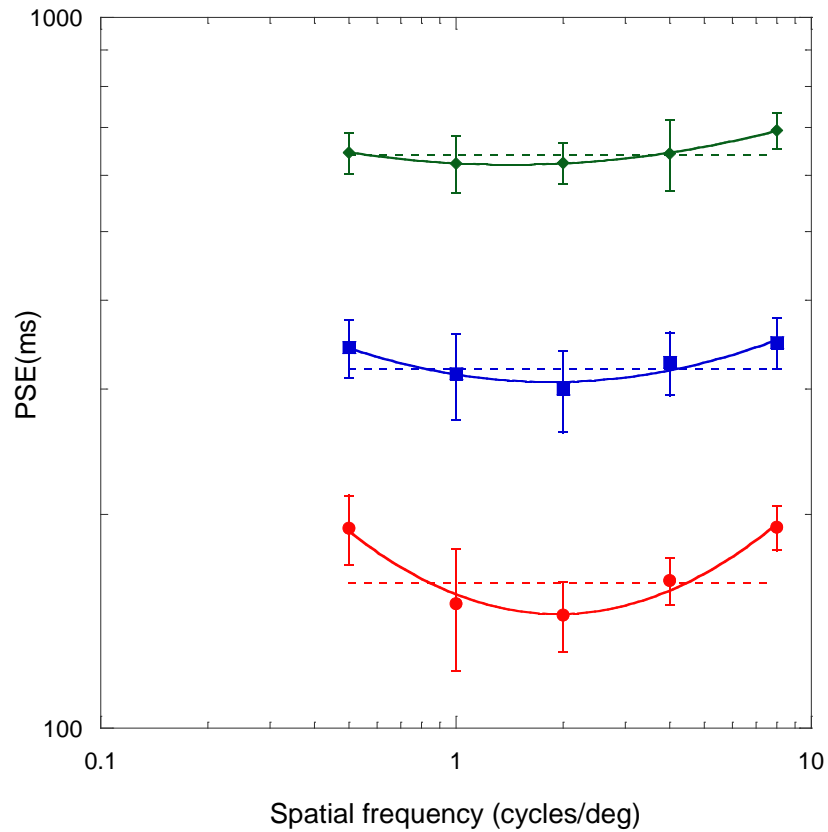


Figure 58 Individual PSEs for five different spatial frequencies and three different baseline durations of 160ms (red), 320ms (blue) and 640ms (green). The data are presented on a log scale. Dotted lines represent the veridical duration of the auditory standard. Error bars show the standard deviation of the group mean.

The spatial frequency dependence of duration perception found in Experiment 2 appears to peak between 1.4 and 2 c/deg ($f = 1.94$ for 160 ms, 1.86 for 320 ms and 1.42 for 640 ms). The minimum of the curve is consistently 10-20 ms lower than the veridical duration ($h = 145$ for 160 ms, 307 for 320 ms and 620 for 640 ms), showing that mid-range spatial

frequencies are perceived as *longer* than the auditory standard. Around 1 or 4 c/deg performance is veridical to the auditory standard, while at the extremes perceived duration is 20-50 ms *shorter* than veridical (i.e. the actual duration of the auditory standard).

A possible interpretation of these results might be that the increase in perceived duration for mid-range frequencies is due to the 2cpd Gabors being in the middle of the range used. In order to test for this we repeated Experiment 2 using Gabors centred on 320ms duration, for three spatial frequencies (0.5cpd, 2cpd, and 8cpd).

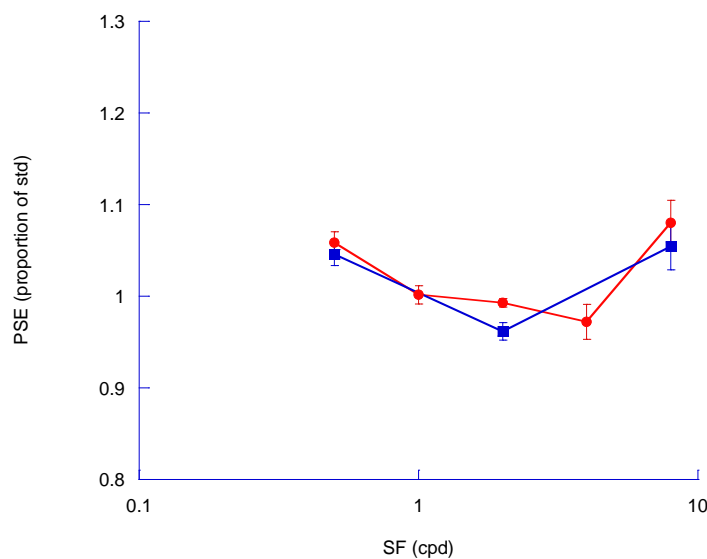


Figure 59 Representative data from a single subject following a control experiment in which three different spatial frequencies were tested independently of one another in separate blocks. As with the main experiment 2cpd Gabors are perceived as being longer in duration than their 0.50cpd and 8cpd counterparts. The y axis denotes PSE as a proportion of 320ms (1=320ms); the x axis denotes spatial frequency.

This time each spatial frequency was tested independently of the others. Data from a representative observer is shown in Figure 59.

The original data for all 5 spatial frequencies for this subject are shown in red, whilst the data from the independently tested spatial frequencies is shown in blue. It may be seen that, as in the main experiment 2cpd frequency Gabors are perceived as being longer in duration than either 0.5cpd or 8cpd Gabors.

8.11 Discussion

The results demonstrate two key findings: Firstly, oddball-related temporal expansion cannot be solely attributable to perceived differences between standard and oddball stimuli: when high level content is minimised, yet (grossly supra-threshold) low-level differences persist; we failed to reproduce the temporal expansion found elsewhere in the literature. Secondly, the data clearly show a significant bias towards perceiving mid-range spatial frequencies as longer in duration than high or low frequencies. The effect occurs when visual stimuli of different spatial frequency are compared to each other (Experiment 1) or are compared cross-modally to an auditory standard (Experiment 2).

Within the time perception literature a widely reproduced finding is that sounds are typically perceived as being longer than lights, despite having the same physical duration (Behar et al., 1961, Goldstone, 1968, Goldstone et al., 1974, Walker et al., 1981c, Wearden et al., 1998). The data suggest that this

bias is dependent upon the frequency content of the visual stimulus. At mid-range spatial frequencies, the perceived duration of the visual stimuli actually *exceeded* the duration of the auditory standard. However, it is noteworthy that previous work comparing the perceived duration of sounds and lights invariably used uniform geometrical stimuli whose spectral content would have been dominated by lower spatial frequencies. The finding that these visual stimuli are perceived as shorter than sounds is, in fact, consistent with our data for low spatial frequencies. If we consider the left hand portion of the curves shown in Figures 57 and 58, it can be seen that, for this particular choice of visual stimuli, sounds will indeed be perceived as being longer than lights (as shown by the vertical elevation of the data relative to the horizontal dashed lines). What the data highlights, however, is that such a finding is not a universal one, but depends critically upon the spatial frequency content of the visual stimulus.

Hughes *et al* (1992) investigated the perceived duration of images after low- or band-pass spatial filtering. They found that images with broader spatial frequency content were perceived as being longer in duration than those images with a narrower range of frequencies, regardless of whether the images contained predominantly low or high frequencies. Since widening the pass-band of the filters used to create the stimuli would necessarily cause more mid-range spatial frequencies to be included in the final image, this finding could be explained by the spatial frequency dependence we have demonstrated in this study and may not necessarily be an effect of broader spatial frequency spectra.

Kaneko and Murakami (2009) systematically varied the spatial frequency of drifting grating stimuli in order to investigate whether perceived duration was dependent upon the temporal frequency or the speed of a stimulus. They concluded that perceived duration increased with speed, but since they compared the duration of their moving gratings to a static comparison of the same spatial frequency, they essentially factored out any effect of spatial frequency *per se* on perceived duration.

Several studies in the visual persistence literature have investigated the effect of spatial frequency on the persistence of very short (<100 ms) duration visual stimuli. Over a similar range of spatial frequencies to ours, these studies have found that perceived duration of the stimulus increases monotonically with spatial frequency (Long et al., 1981, Meyer et al., 1977). The persistence of the *afterimage*, on the other hand, either decreases with increasing spatial frequency (Long et al., 1980, Long et al., 1981) or is band-pass, depending on mean luminance (Ueno, 1983). Methodological issues and problems of definition have clouded the persistence literature (Nisly et al., 1989, Bowling et al., 1982) and the emphasis on persistence beyond stimulus offset has neglected factors influencing stimulus *onset*. Baro *et al* (1992) demonstrated that reaction times to stimulus offset increased as spatial frequency increased, echoing previous findings. However, they also demonstrated that reaction times to stimulus *onset* increased with spatial frequency at the same rate, implying a constant perceived duration across spatial frequency. Finally, in all of these previous studies, no attempts were made to control for the visibility of the different spatial frequencies (Long et al., 1981), which makes comparison to our data rather difficult.

The envelope size of the Gabor stimuli was not varied and since the size of receptive fields scales with spatial frequency (Devalois et al., 1982), more individual detectors would be activated by the high spatial frequency stimuli than the low spatial frequency stimuli. The greater number of detectors, and therefore greater neural energy expenditure, could contribute to the greater perceived duration. However, this would predict a linear increase, not the u-shaped function we obtain. Future work could utilise stimuli of constant bandwidth to investigate this issue. Biases in temporal processing may be explained by an internal “clock” running faster or slower (Galazyuk et al., 1997). Another possibility is that – given accumulated experience about the possibility of commonly encountered durations – observers may have prior assumptions about the durations of certain stimuli, just as we have a tendency to impose shading patterns consistent with the “light from above” prior (Sun et al., 1998). The perception of mid-range spatial frequencies being longer in duration may then conceivably reflect a higher incidence of longer physical durations for images dominated by these frequencies. However, biases in both these mechanisms would be manifest as a constant proportional bias (Jazayeri et al., 2010), and with respect to the Bayesian explanation, it is unclear whether certain spatial frequencies would be physically present longer than others in natural vision. The bias in perceived duration that we find across spatial frequency is constant in millisecond terms, and therefore appears to reflect greater persistence for mid-range spatial frequencies. It is tempting to implicate low level factors such as intensity or contrast sensitivity to explain this bias in stimulus persistence, but by

equating the visibility of our stimuli on an observer-by-observer basis, these considerations are unlikely to form a convincing explanation for our effects.

It is noteworthy that no evidence of any consistent “subjective time dilation” in response to oddballs was found. The expansion and contraction effects we obtain in Experiment 1 appear to be approximately equal and opposite and appear to be explained entirely by the difference in perceived duration across spatial frequency observed in Experiment 2. Although stimuli equated for visibility were used, the standard was repeatedly presented and the oddball only infrequently presented, which may have resulted in some contrast adaptation in the spatial frequency channels tuned to the standard. Differences in contrast or intensity may lead to biases in perceived duration; the most common finding being that higher contrast/intensity is perceived as longer. However, if a consistent expansion in the perceived duration of oddballs as a result of arousal (e.g. Ulrich et al., 2006(a) information processing (e.g. Tse et al., 2004), repetition suppression (e.g. Eagleman et al., 2009) or contrast adaptation existed over and above the spatial frequency differences noted here, we would expect an overall reduction for all oddball PSE values shown in Figures 53 and 54. The extent of this reduction would be then be modulated via changes in spatial frequency. We do not see this. In addition, the differences in perceived duration we observe are also much smaller in magnitude (around 5% in either direction) than that usually reported in oddball studies. The primary difference between the current study and previous studies is the nature of the stimulus, which in our study is a narrowband, low-level stimulus. Previous studies have variously used dynamic, broadband or cognitively-engaging natural images and it may be

that our stimuli, chosen to selectively target low-level visual mechanisms are not “high-level” enough to evoke subjective time dilation. This suggests that “subjective time dilation” effects are essentially high-level in nature, with a neural locus beyond V1, and necessitate the use of complex, dynamic or cognitively engaging stimuli.

Having demonstrated a persistent bias in the perceived duration of equi-visible gratings of different spatial frequency, it is tempting to contemplate the perceived duration of a compound grating or plaid stimulus composed of multiple frequency components. Would perceived duration be computed in a winner take all fashion, with the longest (or shortest) duration dominating, or is perceived duration the mean of the different component durations?

Chapter 9

An investigation into duration adaptation

9.1 Introduction

Adaptation has been used successfully in the past to help reveal the neural underpinning of sensory processing. A range of stimulus features are now generally accepted as being processed using spatial or orientation “channels” – groups of neurons which are maximally sensitive to a particular stimulus value and whose response diminishes for example spatial frequency or orientation moves further from its preferred value (Blakemore et al., 1969(a), Blakemore et al., 1970). It has also been suggested that the perception of temporal extent may be mediated by similar mechanisms. (See Chapters 3.2.2 and 4.4 for details).

The experiments described here employ adaptation techniques and cross modal temporal judgments involving the method of constant stimuli, to test predictions made by channel-based (CB) models of temporal perception.

9.2 Materials and Methods

9.2.1 Participants

Nine observers (four authors and five naïve) participated in the experiments.

9.2.2 Stimuli

The visual stimulus was a 100% contrast isotropic Gaussian luminance blob of white light ($\sigma = 2.26$ degrees at a viewing distance of 57 centimetres) displayed against a uniform grey background (mean luminance 47cdm^2). The blob was presented at the centre of a gamma-corrected monitor screen (Sony Trinitron GDM FW900) which was driven by an Apple Mac Pro desktop computer running Mac OS 10.5. The visual stimulus was generated using Matlab 7.7 (Mathworks, USA) and psychophysics Toolbox 3 (<http://www.psychtoolbox.org>). The auditory stimulus consisted of a burst of white noise presented via Sennheiser HD 280 headphones. Delivery of visual and auditory stimuli and the collection of observer's responses were controlled from within Matlab using custom software. The physical durations of visual and auditory stimuli were given rectangular onset-offset profile were verified via simultaneous capture on a dual-channel oscilloscope.

9.3 Procedure

Observers adapted to sequences of visual or auditory stimuli with a fixed duration before making two interval, forced choice duration discrimination as to 'which had the longer duration - test or reference stimulus?' The test stimulus arose from the adapted sensory modality stimuli, whereas the reference stimulus arose from the non-adapted modality (Figure 60). The duration of the reference stimulus remained at 320ms, while test stimulus duration varied in seven logarithmically spaced steps from 237 to 421ms, which were randomly interleaved within a method of constant stimuli. Adapting duration was either 0 ('no adapt' baseline condition): (Figure 61, red data; Figure 62, blue data), 40,80,160,240,400,640,1280, or 2560ms and remained constant within each experimental block. Following an initial adaptation period comprising 100 adapting stimulus presentations, a 2000ms pause signalled the start of the 'top-up' phase, which constituted the presentation of a further four adapting durations followed by reference and then test stimulus presentations. Receipt of the subject's duration discrimination judgment (via keyboard) triggered the presentation of the next top-up and test cycle. The inter-stimulus interval (ISI) between adapting, top-up, reference and test stimuli was randomly jittered in the range 500-1000ms. Each block contained 10 repetitions of each test duration and three blocks were added together to give a total of 30 repetitions per condition. The presentation order of each block was selected by the presentation software in a pseudorandom order.

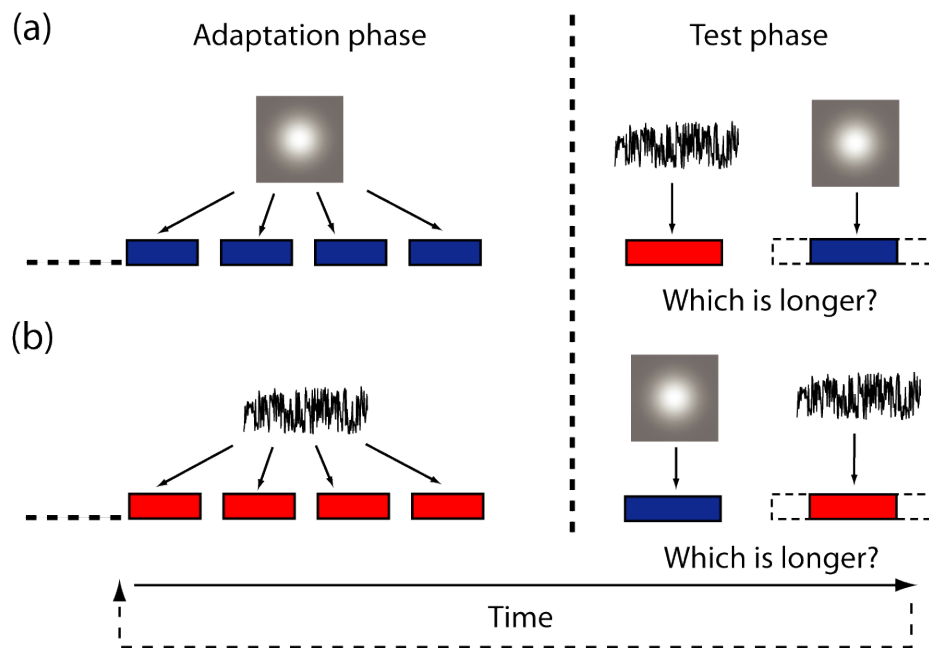


Figure 60 A schematic showing the paradigm used for the experiments. Observers adapt to either (a) Gaussian blobs (in blue) or (b) bursts of white noise (in red). The adaptation phase consists of 100 stimuli of identical duration (not shown) while the test phase consists of a reference stimulus from the opposing modality followed by a test stimulus (from the modality matching the adapting stimuli) of variable duration. In this example, adaptation stimuli are of a relatively short duration (e.g. 160ms) relative to the moderate duration reference (e.g. 320ms). The last four adaptation stimuli are repeated between test phases to form a “top up” phase. For simplicity, the ISI is shown here as fixed, whereas in reality it varied randomly.

9.4 Results

Psychometric functions comprising of observers' duration discrimination, judgements were plotted showing the proportion of 'test longer than reference' responses as a function of test duration. These functions were fitted with a logistic of the form

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

Where μ is the test duration value corresponding to the point of subjective equality (PSE;-the 50% response level on the psychometric function) and θ provides an estimate of duration discrimination threshold (approximately half the offset between the 27% and 73% response levels). In this way, PSE values were obtained for all observers. The PSE data shown in figures 62, 63 and 64 was fitted with a curve based on the first derivative of a Gaussian, namely

$$PSE = y_{pos} + \left[\left(A * \log \left(\frac{D}{x_{pos}} \right) \right) * e^{-\left(\frac{\left(\log \left(\frac{D}{x_{pos}} \right) \right)^2}{2\sigma^2} \right)} \right]$$

where D is the adaptation duration, σ the standard deviation of the Gaussian A , a constant related to the amplitude of the function and (x_{pos}, y_{pos}) the origin of the function. (Note that when Duration is equal to x_{pos} , PSE is equal

to y_{pos}). The maxima and minima of this function occur at adaptor durations $\pm\sigma$ log units from the origin, i.e. $\log(D/x_{\text{pos}}) = \pm\sigma$.

The half-amplitude of this function, which represents the magnitude by which the PSE deviates from baseline (ie the size of any illusion), is therefore given by

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

9.4.1 Experience dependent duration plasticity

Psychometric functions for a representative naive observer are shown in Figure 61. The functions correspond to two adapting durations (160ms and 640ms) that represent a 1 octave difference from the centre of the range of the test stimuli (320ms). The lateral separation in opposite directions from the ‘no adapt’ baseline condition (Figure 61, red data) of the green and blue functions show that adaptation clearly modulates the proportion of ‘test longer’ responses in a repulsive fashion.

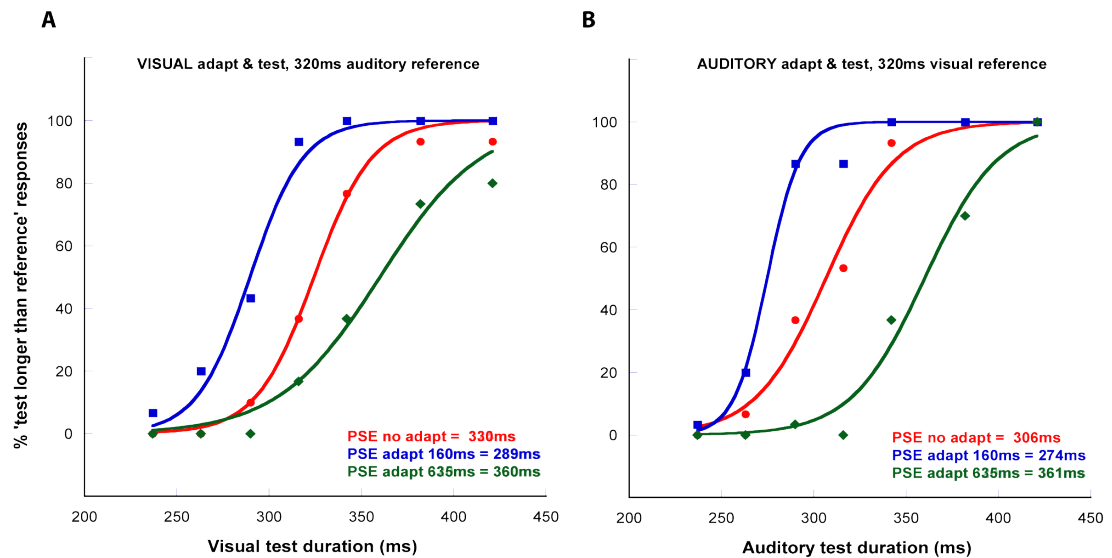


Figure 61 Sample psychometric functions from a single naive, representative observer (LEW) derived from duration discrimination judgements as to "which was longer, test or reference stimulus?" These functions correspond to judgements made in the absence of adaptation ("no adapt" red data) or following adaptation to 160ms or 640ms (a) visual and (b) auditory duration stimuli (blue and green data, respectively). The effects of adaptation are quantified by differences in the point of subjective equality (PSE): the physical test duration corresponding to 50 per cent "test longer" responses.

Specifically, adapting to relatively short visual or auditory durations (160 ms; Figure 61, blue data) induces an expansion in the perceived duration of test durations subsequently viewed (Figure 61a) or heard (Figure 61b). The magnitude of this effect is reflected in the physical test duration corresponding to perceived equivalence between test and reference durations (the PSE). For example, after adapting to 160 ms durations, visual PSE shifts from 330 to 289 ms, while auditory PSE shifts from 306 to 274 ms.

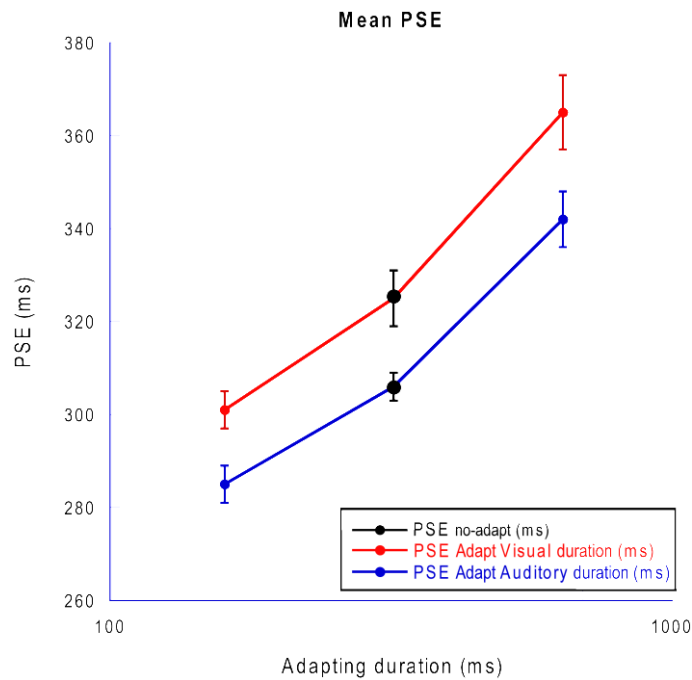


Figure 62 Average PSE shifts following adapting to visual (red) and auditory (blue) durations of 160ms (left) and 640ms (right). (n=9).

A reciprocal pattern is observed following adaptation to relatively long durations (640 ms), where test durations undergo perceptual compression (Figure 60, green data). Average PSE shifts show this effect to be consistent across observers (see Figure 62). A repeated measures ANOVA confirmed the significant effect of adapting duration on PSE ($F_{2, 16} = 99.3$, $P < 0.001$). It also confirmed that the PSEs for the visual condition were significantly higher than for audition ($F_{1, 8} = 17.4$, $p < 0.005$). This is consistent with sound being perceived as longer than vision. There was no significant interaction effect ($(F_{2, 16} = 0.40$, $p > 0.1)$) indicating that the difference between the two senses was consistent across adapting durations.

This pattern of repulsion-type after-effects is broadly similar to that observed following adaptation to consistent spatial information (DeValois et al., 1990). For example, prolonged viewing of visual stimuli of a relatively high spatial

frequency induces a decrease in the perceived spatial frequency of subsequently viewed stimuli (Blakemore et al., 1969(a), Blakemore et al., 1969(b)). This parallel suggests that a channel-based (CB) framework may be consistent with the duration after-effects shown in Figure 60. However, a further prediction of CB models concerns the relationship between after-effect magnitude and the degree of similarity of adaptation and test stimuli. This is exemplified by the finding that the influence of adaptation to consistent motion (Schrater et al., 1998), orientation (Gibson et al., 1937) and spatial frequency (Blakemore et al., 1970) is constrained to situations where adapt and test stimuli fall within a limited perceptual distance of one another. This distance is typically linked to the degree of sensitivity associated with the system's individual component channels (i.e. their bandwidth). In many cases, these psychophysical measurements map closely onto the underlying response properties of neurons at multiple scales of the visual system (Kohn, 2007).

9.4.2 Tuned duration after-effects

The possibility of duration-tuned mechanisms in humans was investigated by systematically altering the duration of the adapting stimuli while keeping the range of test stimuli constant. Average PSE values ($n = 9$) were extracted from the psychometric functions corresponding to each visual (Figure 63a) and auditory (Figure 63b) adaptation duration. Relative to the central 'no adapt' data point (in blue), increasing or decreasing the duration of the

adapting stimuli induces a decrease or increase in PSE (in red), respectively. This reflects a relative contraction and expansion of perceived duration, which appears to increase in an approximately linear fashion over a limited range of adapter durations. Beyond this range, adaptation magnitude declines such that the longest and shortest adapters (40ms and 2560 ms) induce changes in perceived duration approaching those observed in the no adapt condition.

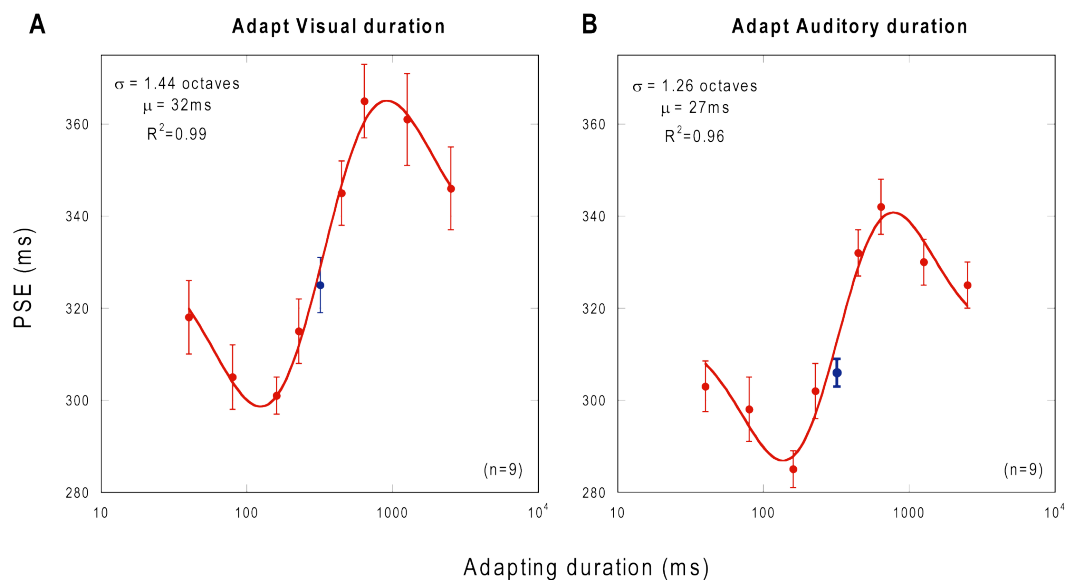


Figure 63 PSE data for all subjects, for a "no adapt" condition (in blue) and following adaptation to (a) visual and (b) auditory stimuli with 40, 80, 160, 240, 400, 640, 1280 or 2560ms durations (in red). Data are fitted with a curve based on the first derivative of a Gaussian, which provides two important parameters: μ , the functions half amplitude (the magnitude by which PSE deviates from baseline, or "aftereffect magnitude") and σ , standard deviation of the function (the temporal tuning of the adaptation). Error bars denote the standard error of the mean.

This pattern of results is markedly similar across visual and auditory domains (Figure 63). In order to characterize these effects, a curve based on the first derivative of a Gaussian was fitted to the data that allowed extraction of several important parameters.

While the amplitude of the visual and auditory functions (reflecting the magnitude of the adaptation effect) is similar, the bandwidths of the functions (in octaves) are slightly broader for vision than audition (1.44 versus 1.26). In other words, both modalities appear to possess approximately equivalent degrees of flexibility in response to duration adaptation, yet vision shows a greater tolerance to discrepancies between the duration of test and adaptor. Consistent with earlier reports (Behar et al., 1961, Goldstone et al., 1974, Walker et al., 1981b, Wearden et al., 1998), auditory durations are perceived as longer than their (physically identical) visual counterparts, irrespective of adaptation. This is reflected in the vertical offset between the two datasets shown in Figure 63, with a higher PSE indicating relatively shorter perceived duration.

9.4.3 Scaled, self-similar duration channels

In addition to the tuning features described above, channel-based perceptual systems are further characterized by a trend towards banks of overlapping channels that form self-similar, scaled versions of one another. For example, the bandwidth of channels responsible for processing auditory pitch (Patterson, 1976, Zwicker, 1961) or visual spatial frequency (Blakemore et al.,

1970, Sachs et al., 1971, Devalois et al., 1982) typically form a fixed proportion of the frequency to which channel is maximally responsive. When expressed in logarithmic terms, this gives rise to tuning functions that are approximately equivalent in appearance across a large range of stimulus parameters. Given that our range of test durations (237-421ms) contains substantial overlap with biologically significant durations such as those thought to be critical for speech perception (Ackermann et al., 1997, Schirmer, 2004), effects shown in Figures 60 and 62 may reflect duration mechanisms that are peculiar to this test range. Alternatively, if duration channels form a generalized feature of temporal judgements in the 'automatic' range (Karmarkar et al., 2007, Lewis et al., 2003), comparable versions of tuning data from figure 63 should be elicited by testing at different sub-second ranges.

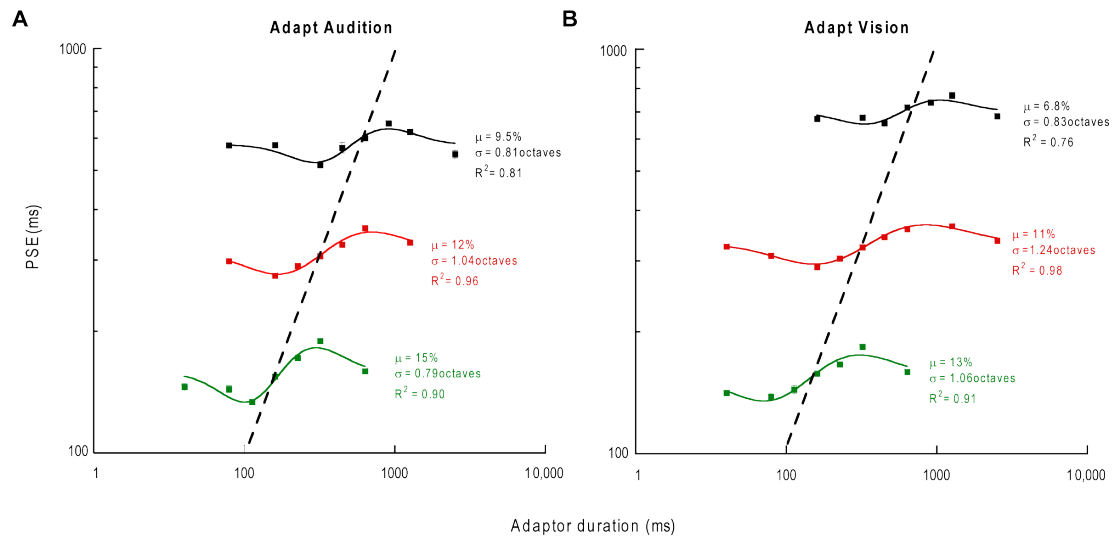


Figure 64 shows tuning data for representative observer DW showing PSE values as a function of (a) auditory and (b) visual adapting duration for test duration ranges centred on 160ms (green), 320ms (red) and 640ms (black). Note that red data points represent this observer's 320ms test range data, which form part of the group average data shown in Figure 62. The dashed black line represents a line of unit slope and illustrates the fact the three curves can be superimposed on top of one another by sliding along this line, indicating scaled self-similar mechanisms operating across test duration ranges. Error bars indicate the standard error of the mean.

This hypothesis was tested by examining the effect of duration adaptation on two further ranges of test durations. These experiments were identical to the main experiment except that the test duration ranges were centred on 160ms and 640ms, and were coupled with adaptation ranges spanning a three-octave range centred on the middle of the test duration range. Results for one representative observer are shown in Figure 64. For both modalities, longer (640 ms, black curve) and shorter (160 ms, green curve) test range data show a marked degree of similarity to the 320 ms range data (in red, as per Figure 63). Specifically, despite small variations in bandwidth and amplitude values across the different test ranges, a remarkable degree of

similarity is evident between the three functions. This pattern of results is replicated across observers as shown in Figure 65. Channels characterized by scaled, self-similar bandwidths are entirely consistent with the data shown in Figures 64-65.

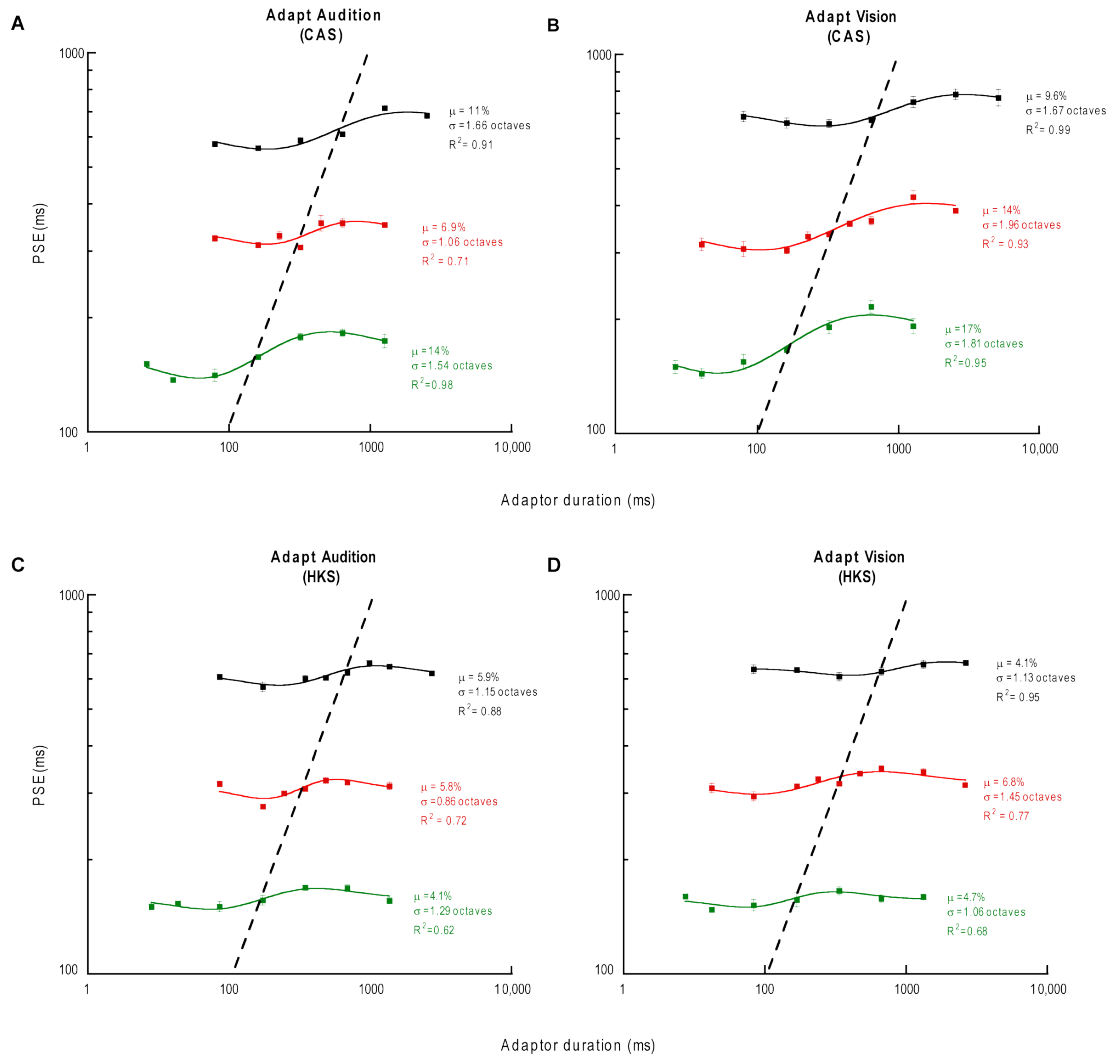


Figure 65 Data for two further subjects (CAS and HKS) showing PSE values as a function of (a) auditory and (b) visual adapting duration for test duration ranges centred on 160ms (green), 320ms (red) and 640ms (black). Note that red data points represent these observers 320ms test range data, which form part of the group average data shown in Figure 62. The dashed black line represents a line of unit slope and illustrates the fact the three curves can be superimposed on top of one another by sliding along this line, indicating scaled self-similar mechanisms operating across test duration ranges. Error bars indicate the standard error of the mean.

9.4.4 Modelling the effects of duration adaptation

Adaptation-induced biases in perception are typically explained using a common set of assumptions: (i) stimulus properties are encoded by populations of neurons with distinct (though typically overlapping) tuning curves; (ii) adaptation selectively changes the responses of these neurons; and (iii) downstream mechanisms that decode ('read out') the activity of the population are unaware of these changes (for recent reviews see (Schwartz et al., 2007, Series et al., 2009)). To determine whether it is possible to account for the effects of duration adaptation in a similar manner, a simple population coding model comprising sets of dedicated, modality-specific time channels was constructed. The intention was to establish a model capable of quantitatively describing our psychophysical data with the smallest set of assumptions possible.

To begin, a population of neurons with log-Gaussian duration tuning for each sensory modality was generated (Figure 66). Physiological evidence has previously been reported for this form of duration tuning across a range of neural structures (Duysens et al., 1996, Yumoto et al., 2011, Casseday et al., 1994b, Aubie et al., 2009, Galazyuk et al., 1997, He et al., 1997). Preferred durations were arbitrarily set to range from 1 to 1000 ms in equal log steps. The number of neurons (n) and the standard deviation (σ) of the tuning functions (fixed for each modality) were varied for different situations. Adaptation was modelled as a selective modality-specific reduction in response gain that was maximal at adapted duration (A_{\max}) and fell off with log-Gaussian profile (width set by A_{σ}).

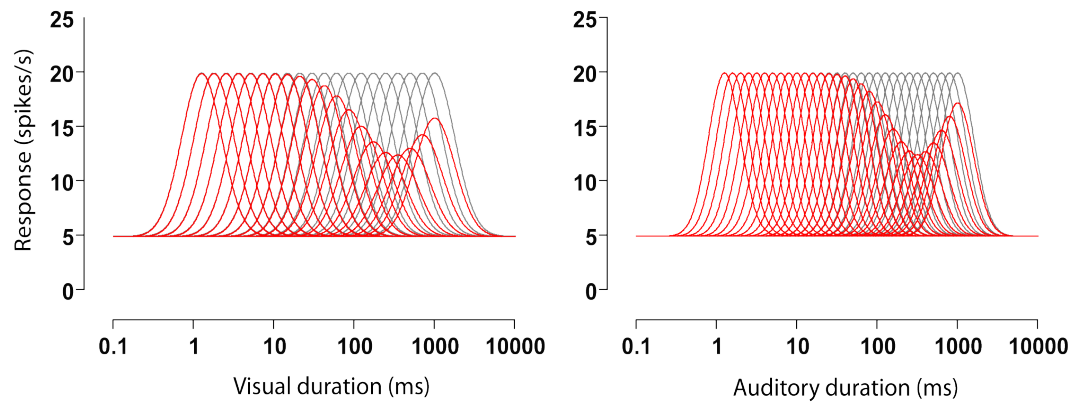


Figure 66 Model tuning curves for (a) visual and (b) auditory durations channels, with (red) and without (black) adaptation to a fixed duration.

Simulations mirrored the trial-by-trial structure of the psychophysical experiment, with a variable test stimulus presented to the adapted modality and a fixed 320 ms reference stimulus presented to the other modality. Neuronal responses were sampled from independent Poisson distributions centred on the value of each tuning curve for a given stimulus. We used a maximum-likelihood decoder (Jazayeri et al., 2006) to generate a binary response on each trial. Figure 67 shows shifts in the PSE produced by the best-fitting model, alongside the corresponding empirical data. Clearly, the model is able to reproduce the repulsive shifts in perceived duration caused by adaptation and provide a reasonable approximation of the tuning of this effect ($R^2 = 0.9$).

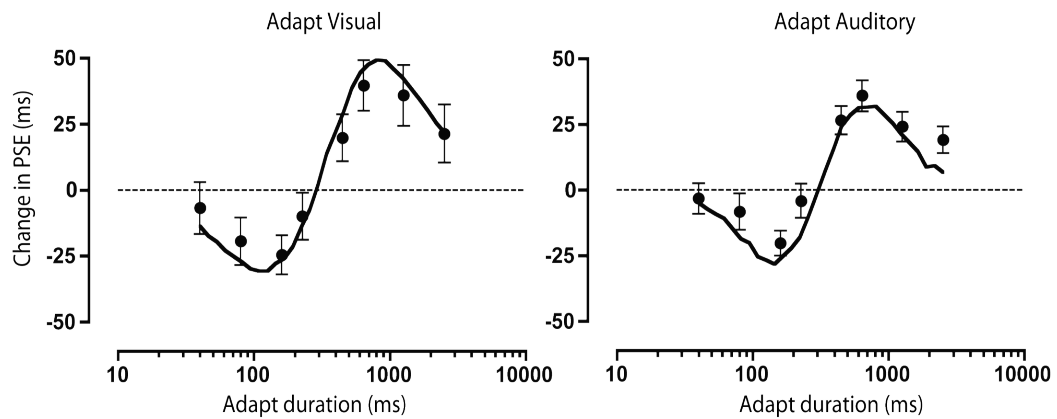


Figure 67 Comparison of experimental and model data. Data points show mean shifts in the PSE as a function of adaptor duration (replotted from Figure 62). Solid lines show the predictions of the best fitting models. (a) visual: $n=20$; $\sigma=0.25$ (log units); $A_{max}=50\%$, $A\sigma = 0.45$ log units). Auditory; $n=30$; $\sigma = 0.2$ log units; $A_{max} = 50\%$, $A\sigma = 0.35$ (log units).

9.5 Duration adaptation or anchoring?

Another possible explanation for these effects could be that subjects were simply ignoring the cross modal reference stimulus and employing a method of single stimulus strategy. In this scenario subjects would base their judgments on all the previous stimuli, including the adaptors and judge the test stimulus against an internally generated mean. However, under these circumstances we should not expect to see the tuning effect produced by adapting to the different durations. In addition, the control experiment described in Chapter 10.1 has subjects reproducing a standard duration following adaptation. This experiment produces comparable effects, in a paradigm where using the method of single stimulus is not possible. Finally it

is worth noting that the only duration perception experiment to use an anchoring paradigm produced cross-modal effects and not the intra-modal effects produced in the experiments described in this chapter (Behar et al., 1961).

9.6 Duration adaptation or temporal frequency adaptation?

9.7 Introduction

Recent evidence suggests that the perception of moderately paced rhythmic auditory patterns can be slowed down or speeded up via prior exposure to relatively fast or slow tone sequences (Becker et al., 2007). In the experiments described previously (Chapter 9.1-9.4), observers adapted to filled durations rather than rhythmic sequences. However, the combination of stimulus duration and an average ISI of 750 ms (jittered between 500 and 1000 ms) provided observers with an average temporal frequency (TF) that varied with the duration of the adapting stimulus. For example, adapting to 160 ms stimuli introduces an average TF of 1.1 Hz, whereas a 640 ms adapting stimulus provides an average TF of 0.72 Hz. In order to ascertain whether TF after-effects contribute to the effects presented thus far, a control experiment was conducted where visual adapting duration was fixed at 160 ms but average ISI was manipulated to provide a TF of 1.1. Hz, (see Figure

68). If our adaptation effects are driven by a TF-based mechanism, we would expect to see equivalence between the 1.1 Hz (160 ms duration stimuli) condition and the 1.1 Hz (640 ms duration stimuli) condition. However, if our effects reflect genuine duration adaptation, the 1.1 Hz (160 ms) should share similarity with the 160 ms data shown in Figures 61 - 65.

9.8 Methods

9.8.1 Subjects

The subjects for this experiment were 4 non-naive experienced observers (DJW, JH, CAS and HKS)

9.8.2 Stimuli

The stimuli were the same as those used in the visual adaptation experiments described previously.

9.8.3 Procedures

The procedures used were the same as those described previously (Figure 60) except that subjects adapted to 160ms visual stimuli whose inter stimulus

interval was lengthened to be jittered between 1135ms and 1635ms, giving an average inter stimulus interval of 1385ms (see Figure 68). These values produced a temporal frequency of 0.72Hz, which was the same as the adapt 640ms condition described earlier. This experiment designed allowed us to establish whether temporal frequency is responsible for, or has any influence on, the results produced in the experiments described earlier.

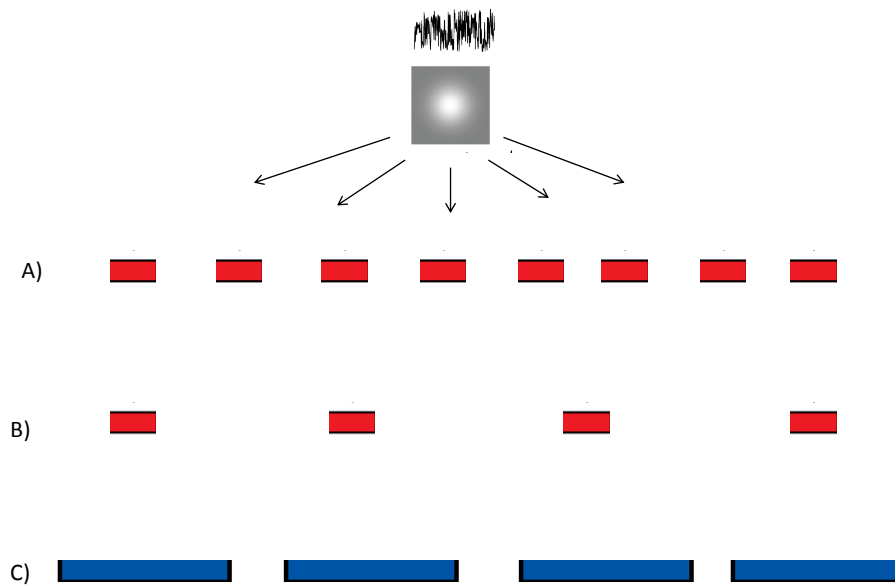


Figure 68 A schematic showing the paradigm used for the experiment. A) Shows the original 160ms adaptation stimuli. B) Denotes the increased ISI which produces a temporal frequency matching that of C) the 640ms adaptation stimuli. (Not to scale).

9.8.4 Results

The results are shown in Figure 69 where adaptation-induced shifts in PSE are plotted – relative to the 320 ms baseline condition – for the two different

TFs and adapting durations. PSE magnitude is plotted relative to the “no adapt” baseline conditions produced in the experiments described in earlier and reproduced in Figure 63a (blue point). Positive and negative values indicate adaptation induced contraction and expansion, respectively. The visual adaptation data from the 160ms (light grey bar) and the 640ms (white bar) is plotted alongside data from the 160ms adaptation condition in which the inter stimulus interval was lengthened to produce a temporal frequency which is the same as the 640ms adaptation condition (0.72Hz). The data show that when the 160ms adaptation condition is changed to produce a temporal frequency which matches that of the 640ms condition, the result is still a perceived temporal expansion and not the temporal contraction produced by the 640ms adaptation condition. Clearly, the closest match in after-effect magnitude and polarity is between the 0.72 Hz (160 ms) and 1.1. Hz (160 ms) conditions. A paired sample t test shows that there is no significant difference between the adapt 160ms conditions ($t(3) = 1.2425$, $p=0.3023$). This finding confirms the underlying importance of event duration, rather than inter-event temporal frequency, in generating the aftereffects presented earlier.

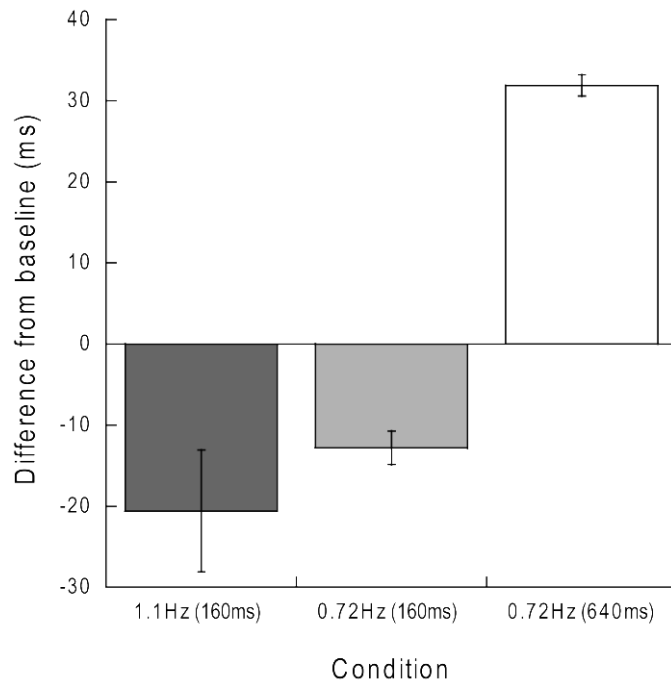


Figure 69 Data from the temporal frequency experiment. PSE values are expressed relative to the "no adapt baseline condition (Figure 60a blue data point) with positive and negative values indicating adaptation-induced contraction and expansion of perceived duration, respectively. 160ms (light grey bar) and 640ms (white bar visual adaptation data from figure 60a is plotted alongside data from a control condition where the ISI separating 160ms adaptation stimuli was lengthened (see Methods for details) to provide a temporal frequency of 0.72Hz (n=4). Error bars represent the standard error of the mean.

9.9 Discussion

In the current study, evidence that human estimates of visual and auditory temporal extent are mediated by a series of bandwidth-limited duration channels is presented. Specifically, adaptation to fixed auditory or visual duration induces sensory-specific distortions of subsequently heard or

viewed durations. The temporal spread of these distortions is limited by the temporal proximity of adaptation and test stimuli, a feature that underscores one of the key similarities between our duration-based effects and the classic literature characterizing CB visual (DeValois et al., 1990) and auditory (Irvine et al., 2005) processing. The fact that our data are well predicted by a generic CB model - without recourse to any novel features specific to temporal perception – emphasizes the similarities between established forms of CB perception and the effects presented in the current study.

9.9.1 Psychophysical context

A significant aspect of the data is the seeming ability of recent experience to selectively initiate both expansion and contraction of perceived duration. This bi-directionality differentiates our effects from other recent duration-based phenomena where sensory history also appears to play a role. For example, perceived duration can be manipulated via prior exposure to dynamic visual stimuli such as flickering patches (Ayhan et al., 2009b, Johnston et al., 2006) or drifting gratings (Burr et al., 2007a, Johnston et al., 2006). Similarly, it has recently been argued that perceived visual duration depends on the extent to which a stimulus is deemed to be repetitive (i.e. its relative novelty) (Pariyadath et al., 2008). In both instances, experimental manipulations induce a unidirectional contraction of perceived duration but, as yet, have not shown reciprocal effects.

The CB framework provides an explanation for earlier reports showing that repeated stimulation (Walker et al., 1981a) or perceptual anchoring (Behar et al., 1961) can influence subsequent duration judgements. In addition, emerging evidence from perceptual learning experiments suggests that training-related increases in duration discrimination sensitivity are tied to durations close to the centre of the trained duration range (Bartolo et al., 2009). Consistent with the data shown in Figures 64 and 65, the magnitude and bandwidth of these learning effects are approximately constant when expressed relative to the trained range (3-4% and 8-11%, respectively, (Bartolo et al., 2009). Similarly one of the defining characteristics of duration judgements is the proportional relationship between duration discrimination threshold and mean estimated duration (Weber's law for duration). Both of these effects show a degree of proportionality consistent with the data shown in Figures 64 and 65, and sit comfortably within a CB framework. Specifically, because channel bandwidth appears to vary in proportion to preferred duration, a system using these channels should show precisely the kind of Weber's law behaviour that is so often observed throughout the duration perception literature (Getty, 1975, Jazayeri et al., 2010, Westheimer, 1999, Miyazaki et al., 2005). Interestingly, the amplitude of the effects show a small but consistent tendency to decline with increases in test duration range (Figures 64 and 65). On first inspection, this effect is perhaps suggestive of smaller levels of response gain reduction at longer test duration ranges. However, it is perhaps more likely to reflect an artefactual feature introduced by increases in the total elapsed time between successive test stimulus presentations: longer test durations are paired with proportionally longer

adaptation stimuli, which have unintended consequence of lengthening test/re-test interval (Figure 60). As such, it is reasonable to speculate that some degree of temporal decay is operating at the longer test duration ranges (Figure 64 and 65). A further control experiment found that adaptation failed to influence perceived duration when our train of adapting stimuli was replaced with a single adapting stimulus (See Chapter 10.11). This finding appears to distance the effects from rapid, attention-dependent adaptation phenomena for which neural loci are thought to reside in extra-striate areas of the cortex (Kovacs et al., 2007, Kovacs et al., 2008, Mueller et al., 2009, Roach et al., 2009).

9.9.2 Neural basis

To model the results, a population coding framework was implemented in which stimulus duration is represented by the pattern of activation across a number of band pass tuned channels. A critical property of this framework is that stimuli of a particular duration stimulate (and therefore adapt) channels in a selective manner, allowing us to account for both the bidirectional (i.e. compressive and expansive) and tuned characteristics of the observed after-effects. While band pass tuning of responses as a function of certain stimulus attributes is relatively common in sensory neurons, realizing this form of selectivity in the time domain poses unique challenges. Consider a collection of channels in which each selectively responds once a particular time interval has elapsed following stimulus onset. Because of the unidirectional flow of time, the presentation of a stimulus will elicit a ‘domino effect’ in which channels respond successively one after another. In principle,

repeated presentation of brief adapting stimuli might selectively adapt channels tuned to brief intervals, providing a basis for explaining expansions of perceived duration. However, as these same channels would also respond to each presentation of a longer adapting stimulus, achieving the selective adaptation required to produce compressions of perceived duration is problematic.

A simple mechanism that avoids this problem is a form of coincidence detection, in which channel activity is driven by simultaneous occurrence of sub-threshold excitatory events linked to stimulus onset and offset (Aubie et al., 2009, Faure et al., 2003). Within this scheme, different duration preferences can be generated by varying the latency of the onset event. Neurophysiological evidence for this type of tuning has been documented in the auditory midbrain (inferior colliculus) (Casseday et al., 1994b, Faure et al., 2003, Brand et al., 2000), primary auditory cortex (Galazyuk et al., 1997, He et al., 1997) primary visual cortex (Duysens et al., 1996) and prefrontal cortex (Yumoto et al., 2011). Neurons within these areas display a phasic burst of spiking activity at stimulus offset, the magnitude of which is tied to the time elapsed since stimulus onset (see Figures 70 and 71). Although the number of these neurons found is few, they do nonetheless show that such a scenario is possible. In addition it should be said that an advantage of a channel based population response model is that it should be capable of representing a wide range of durations with relatively few neurons.

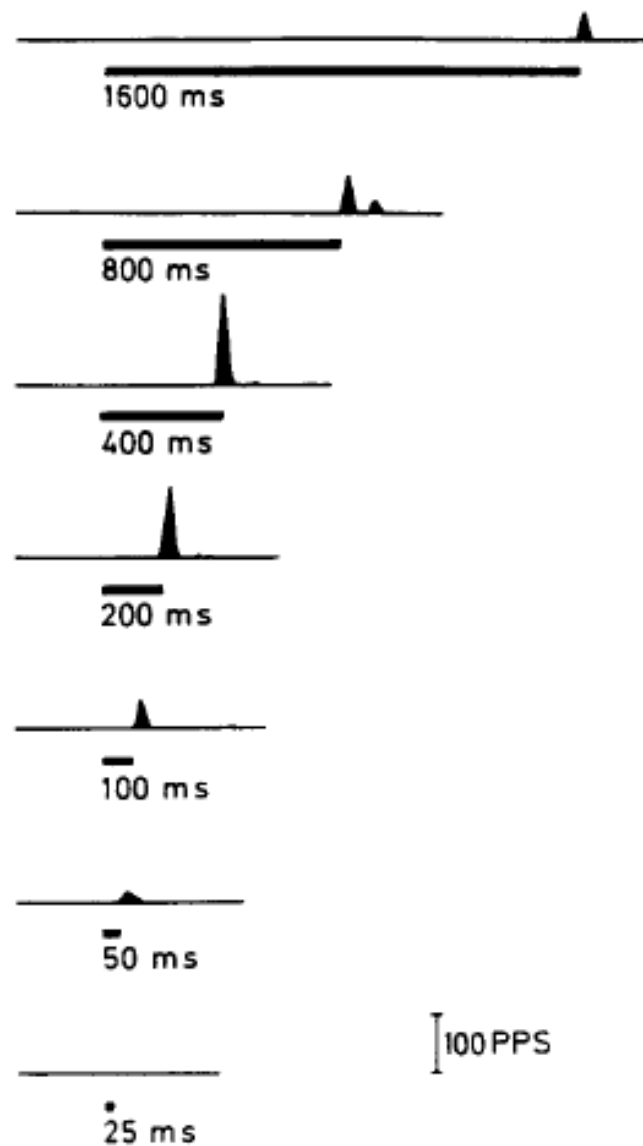


Figure 70 Schematic showing an area 17 (cat) cell with a duration tuning for off responses. This cell shows a maximum response at 400ms (Duysens et al., 1996).

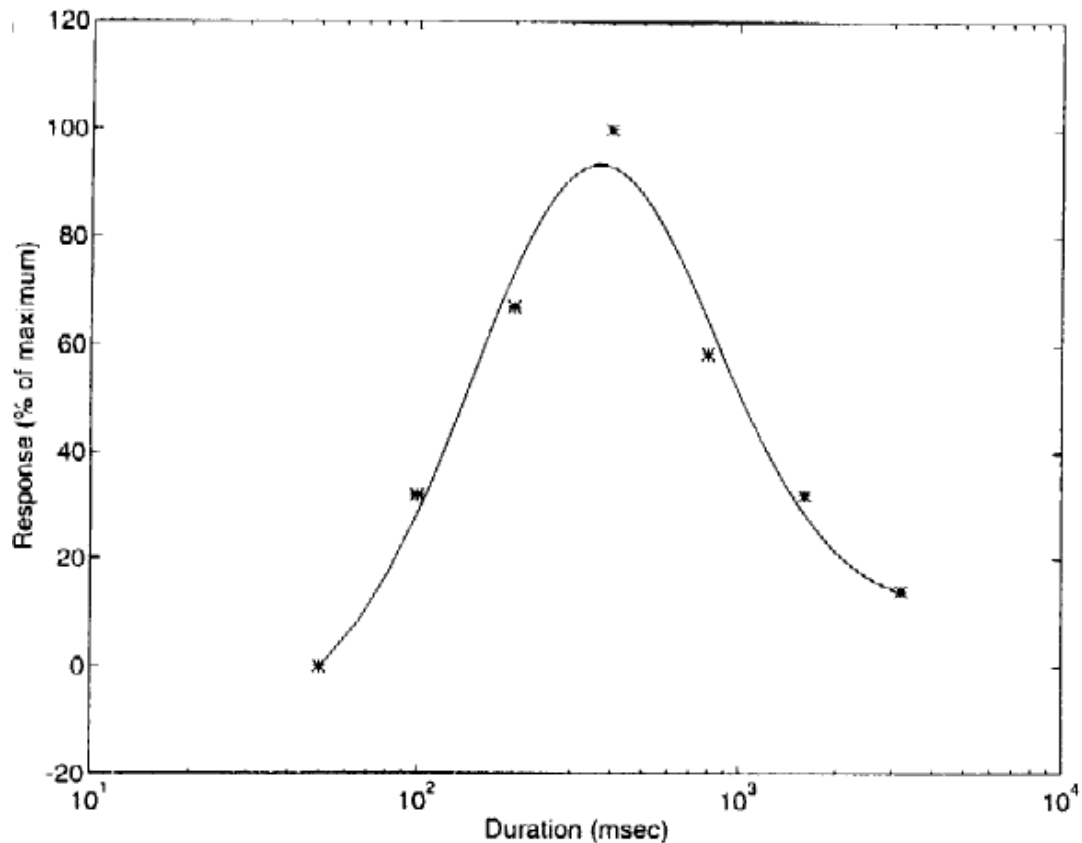


Figure 71 The resulting duration tuning curve for the area 17 cell denoted in Figure 69 (Duysens et al., 1996).

A variety of interval-coding mechanisms have been proposed, not all of which rely on dedicated timing channels (Ivry et al., 2008, Mauk et al., 2004). Recent years have seen the emergence of distributed timing models, referred to as population clocks (see section 3.2.3), which rely on time-dependent changes in the state of neural networks (Karmarkar et al., 2007, Buonomano et al., 2010, Buonomano et al., 1995). This approach offers considerable flexibility, permitting the continuous coding of elapsed time as well as an ability to discriminate between more complex temporal patterns. Because timing is represented in the dynamics of the entire network, it is not immediately obvious how our finding of selective duration after-effects could

be accommodated within this framework. In some implementations of population clock models, different network states are read by output neurons that receive inputs from all the neurons in the network (Buonomano et al., 2010). Feasibly, these output neurons could provide the basis of adaptable duration channels. However, neurophysiological evidence for this process is limited and it remains to be seen whether such a scheme could produce duration selectivity that overcomes the cascading activation problem discussed earlier.

One of the key advantages of a CB system is that the overlapping nature of these channels (Figure 66) negates the need for the system to accommodate a large (potentially infinite) number of channels corresponding to every conceivable duration: by comparing differential activation levels across channels and extracting the population response (Levinson et al., 1976, Mather, 1980), the system can interpolate between neighbouring channels' preferred durations. As outlined earlier, in addition to offering metabolic savings, such a system also affords high-resolution, low-ambiguity estimates of duration. However, this efficiency comes at a cost to the nervous system; sustained activity within individual channels induces repulsive biases in the population response to subsequently presented durations.

9.10 Conclusions

By using sensory adaptation, these experiments have revealed a pattern of temporal perception that is indicative of a perceptual system underpinned by a range of overlapping duration-sensitive channels. The data presented suggests that when formulating estimates of temporal extent, the human nervous system applies some of the same computational principles that are used in the processing of many of the fundamental, yet non-temporal, properties of the world around us.

Chapter 10

Stimulus specificity of duration adaptation

10.1 Is duration adaptation truly intra modal?

10.1.1 Introduction

Chapter 9 demonstrates that adapting to a short visual or auditory stimulus produces a repulsive aftereffect such that an intermediate duration stimulus is perceived as being longer than it actually is. In a similar way, if we adapt to a relatively long stimulus the perceived duration of the intermediate stimulus is shortened. This effect appears to be confined to the modality of the adapting stimulus, is true for a range of sub second intervals and is bandwidth limited. In this chapter the conditions necessary to produce this effect are investigated in a series of follow up experiments.

Studies investigating duration adaptation fail to agree on whether it only occurs intra-modally (Walker et al., 1981a) or whether there is a cross modal transfer of the aftereffect (Behar et al., 1961). Our original experiment could have had two possible outcomes which could have led to four possible conclusions. Firstly, pre and post-adaptation PSE values could have been equivalent. This would lead to the conclusion that either adaptation had no effect on the test or the standard stimuli or that they were both affected to the

same degree. The latter would at least have lent some support to the central timer hypothesis. The experiments described in Chapter 9 show that this is clearly not the case: PSE's were strongly modulated by adaptation. This suggests that either the adapting modality was affected whereas the standard modality was not (i.e. sensory specificity), or that both modalities were affected by adaptation stimuli but to different degrees. In this scenario, partial aftereffect would cause duration discrimination judgments to underestimate the transfer between the senses described in Chapter 9.

On the other hand, if the senses adapt to time independently of one another, it becomes difficult to justify the presence of a single, supramodal, timing mechanism responsible for all senses. In order to investigate this further, two of the conditions from the duration adaptation experiment were repeated but on this occasion a reproduction method was used to measure the adaptation effect during intra-modal and cross-modal conditions. Although reproduction tasks are associated with issues surrounding their criterion-dependent nature (Wearden et al., 2008) by removing the relative nature of the inter sensory comparisons made in Chapter 9, a more absolute measure of perceived time is made available. If adaptation is wholly sensory specific, we should expect duration reproduction paradigms to find similar degrees of adaptation to those found with 2AFC and no adaptation effect for cross modal conditions in which subjects adapt in one sense and reproduce test durations from the opposite modality.

10.2 Methods

10.2.1 Subjects

The subjects for this experiment were 4 non-naive experienced observers (DJW, JH, CAS and HKS)

10.2.2 Stimuli

The stimuli were the same as those used in the experiments described in Chapter 9, Section 3.

10.2.3 Procedures

The procedures used were the same as those described in Chapter 9, Section 3, except that after adapting to the visual or auditory stimulus, subjects were presented with a test stimulus, and were asked to reproduce its perceived duration by pressing a computer key (see Figure 72). There were four adaptation conditions, two cross-modal and two intra modal. These were as follows: (a) adapt vision-reproduce audition, (b) adapt audition-reproduce vision, (c) adapt vision-reproduce vision, and (d) adapt audition-reproduce audition (see Figure 72). For each of these conditions the subjects adapted to durations of 160ms and 640ms (in separate blocks) then

reproduced seven test durations which were logarithmically spaced and ranged from 237ms to 421ms (randomly interleaved within a MOCS).

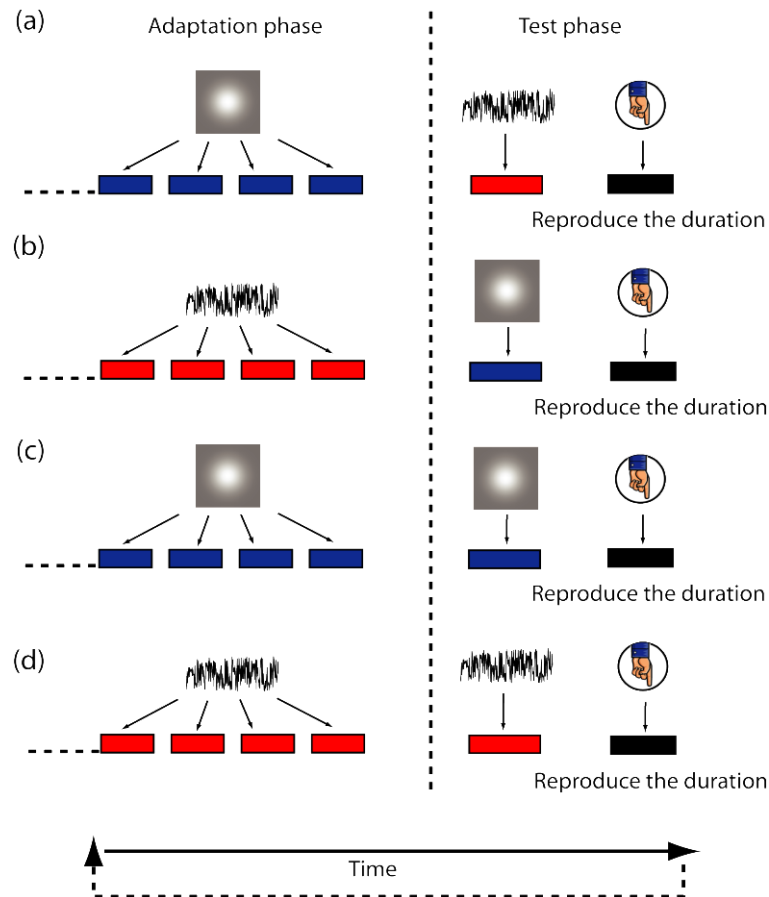


Figure 72 Conditions for the reproduction experiment. a) Adapt visual duration-reproduce auditory duration, b) Adapt auditory duration-reproduce visual duration c) Adapt visual duration-reproduce visual duration and d) Adapt auditory duration- reproduce auditory duration.

10.3 Results

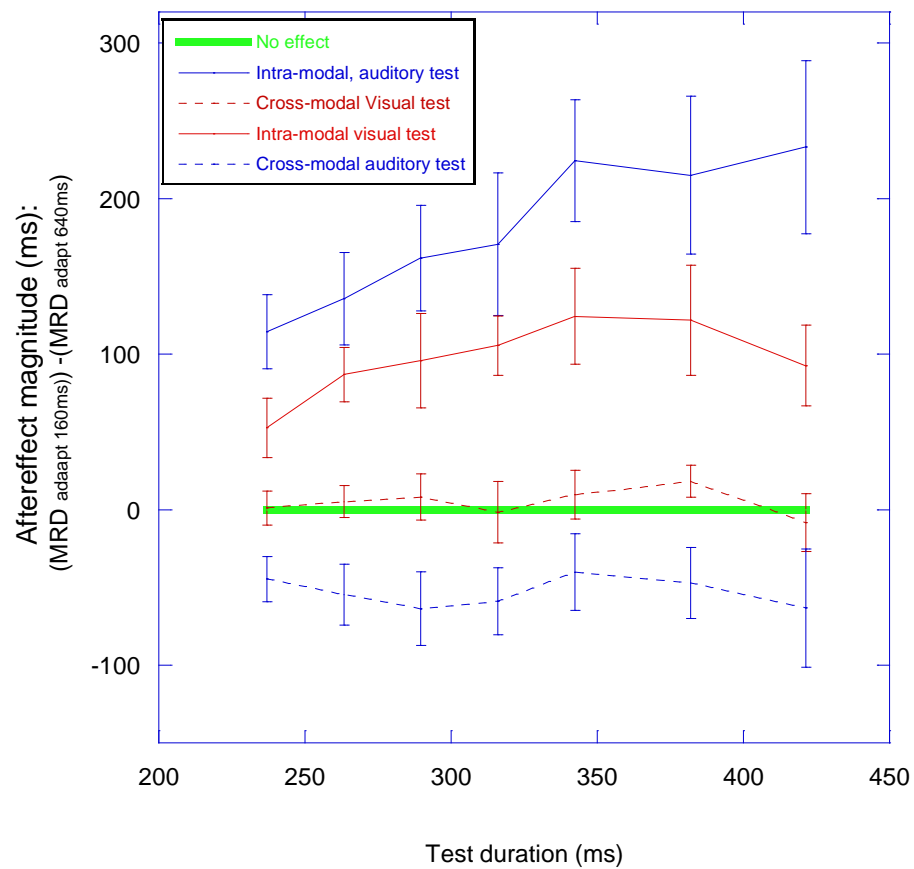


Figure 73 shows the average effect sizes for the four observers. Effect size is defined as the reproduced duration post 160ms adaptation, minus reproduced duration post 640ms adaptation. Effect size is plotted against test duration for cross modal and intra modal conditions. The solid lines represent intra modal adapt and reproduction conditions, broken lines represent cross modal adapt and reproduction conditions. A rebound effect such as that produced in the experiments described in Chapter 9 is shown as a positive effect size in this figure.

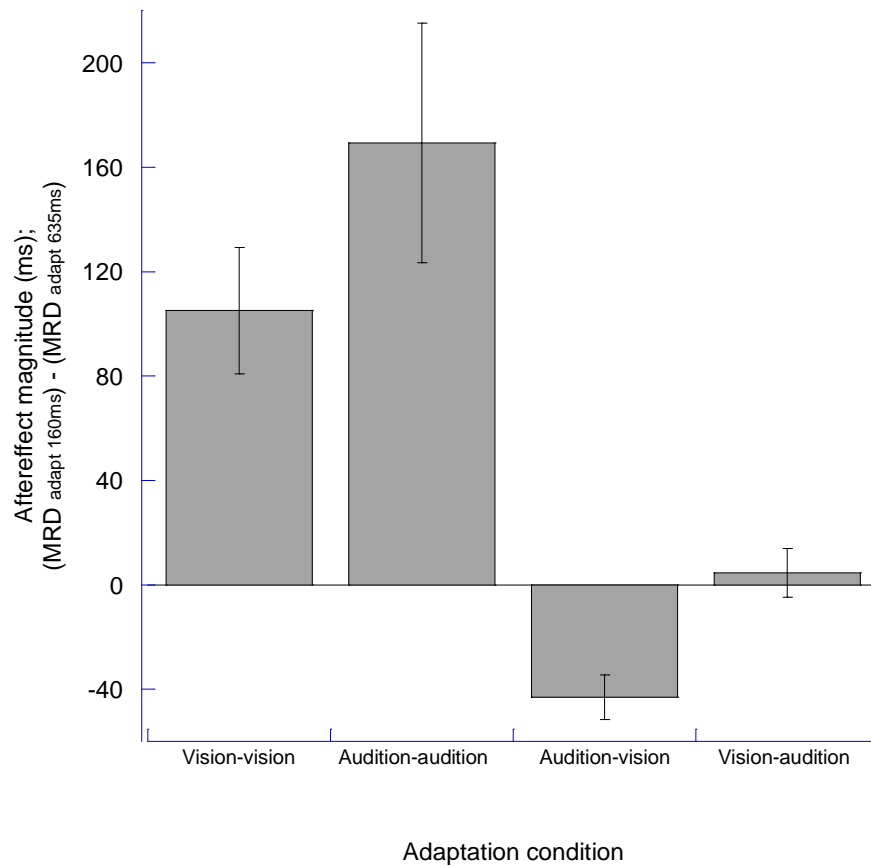


Figure 74 The results of the reproduction experiment expressed in bar graph form. The aftereffect magnitude is calculated by subtracting the MRD (mean reproduced duration) produced when subjects adapted to 640ms from the MRD produced when observers adapted to 160ms. Thus a repulsive aftereffect similar to those produced in the experiments described in Chapter 9 should result in a positive value.

The reproduced durations for each observer were averaged to give a mean reproduced duration (MRD) for each of the four conditions shown in Figure 73. The resulting data is shown in Figure 74.

Aftereffect magnitude is calculated by subtracting the reproduced duration following adaptation to a 640ms duration stimulus from the reproduced duration when adapting to a 160ms stimulus from its counterpart. A rebound effect of the type seen in the experiments described in Chapter 9 will therefore produce positive values. It can be seen that for the intra-modal reproduction (Figure 73- solid lines) conditions a robust adaptation effect is seen for both vision and audition in line with the results of the experiments described in Chapter 9.

For the cross-modal conditions (Figure 73-dashed lines) however, there is no effect for adapt auditory duration - reproduce visual duration and a small effect in the opposite direction for the adapt visual duration - reproduce auditory duration condition. A repeated measures analysis of variance revealed that there was no significant effect of test duration upon aftereffect magnitude for this condition ($F_{6,18}=0.82$, $p>0.1$). However, the aftereffect magnitude was significantly different from zero ($t(27)=6.39$, $p<0.001$). To date a satisfactory explanation for this remains elusive. Figure 74 presents the data from Figure 73 averaged across test duration.

10.4 Discussion

The positive effect with intra modal adaptor and test stimuli is in line with the results obtained using the method of constant stimuli described in Chapter 9. Thus duration aftereffects are produced regardless of which psychophysical method is used. This confirms the robustness of the effect. This contrasts

sharply with the lack of effect found in the cross modal adaptor and test stimuli conditions. The results demonstrate that the adaptation effects found in Chapter 9 occur only when adapt and test stimuli are from the same modality. In addition these adaptation effects do not transfer from vision to audition or vice versa. This confirms the findings of Walker et al (1981a), but is in contrast to those of Behar and Bevan (1961) who found that adaptation effects transferred between modalities. Since our experiments and those of Walker et al used very short (sub second) durations, and those conducted by Behar and Bevan used durations of several seconds, this may be taken as further evidence that different mechanisms are involved for sub second timing than for timing stimuli of several seconds.

10.5 The role of visibility

10.6 Introduction

It has been suggested that stimulus visibility may influence the perceived duration of a visual stimulus (Brigner, 1986, Terao et al., 2008). Therefore it is possible that retinal adaptation to the brightness of the visual adaptor stimuli used in the duration adaptation experiment had an effect on the visibility of subsequently presented test flashes. Specifically, the repetition of the adapting stimuli in retinotopic space could reduce the visibility of subsequent visual test stimuli resulting in a reduction of perceived duration. This could influence all perceived durations but the effect could be expected to be greatest for the long adapting stimuli. Under these circumstances we would expect a differential enhancement of duration contraction, due to the differing visibility of test stimuli, between long and short adapting stimuli, which could conceivably result in an exaggeration of the difference in perceived duration following adaptation to long and short visual stimuli. This experiment sets out to establish whether any reduction in apparent stimulus visibility could have affected the results described in Chapter 9.

10.7 Methods

10.7.1 Subjects

The subjects for this experiment were 3 experienced, non-naive observers (DJW, CAS and HKS).

10.7.2 Procedure

The procedure for this experiment was the same as those used in the main adaptation experiments described in Chapter 9. Subjects adapted to a visual stimulus of 640ms, subsequently they were asked to judge the relative durations of a burst of white noise of 320ms duration and a Gaussian blob of varying durations all of which were centred on 320ms. On this occasion however, for each trial, the adapting visual stimulus consisted of interleaved black and white Gaussian blobs during the adaptation phase. In this way, the visibility of the test stimuli should not be affected by adaptation to the luminance of the visual stimulus presented during the adaptation phase. The 160ms adaptation condition was not used for this experiment because the effect of retinal adaptation on post adaptation stimulus visibility would be expected to be greater with longer duration adapting stimuli.

10.8 Results

Figure 75 shows the PSEs for three subjects after adapting to the single polarity stimulus and after adapting to the alternating polarity stimulus.

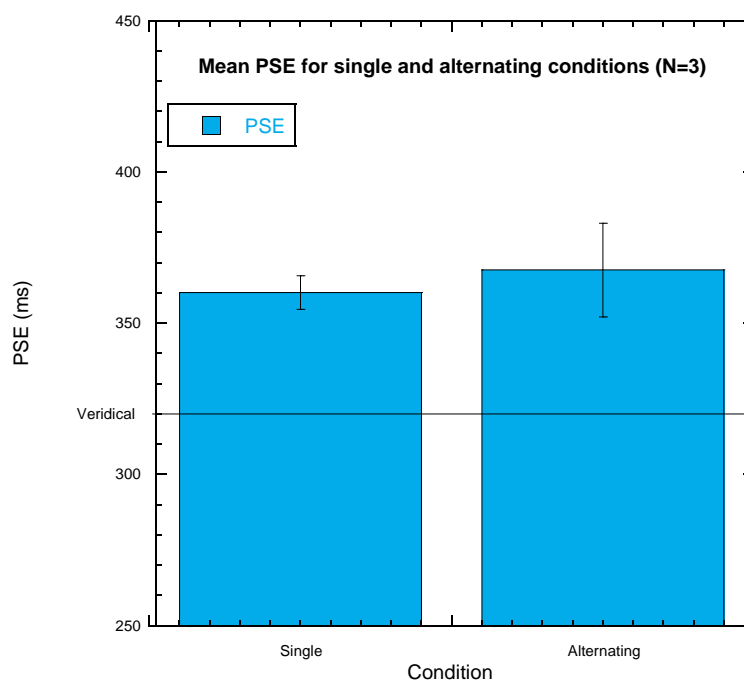


Figure 75 Mean post adaptation PSEs following adaptation to a single (left) and alternating (right) visual stimuli.

Adapting to 640ms stimuli resulted in subsequent visual stimuli being perceived as shorter, therefore in order to be perceived as equal to the 320ms auditory stimulus the visual stimuli need to be, on average around 40ms longer (around 360ms). A repeated measures t test was run on the

data showing that the results obtained with alternate positive and negative stimuli were not significantly different from those with positive stimuli only. ($t(2)=-0.304$, $p=0.79$ (NS)).

10.9 Discussion

The visibility of the visual markers of empty intervals centred at around 100ms has been shown to influence the perceived duration of the interval (Terao et al., 2008). This has also been found for filled intervals of 1000ms (Brigner, 1986). Specifically reduction of the visibility of the visual stimulus produced a contraction of perceived duration. The question this experiment sought to answer was whether any reduction in visibility resulting from adapting to the longer visual stimuli in the experiment described in Chapter 9.3, could have increased the magnitude of perceived duration contraction found. If this were the case, then the alternate reversing of polarity in this experiment would be expected to reduce or neutralise any reduction in visibility and therefore reduce duration contraction relative to the single white stimulus condition. However the analysis of the data shows that if the stimulus visibility is reduced, the effects are too small to have any significant impact on the data presented in Chapter 9.

10.10 Adaptation or aftereffect?

10.11 Introduction

Although the preceding experiments suggest a low-level adaptation mechanism an alternative interpretation of these results is that they are mediated by high level mechanisms akin to those responsible for shape after effects. (Suzuki, 2001). For example when images of squares briefly presented are preceded by concave or convex shapes presented for 134ms or less their perceived shape is distorted. An example of this is shown in Figure 76.

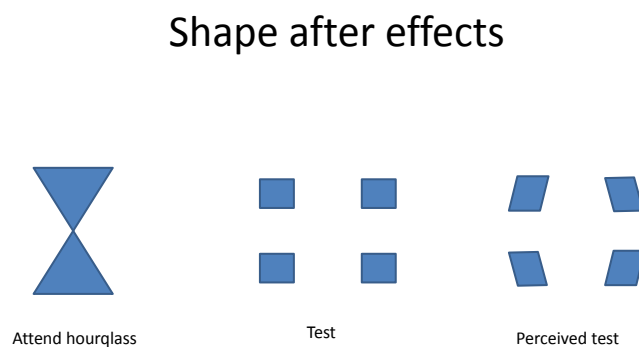


Figure 76 an example of a shape after effect. If after brief attention to the hour glass we present square shapes they are perceived to be distorted in shape as seen on the far right of the schematic.

These effects are thought to occur at a higher level than V1 and do not resemble spatial frequency aftereffects which require an extended period of adaptation. In order to examine this issue further we repeated part of the main experiment outlined in Chapter 9.3, for two adaptor durations, one “short” (122ms) and one long (518ms). This time however, just one adaptor duration was presented before the standard and test stimuli.

10.12 Methods

10.13 Subjects

The subjects for this experiment were four non-naive experienced observers (DJW, JH. CAS and HKS) and 1 naive observer (BB)

10.14 Stimuli

The stimuli were the same as those used in the experiments described in Chapter 9.

10.15 Procedures

The procedures for this experiment were exactly the same as those used in Chapter 9, except that a single duration was presented before each trial instead of one hundred initial adapts and four top ups. The experimental design is shown in Figure 77.

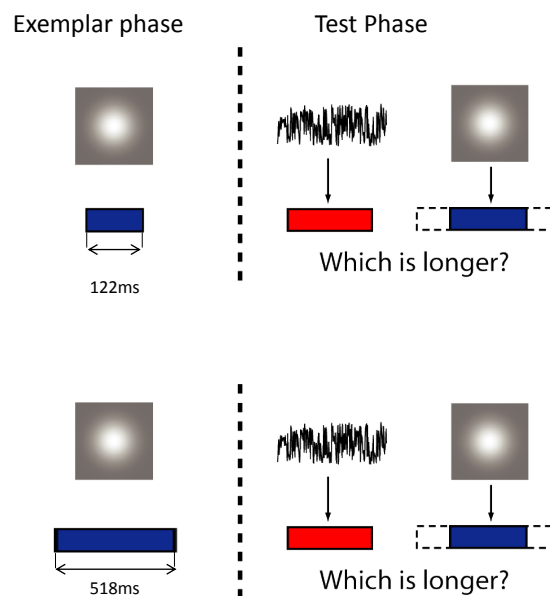


Figure 77 Modified method for the experiment, showing a single adapting duration instead of the one hundred adaptation durations and four top ups used in the experiments described in Chapter 9.

10.16 Results

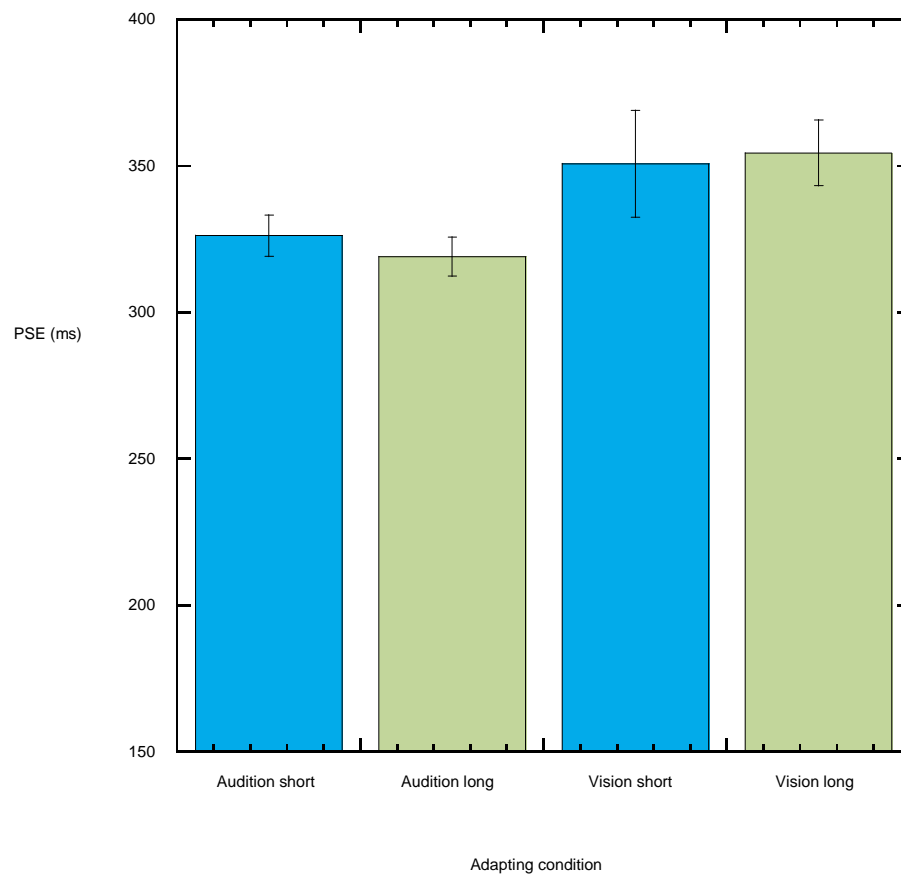


Figure 78 The point of subjective equality between the standard 320ms stimulus duration and test stimulus durations from the opposite modality for all subjects are shown plotted against the different adaptation conditions. Since the bars are all around 320ms the graph shows no significant effect. The error bars denote standard error.

Psychometric functions comprising observers' duration discrimination, judgments were plotted showing the proportion of 'test longer than reference'

responses as a function of test duration (see section 7.2 for details). PSE values (between reference and test durations) were extracted and mean values for all observers ($n=5$) calculated. These values are shown in bar graph form (Figure 78).

By comparing the PSEs shown in Figure 78 it can be seen that there is no significant effect of the single duration ($t(4) = 1.1158$, $p=0.327$ audio, $t(4) = -0.35266$, $p=0.7421$ visual). A two-factor repeated measures ANOVA shows that the difference between short and long is not significant ($F(1,4)=2.404$, $p=0.841$). Although we can see a small difference between the PSEs for visual and audio test stimuli in line with studies which have found that auditory stimuli tend to be perceived as longer than visual stimuli, the ANOVA shows this fails to reach significance ($F(1,4)=5.499$, $p=0.079$). There is no interaction between Audition/Vision and Short/Long, ($F(1,4)=2.404$, $p=0.196$).

10.17 Discussion

The results appear to rule out a higher level response as described by Suzuki (2001) and show the importance of prolonged adaptation rather than brief exposure to a stimulus. If prolonged adaptation was not the crucial factor in the duration effects, then a similar distortion of perceived duration to that seen in the earlier experiments would have occurred. The results of this experiment appear to confirm that the adaptation effects described in

Chapter 9 are of a similar nature to adaptation aftereffects found with spatial frequency.

10.18 Adaptation and spatial frequency

10.19 Introduction

According to the intrinsic timing models of duration, our perception of short periods of time is processed using information provided by the neural networks activated by the stimulus being timed. As discussed at length in Chapter 3.2.3, this could be effected via “energy readout” mechanisms in which stimulus duration is encoded in the level of neural activity or by interpreting the changes in these neural networks in which different durations are represented by particular time dependent spatial patterns of network activity (Karmarkar et al., 2007, Ivry et al., 2008). As the neurons involved in representing a stimulus are also said to process short periods of time, these models predict a degree of specificity for time perception.

A well-established theory of visual processing states that different spatial frequencies are processed by different “channels” – neurons which are maximally sensitive to particular spatial frequencies (Blakemore et al., 1969(a), Blakemore et al., 1970). If different spatial frequencies are processed by distinct groups of cortical neurons, might these neurons also be responsible for the intrinsic timing? If so, would it be possible to adapt the duration of a particular spatial frequency, whilst the perceived duration of a distant spatial frequency remains veridical? Put another way, is the adaptation effect confined to the activated spatial channel or does it transfer

across spatial channels? We investigate this possibility by manipulating stimulus characteristics between adapt and test phases.

10.20 Methods

10.20.1 Subjects

Five observers took part in the experiment, 3 experienced (CAS, DJW, and HKS) and 2 observers naive to the purpose of the experiment (AK, and KM).

10.20.2 Stimuli

Visual stimuli comprised Gabor patches presented on a Sony Trinitron GDM-FW900 monitor (See Chapter 7.3 for details). The Gabor patches were composed of a sinusoidal grating carrier (2cpd or 8cpd) multiplied by a spatial Gaussian envelope. The background was a mean luminance grey. The Gaussian envelope had a standard deviation of 2.7 degrees of visual angle for both carrier frequencies in order that both Gabors should be of the same size. The onset and offset profiles of the Gabors and the bursts of white noise were rectangular. Auditory stimuli comprised white noise bursts presented over headphones (see Chapter 7.3 for details)

10.21 Procedures

10.21.1 Adaptation phase

Subjects began a block of trials by entering the adaptation duration for that block. Blocks were allocated in a pseudo-random fashion. For each block, subjects were exposed to a series of 100 presentations of a 2cpd Gabor patches. The duration of these patches was either 160ms or 640ms duration which was fixed within a block. The inter-stimulus interval between adapting stimuli varied randomly between 0.50 and 1.0 seconds. This was followed by a short pause of 2 seconds which alerted the subject that the trial was about to begin.

10.21.2 Test phase

Each trial was preceded by 4 top up adaptation stimuli in an attempt to maintain adaptation levels between trials. On test trials the subjects were asked to make a 2 alternative forced choice duration discrimination judgment between a 320 ms burst of white noise whose duration was jittered in the same fashion as described in Section 8.8, and the test Gabors which were randomly presented for one of seven durations. These were proportional to the reference stimulus in log steps (0.775, 0.85, 0.925, 1, 1.075, 1.15, or 1.25). Subjects indicated which was the longer of the two stimuli (the auditory reference or visual test) by pressing the appropriate arrow key on the

computer keyboard. This response initiated the next top up/test cycle until each of the test durations had been presented 10 times. The adaptor durations were alternated between blocks. Subjects rested for a minimum of 10 minutes between blocks of trials in order that any adaptation from one block should not be carried over to the subsequent block. The test Gabor spatial frequencies were either 2 cpd or 8 cpd for each block. Subjects completed four blocks for each of the adaptor durations and spatial frequencies mentioned in the adaptation and test phases. The presentation order of each adapt-test pairing was randomised.

10.22 Results

Psychometric functions were plotted showing the percentage of test longer than reference responses in relation to test duration. The functions were fitted with a logistic of the form

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

(See Chapter 7.3), where μ is the test duration corresponding to the point of subject equality ($p=50\%$) between the “reference” stimulus and “test” stimulus extracted for each subject and for each adaptor duration and θ is an estimate of duration discrimination threshold (approximately half the offset between the 27% and 73% response levels. PSE values were extracted for each subject and each adapt-test configuration.

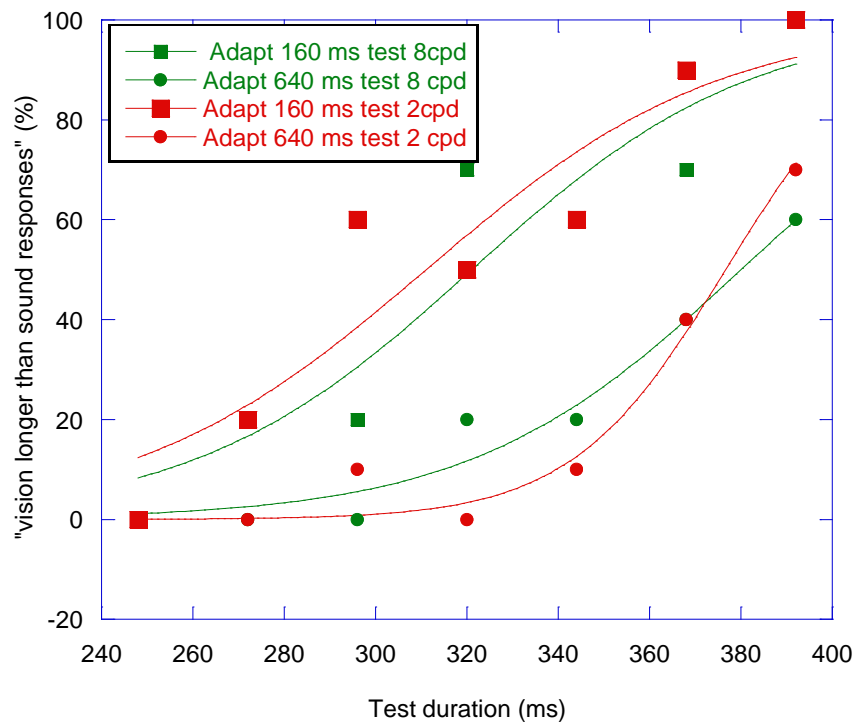


Figure 79 Sample data from subject DJW. The Y axis denotes the percentage of vision longer than sound responses. The X axis denotes the average duration of the test stimulus . The adapt 160ms conditions are denoted filled squares. The adapt 640ms conditions are denoted with filled circles. Test with 2cpd conditions are shown in red, test with 8cpd are shown in green. For all conditions the adapting stimulus had a spatial frequency of 2cpd.

Figure 79 shows sample data from representative subject DJW. The Y axis denotes the percentage of vision longer than sound responses. The X axis denotes the average duration of the test Gabors.

Figure 79 clearly shows that repulsive aftereffects in perceived duration are not specific to the Gaussian blob stimuli employed in earlier chapters. Specifically, adapting to 160 ms and 640 ms duration stimuli caused

subsequent test durations to be expanded or contracted respectively. This is shown by the horizontal separation between the red and green functions representing the 160 ms and 640 ms adapting functions. Of critical importance to the current study is the finding that these effects show no dependence on spatial frequency. For example, adapting to relatively long 2cpd stimulus induces a compression in the duration of both 2cpd and 8cpd stimuli (red and green circles in Figure 79). This is illustrated by the similarity in the horizontal separation between the red functions and their green counterparts (Figure 79).

It is possible that adapting to the 2cpd Gabor could alter the perceived spatial frequency and/or visibility of the test stimuli which in turn could affect perceived duration. Therefore we calculated duration aftereffect magnitude (DAM); the difference in PSEs between conditions where any confounding effects were common to both conditions and the only difference was test spatial frequency (see Figure 80).

$$\text{DAM}_{\text{same SF}} = (\text{PSE}_{\text{adapt 2cpd, 640ms test 2cpd 320ms}}) - (\text{PSE}_{\text{adapt 2cpd, 160ms test 2cpd, 320ms}})$$

$$376.640 \text{ ms} - 311.07 \text{ ms} = 65.568 \text{ ms}$$

$$\text{DAM}_{\text{different SF}} = (\text{PSE}_{\text{adapt 2cpd, 640ms test 8cpd 320ms}}) - (\text{PSE}_{\text{adapt 2cpd, 160ms test 8cpd, 320 ms}})$$

$$380.032 \text{ ms} - 320.96 \text{ ms} = 59.072 \text{ ms}$$

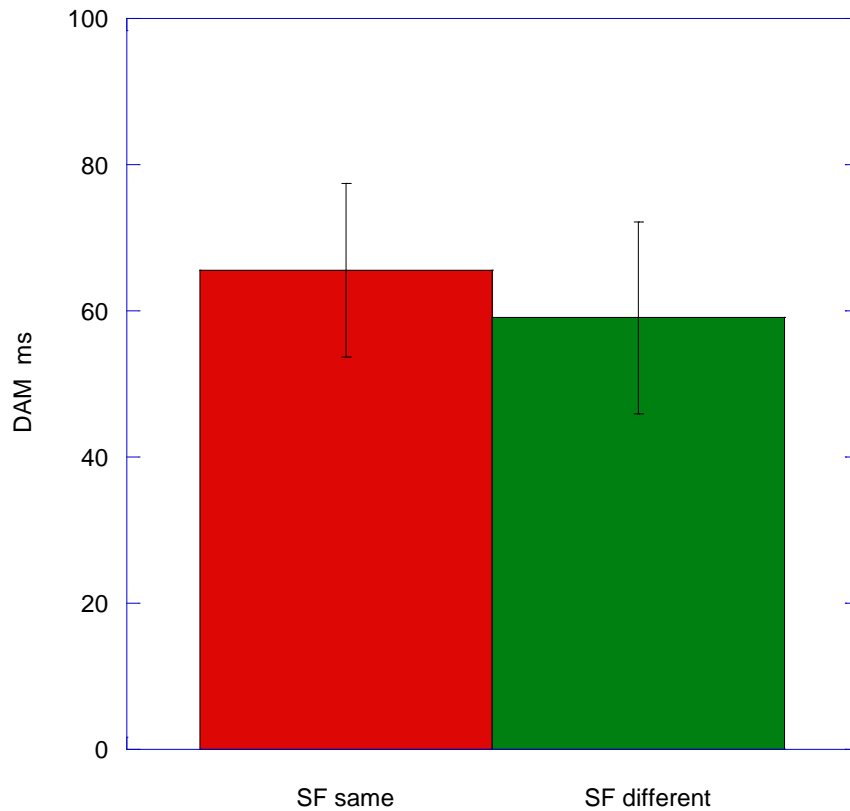


Figure 80 Data for the same subject shown as a bar graph. The y axis denotes the Duration aftereffect magnitude between adapt 640ms and adapt 160ms. The x axis denotes the adapting conditions where spatial frequency was held to a constant 2 cpd across adapt and test phases (red bar) or was increased from 2cpd (adapt) to 8cpd (test) (green bar).

Mean data taken from all 5 observers shows a similar effect for the different spatial frequencies and adapting conditions to that found for subject DJW (see Figure 81 and calculations).

$$\text{DAM}_{\text{same SF}} = (\text{PSE}_{\text{adapt 2cpd, 640ms test 2cpd 320ms}}) - (\text{PSE}_{\text{adapt 2cpd, 160ms test 2cpd, 320ms}})$$

$$357.824\text{ms} - 309.334\text{ms} = 48.490\text{ms}$$

$$\text{DAM}_{\text{different SF}} = (\text{PSE}_{\text{adapt 2cpd, 640ms test 8cpd 320ms}}) - (\text{PSE}_{\text{adapt 2cpd, 160ms test 8cpd, 320 ms}})$$

$$370.784\text{ms} - 332.928\text{ms} = 37.856\text{ms}$$

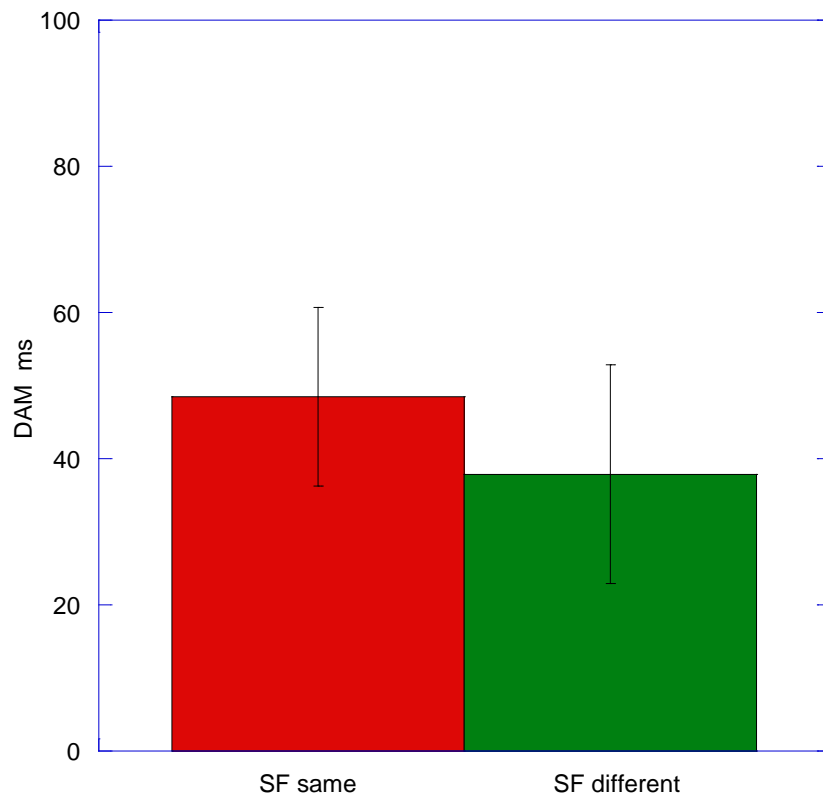


Figure 81 Mean DAM data for all 5 subjects shown as a bar graph. The y axis denotes the DAM between adapt 640ms and adapt 160ms as a proportion of the reference stimulus. The x axis denotes the adapting conditions where spatial frequency was held to a constant 2 cpd across adapt and test phases (red bar) or was increased from 2cpd (adapt) to 8cpd (test) (green bar). Error bars represent standard error.

The fact that the calculations produce a positive differential tells us that duration adaptation is not specific to Gaussian blobs (Section 9) since it also occurs with Gabors. A repeated measures t test was run on the data showing that the results obtained with 2cpd test stimuli were not significantly from those with 8cpd test stimuli. ($t(4) = -.4164$, $p=0.6985$ (NS)).

10.23 Discussion

The current study tested the hypothesis that the neurons responsible for the channel based processing of different spatial frequencies, may also intrinsically process the duration of a visual stimulus. If the perception of spatial frequency and duration is achieved using the same neurons, we should expect maximal adaptation effects when adapting and testing stimuli have matched spatial frequencies but differing durations. Clearly, the results of this experiment do not support the hypothesis. In conclusion, this experiment appears to demonstrate that visual duration is not processed using the neurons responsible for channel based spatial frequency perception. It appears that although spatial frequency may be a factor in the perceived duration of a briefly presented Gabor (Chapter 8), duration adaptation is not sensitive to changes in spatial frequency between adapt-test phases. This would seem to suggest that spatial frequency sensitive neurons are not directly involved in producing duration adaptation effects and that said effects occur earlier or later than V1. In other words, duration

adaptation effects occur at a processing stage before or after the extraction of spatial frequency information.

10.24 Is temporal adaptation Retinotopic or Spatiotopic?

10.25 Introduction

For many years the dominant theory for the human perception of time has been that of a centralised supra modal timing mechanism (Creelman, 1962, Treisman, 1963).

More recently, the notion of a central supra-modal clock type mechanism has been called into question by a number of studies which suggest that the timing of short (sub second) durations may be achieved more locally. In the case of the visual system, it has been found that adapting to a 20Hz drifting grating produces a contraction of the perceived duration of a subsequently presented 10 Hz drifting grating stimulus in the same spatial location as the adapting stimulus and nowhere else (Johnston et al., 2006). This effect does not occur if the test stimulus is displaced by as little as 1 degree away from the adapting stimulus (Ayhan et al., 2009b). This degree of specificity led the authors to suggest an early part of the visual system as the most likely site of the adaptation effect.

However, some related studies of this effect propose that a later stage of the visual system is responsible (Burr et al., 2007a). This view is based upon the findings from experiments investigating whether adaptation occurs retinotopically or spatiotopically (Burr et al., 2007a). In these experiments the adapting stimulus comprised of a 1cpd horizontal test grating which drifted vertically at 20 Hz as the adaptation stimulus. After adaptation, subjects

made a saccade which enabled a test grating (drifting at 10 Hz) to be presented at a number of spatiotopic and retinotopic positions, or a control position. The test gratings were compared with subsequently presented “probe” gratings of variable duration (see Figure 82).

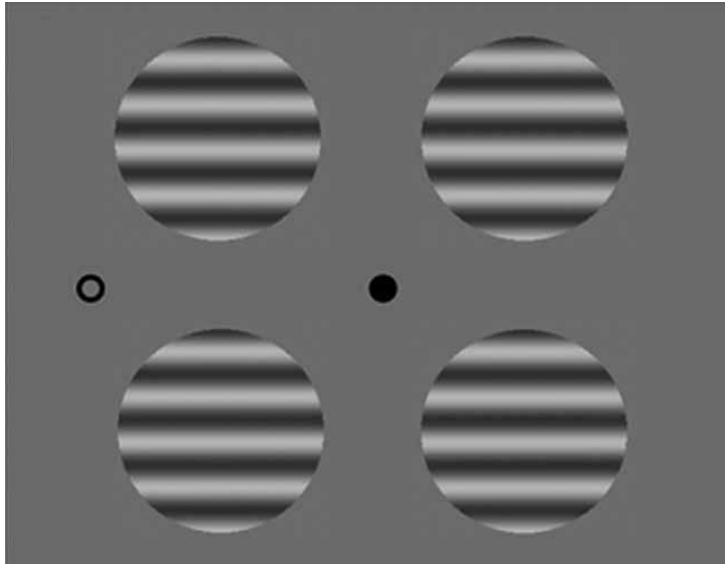


Figure 82 Stimuli used in the Burr et al (2007) study. Subjects fixated on the open circle on the left of the display whilst adapting to the top left hand grating, drifting at 20Hz (changing direction every 2 seconds). The fixation point then changed to the filled circle in the centre of the display and subjects made a saccade to this point. Subsequently a grating drifting at 10Hz appeared for 600 ms at one of three positions, the top left position (spatiotopic), the top right position (retinotopic) or the bottom left position (control). These stimuli were interleaved randomly. After 500ms the probe grating appeared in the bottom right hand position for a variable duration. Subjects were required to make a 2AFC judgement as to which grating appeared for the longest duration the “test” stimulus or the “probe” stimulus (Burr et al., 2007a).

The authors matched the perceived speed of the test stimuli in spatiotopic and retinotopic locations because it has been shown that perceived speed influences apparent duration (Kanai et al., 2006). They found that when the

perceived speeds were matched, time compression was found only spatiotopically.

Since the early part of the visual system encodes visual information retinotopically, the authors concluded that this adaptation effect could take place in one of the higher order areas of the parietal cortex which are thought to encode space in spatiotopic terms (Burr et al., 2007a).

Another study (Bruno et al., 2010) investigating the retinotopic/spatiotopic question reached the opposite conclusion and found retinotopic but little or no spatiotopic adaptation. The authors say that the difference in findings may be partly because the earlier study (Burr et al., 2007a) presented the test retinotopic and spatiotopic stimuli before the probe stimulus in all trials whereas the later study (Bruno et al., 2010) randomised their order. Because the second of a pair of intervals is usually thought to be perceived as being greater than the first (Jamieson et al., 1975, Nachmias, 2006), it is suggested this order effect could interact with the adapting conditions to produce a greater effect post adaptation. The authors say that when the test stimuli were presented in random order a retinotopic effect was found with little or no spatiotopic adaptation.

The effects found have similarities with those described in Section 9.4 because in both cases a period of adaptation produces a subsequent distortion of perceived duration. A crucial difference between them however, is the nature of the adapting stimulus. The present experiment sets out to determine whether these effects are related. If so we should expect to find some spatial specificity to the effects of duration adaptation.

10.26 Methods

10.26.1 Apparatus and stimuli

Visual stimuli were produced on a 22 inch, Sony Trinitron GDM-FW900 monitor. The auditory stimulus was a burst of white noise produced using Seinnheiser HD280 headphones. The visual stimuli consisted of a Gaussian blob of white light, of the maximum luminance available for the monitor (See Section 7.3 for details).

10.26.2 Subjects

There were 8 subjects for this experiment, 4 experienced (CAS, DJW, HKS, and JH) and 4 who were naive (AH, FA, KM and NC).

10.26.3 Procedure

Adaptation was produced using a paradigm very similar to that described in section 9.3. Subjects were presented with a fixation cross 10 degrees to the left of the screen centre (Figure 82). They were instructed to fixate on the cross throughout the experiment. A Gaussian blob was then presented at the centre of the screen for 160 ms in half of the blocks or 640 ms for the other

half. The stimulus was presented 100 times, after which there was a 2 second pause which alerted the subject that the test phase of the experiment was about to begin. This was followed by 4 “top up” presentations of the adapting stimulus. The fixations cross then moved to a position 10 degrees to the right of the screen centre and subjects shifted fixation along with it. After a variable pause of between 500 ms and 1000 ms, a burst of white noise was presented for 320 ms followed by the test stimulus at either the centre of the screen (the same spatiotopic position occupied by the adapting stimulus) or 10 degrees to the right of fixation (the same retinotopic position). The visual stimuli were presented for a range of 7 durations centred around 320 ms from 237 ms to 471 ms. Subjects were asked to indicate whether the auditory or visual stimulus appeared longer by means of the computer keyboard. Each of the visual test durations was presented 5 times in each position (randomly interleaved) within a block. Blocks were repeated 6 times, giving a total of 30 presentations per visual location for each subject. The procedure for the baseline “no adapt” condition was the same except that when subjects fixated the first fixation cross there were no adapting stimuli or top ups. They were however required to make a saccade to the second fixation cross before making the duration judgement.

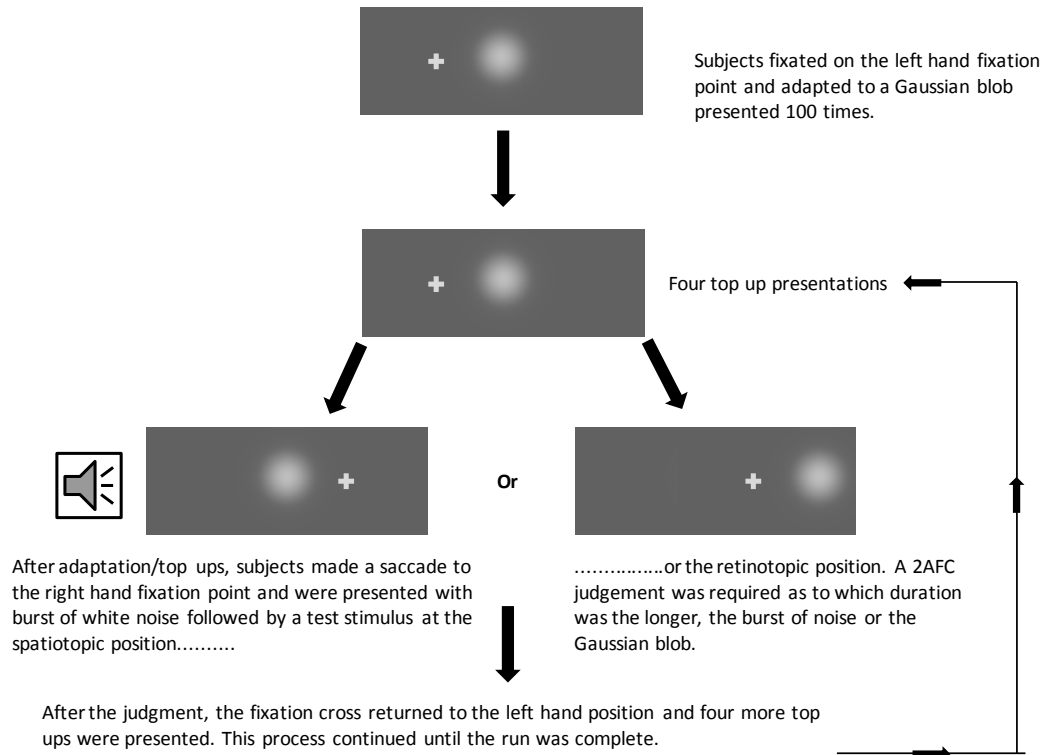


Figure 83 Experiment schematic (see text for details).

10.27 Results

The resulting data was used to produce psychometric functions for each subject in which the number of test longer responses than reference responses was plotted against the test duration values. A logistic function

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

was fitted to the resulting data and the point of subject equality (PSE) ($p=50\%$) between the “standard” stimulus and “test” stimulus extracted for each subject, adaptor duration and spatial location.

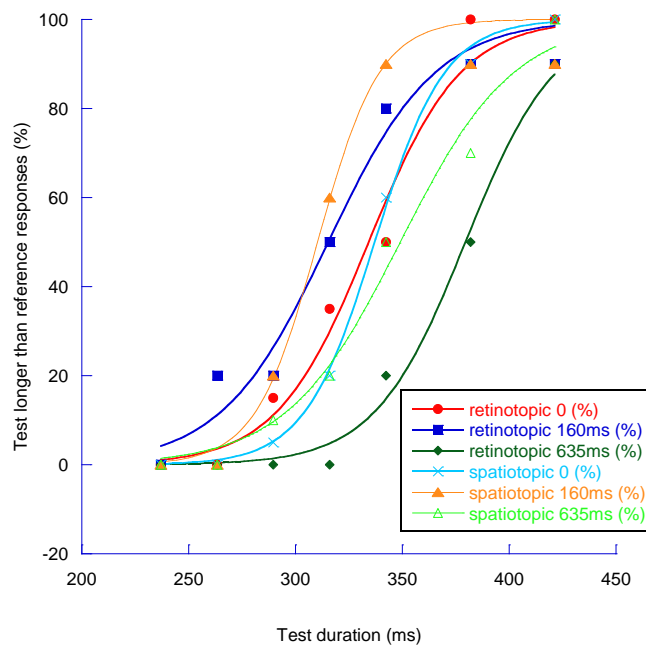


Figure 84 The psychometric functions for subject DJW. The percentages of test longer responses are plotted along the Y axis and the test durations along the X axis. The figure represents 6 conditions – 3 for spatiotopic test stimuli and 3 for retinotopic stimuli. The adaptation conditions for each are no adapt, adapt 160ms and adapt 640 ms. The PSE values for this subject are as follows:-

Retinotopic no adapt 334.36 ms adapt 160 315.38 ms adapt 640ms 371.6ms.

Spatiotopic no adapt 337.16 ms, adapt 160ms 310.41ms, adapt 640ms 348.99ms.

A representative set of psychometric functions for subject DJW is shown in Figure 84. The effect of adaptation for the different conditions may be seen more clearly if we subtract the PSE values for the no adapt conditions from the 160 ms and 640 ms adapt conditions and plot the resulting figures as a bar graph. This is shown for subject DJW in Figure 85.

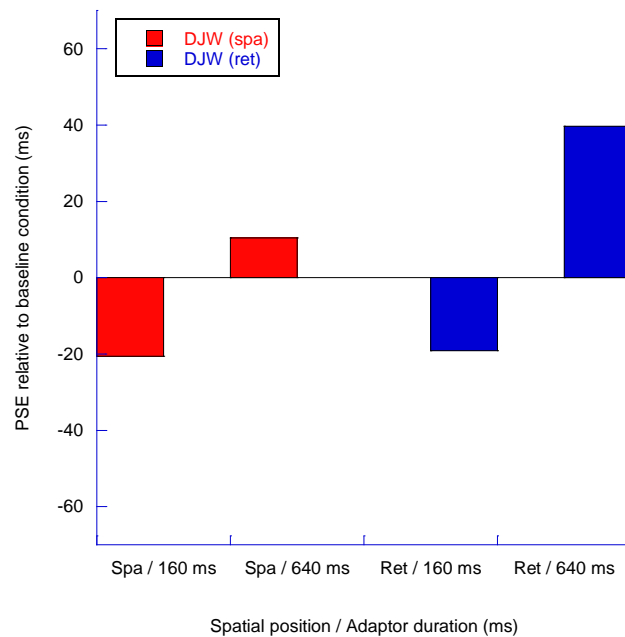


Figure 85 PSE's for subject DJW for the 160 ms and 640 ms adapt conditions minus the PSEs for the no adapt condition. The effects of the adapting stimuli are now shown relative to the no adapt condition. PSE's are plotted on the Y axis and the different adapting conditions on the X axis (Spa = spatiotopic, Ret = retinotopic)

Both spatiotopic and retinotopic conditions produce a distortion of the perceived duration of the test stimuli. When the subject adapts the 160 ms duration, the perceived duration of the test stimuli at the spatiotopic position is greater (negative PSE shifts reflect perceptual expansion of duration). This is a typical rebound effect of the type found in Chapter 9.4. It can be seen that a test stimulus of 20.58 ms less than no adapt PSE is perceived as being equal to the 320 ms standard burst of white noise. Conversely, the 640 ms adapt condition produces a contraction in the perception of the test stimulus. In this case the test stimulus needs to be 10.45 ms longer, relative to the no

adapt condition) before it is perceived as being equal to the standard stimulus.

Turning to the retinotopic condition, a similar set of results are seen insofar as the adapting conditions produce the same type of contraction and expansion of the perception of the test stimuli. However, the retinotopic test stimuli are affected to greater degree. This is particularly marked when adapting to 640 ms which shows over 3 times the effect on PSE of the spatiotopic condition (39.73 ms greater than the no adapt condition). The adapt 160 ms retinotopic condition produces a PSE of -19.13 ms relative to the no adapt condition.

The data was obtained for all 8 subjects and mean values produced. These are shown in Figure 86. It may be seen that the effects produced are similar to those of individual subject DJW. We calculated the overall aftereffect magnitude by subtracting the PSEs for the adapt 160ms conditions from the adapt 640ms conditions. The overall aftereffect magnitude was significant both for the spatiotopic condition ($t(7)=4.807$, $p<0.005$) and particularly for the retinotopic condition ($t(7)=5.852$, $p<0.001$). Nevertheless, despite the trend towards a larger effect size under retinotopic conditions, a paired t-test revealed that this difference did not quite reach statistical significance ($t(7)=-2.22$, $p=0.0618$).

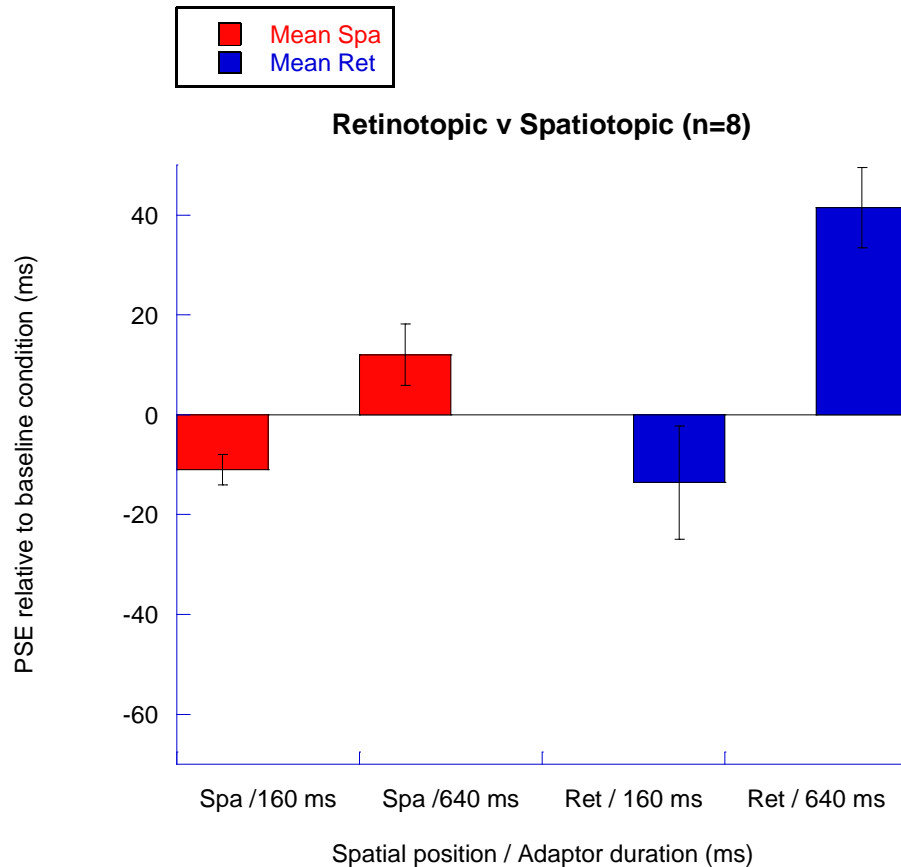


Figure 86 Mean data for all 8 subjects. Showing PSE's of the different adapting conditions relative to the no adapt conditions. PSE's are plotted on the Y axis and the different adapting conditions on the X axis. Error bars denote standard error.

10.28 Discussion

This experiment sought to clarify the nature of visual temporal processing by examining whether the established effects of temporal adaptation were retinotopic or spatiotopic. The speed of a stimulus has been shown to influence its perceived duration (Kaneko et al., 2009, Brown, 1995) as has temporal frequency (Kanai et al., 2006). In this experiment we sought to remove questions of temporal frequency and speed and isolate the question

of retinotopic/spatiotopic adaptation as far as was possible. The results show that both conditions produced an adaptation effect, although the greater effect was for the retinotopic condition. The greatest differential between retinotopic and spatiotopic conditions in our study occurs for the adapt 640ms condition, a possible explanation for this may be that adapting the retina to the longer visual stimulus repeatedly could cause a reduction in the visibility of the stimulus. Since reduced visibility has been shown to produce a perceived temporal compression (Brigner, 1986, Terao et al., 2008) this could be a possible explanation for the difference between the two conditions. It is evident that duration adaptation occurs regardless of whether test stimuli are presented retinotopically or spatiotopically. This is in contrast to the previous studies investigating the retinotopic/spatiotopic question. These studies used a drifting grating to produce a contraction in the perceived duration of subsequently presented gratings which drifted at a different rate. In our study the visual stimuli were Gaussian blobs. It may well be that the sensation of movement of the object of regard is the crucial factor in obtaining the retinotopic/spatiotopic effect found in previous studies. This may lend support to a recent study which suggests that motor and beat (tasks involving movement/prediction as well as time) and sensory timing may involve separate mechanisms to tasks involving purely sensory timing (Buonomano et al., 2010). Alternatively, it may be that the separation between the retinotopic and spatiotopic stimuli was insufficient to produce a significant differential effect. In this case both the retinotopic and spatiotopic stimuli could lie within the same spatial field. This would place the effect at a higher level of the visual system where spatial fields are much coarser. One

way to investigate this possibility would be to repeat the experiment with an increased separation of (say) 20 degrees between the retinotopic and spatiotopic stimuli.

Chapter 11

Duration adaptation precedes sensory integration

11.1 Introduction

Despite the ecological importance of accurate time perception, a broadly reproduced finding is the remarkable extent to which perceived durations are modified by the context in which they are presented. Two recent examples of these misperceptions include duration distortions induced via multisensory integration (see Chapter 5) and those induced via duration adaptation (see Chapter 9). The former occurs when perceived visual duration is pulled in the direction of concurrently presented - but physically discrepant – auditory durations (which we shall call distractor stimuli) (Chen et al., 2009, Klink et al., 2011). The latter arises when recent sensory history contains consistent duration information. Post-adaptation, perception is characterised by repulsive duration aftereffects that are sensory specific, bidirectional and bandwidth-limited around the adapting duration (see Chapter 9).

An unresolved issue is the relative positioning of these perceptual distortions within the processing hierarchy. In principal, duration adaptation could be mediated via neurons known to respond selectively to a narrow range of stimulus durations centred on their preferred duration. These neurons are located in relatively early areas of the visual and auditory nervous systems

including the primary visual cortex (Duysens et al., 1996), and auditory midbrain (Casseday et al., 1994a, Faure et al., 2003, Sayegh et al., 2011, Duysens et al., 1996, Perez-Gonzalez et al., 2006, Brand et al., 2000, Leary et al., 2008). Alternatively, it could have a much later neural locus associated with duration-specific firing patterns observed in higher cortical areas including pre-frontal/frontal (Genovesio et al., 2006, Genovesio et al., 2009, Shinomoto et al., 2011) motor/pre-motor (Lebedev et al., 2008, Mita et al., 2009, Merchant et al., 2011) or lateral intraparietal regions (Leon et al., 2003).

Although the neural locus of multisensory integration-based duration distortions remains unknown, the integration of visual-vestibular directional cues and audio-visual speech components are increasingly being ascribed to higher, extrastriate areas such as dorsal medial superior temporal area (Fetsch et al., 2012, Gu et al., 2008) and the superior temporal sulcus (Nath et al., 2011), respectively. Nevertheless, a recent study also suggests a role for the primary sensory cortices in multisensory integration (Helbig et al., 2012).

If the two illusions (produced by duration adaptation and multisensory integration) are indeed generated by neural mechanisms with distinct positions in the processing hierarchy, the interaction between these illusions should be unidirectional. For example, if multisensory integration-based illusory durations are generated by mechanisms that reside in relatively early sections of the nervous system, these illusions will feed-forward to activate (later stage) duration adaptation mechanisms and thus induce visual duration aftereffects consistent with *perceived* rather than physical visual duration (Figure 87A). On the other hand, if duration adaptation illusions are

generated prior to multisensory integration, illusory visual durations generated by the latter will not be available to the former. In this situation, we would expect multisensory illusory durations to be incapable of activating duration adaptation mechanisms (Figure 87B).











Adaptation stimulus (physical)	Test stimulus	
	Physical	Perceived
(A) 		
(B) 		
 Auditory  Visual  Early DA, late MSI  Early MSI, late DA		

Figure 87 (A) If multisensory integration (MSI) precedes duration adaptation (DA) the visual adapting stimulus (in blue) will be perceptually lengthened via concurrent presentation with a longer auditory duration (red). Adaptation to this expanded visual duration will cause physically identical visual test duration to be perceptually compressed. (B) Alternatively, relatively early DA would be unaffected by concurrent auditory stimuli, thus maintaining perceptual equivalence between visual adapting and test durations and therefore veridical perception of the latter.

Chapter 11 seeks to address this question via a series of interconnected experiments. The first group of these experiments maps out the extent to which visual and auditory durations can be perceptually expanded or contracted via concurrent presentation of durations from the opposite modality (multisensory integration). The second group uses information from the first to investigate the proceeding hierarchy of duration perception.

Specifically, these experiments employ the multisensory conditions which promote maximal distortions of perceived duration and attempt to deploy these distorted durations as adapting stimuli.

11.2 General methods

11.2.1 Participants

Seven observers (four experienced and three naive) participated in these experiments.

11.2.2 Stimuli and apparatus

The visual stimulus was an isotropic Gaussian luminance blob the peak luminance of which was the maximum available. The blob subtended an angle of 2.1° at a viewing distance of 57cm and was presented on a uniform mean luminance grey background. It was presented in the centre of a gamma corrected monitor screen. Experiment 1 was carried out on a Compaq P1220 CRT. Experiment 2 was carried out on a Sony Trinitron GDM FW 900. The auditory stimulus consisted of a burst of white noise of presented via Sennheiser HD 280 headphones (see Chapter 7.3 for details).

11.3 Multisensory integration

11.3.1 Procedure

The experiment consisted of two parts. For the first part, subjects were asked to make a two interval, forced choice duration discrimination judgement between a visual stimulus of variable duration centred on 320ms (200ms, 240ms, 280ms, 320ms, 360ms, 400ms, or 440ms) and a 320ms visual stimulus accompanied by a distractor stimulus consisting of a burst of white noise of variable duration centred on 320ms on an approximately logarithmic scale (180ms, 200ms, 220ms, 250ms, 290ms, 320ms, 360ms, 400ms, 450ms, 510ms or 570ms).

Subjects were asked to ignore the sound component of this bimodal stimulus as far as possible and to simply judge the durations of the two visual stimuli relative to each other. The visual stimuli and bimodal stimuli were presented in random order and subjects indicated whether the first or second visual stimulus was longer by means of a button press. Receipt of the subject's response triggered the next presentation of the stimulus (see Figure 88A).

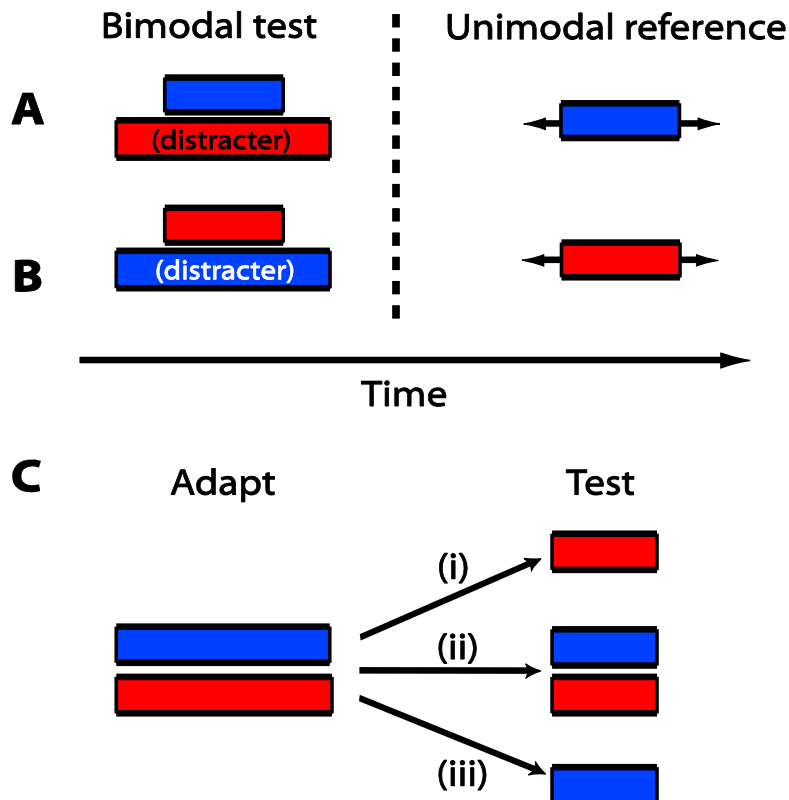


Figure 88(A & B) A schematic of the multisensory integration experiment where subjects were presented with a test stimulus consisting of 320ms duration from one modality (either vision (blue rectangles) or audition (red rectangles)) and a concurrently presented 'distracter' duration from the opposite modality. Following the presentation of the reference stimulus (200-440ms) subjects made unimodal 2AFC duration discrimination judgments ('which had the longer duration, test or reference stimulus?') whilst attempting to ignore the distracter duration. In the example provided above the distracter stimulus has a physically longer duration than the test stimulus but were in fact equally likely to be longer or shorter (see Figure 88 for resultant data) (C) A schematic of a control experiment where observers adapted to concurrently presented auditory and visual stimuli with physically matched durations before reproducing the duration of (i) auditory, (ii) bimodal (again, of matched duration), or (iii) visual test stimuli (see Figure 89B for resultant data and Methods for details).

A block of trials was complete when each auditory stimulus accompanying the blob was repeated 10 times. Each subject completed a minimum of 3

blocks of trials, giving a minimum total of 30 responses per presented duration.

The second part of the experiment was similar to the first except that the modalities were reversed. In other words, subjects were required to make a two interval, forced choice duration discrimination judgement between a burst of white noise of variable duration centred on 320ms (200ms, 240ms, 280ms, 320ms, 360ms, 400ms, or 440ms) and a 320ms burst of white noise accompanied by distractor stimulus consisting of a flash of white light of variable duration centred on 320ms on a logarithmic scale (180ms, 200ms, 220ms, 250ms, 290ms, 320ms, 360ms, 400ms, 450ms, 510ms or 570ms). Subjects were asked to ignore the visual component as far as possible and to simply judge the durations of the sounds relative to each other (Figure 88B).

The auditory and audio-visual combinations were presented in random order and subjects indicated whether the first or second sound was longer by means of the direction arrows on the computer keyboard (left for first interval longer, right for second interval longer). As in the first part of the experiment receipt of the subject's response triggered the next presentation of the stimuli and a block of trials was complete when each visual stimulus accompanying the sound was repeated 10 times. Each subject completed a minimum of 3 blocks of trials.

11.3.2 Results

For all observers, the percentage of ‘reference longer than test’ (Figure 88 A and B) responses for each condition was plotted as a function of reference duration and fitted with a logistic function of the form

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

where μ is the reference duration corresponding to the Point of Subjective Equality (‘PSE’ - the 50% response level on the psychometric function), and θ provides an estimate of duration discrimination threshold (approximately half the offset between the 27% and 73% response levels). In this way, PSE values were obtained for all observers in all conditions. Figure 89 plots these PSE values (averaged across observers) as a function of distracter duration. Blue squares denote the PSE when test and reference stimuli are visual, and the distracter stimulus is auditory (e.g. Figure 88A) whilst red squares denote the converse situation (e.g. Figure 88B). It can be seen that the auditory stimulus has a systematic effect on the perceived duration of the visual stimulus. As the auditory distracter stimulus increases in duration beyond 320ms, so the visual reference duration must be expanded in order to maintain perceptual equivalence with the 320ms visual test stimulus. This expansion is mirrored by perceptual compression of the test stimulus when auditory distracter stimuli are shorter than 320ms. This illusory expansion and contraction is maximal with auditory distracter durations of 510ms (visual PSE = 375ms) and 200ms (visual PSE = 269ms), respectively, beyond which

the PSE returns to baseline (see extreme right and left of the plot, respectively). In stark contrast to these marked distortions is the apparent absence of any effect of a visual distracter on the perceived duration of a simultaneously presented auditory stimulus (red data).

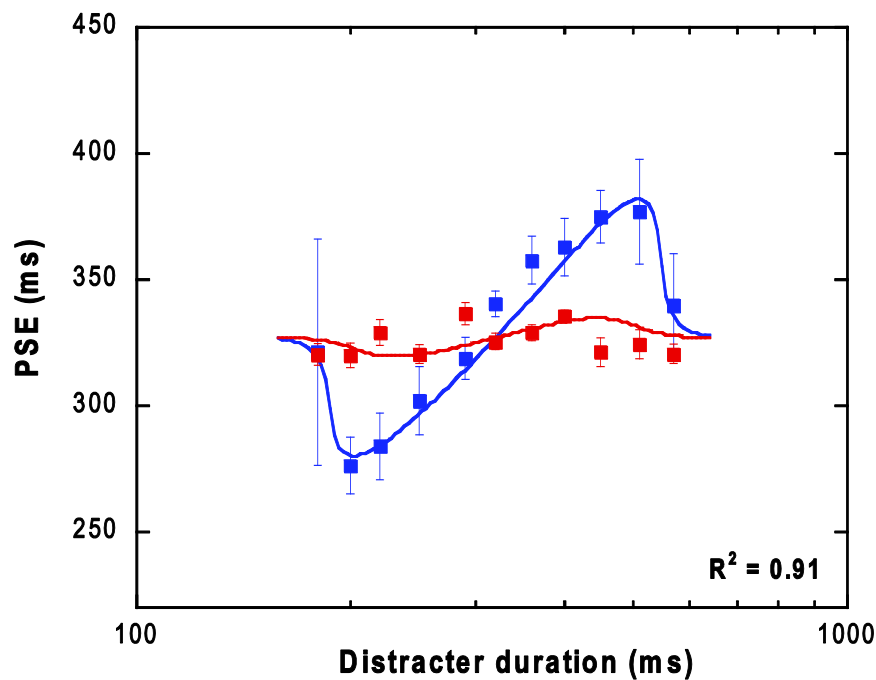


Figure 89 Data from the duration discrimination conditions depicted in Figure 87A&B. Data points represent the point of subjective equality (the physical reference duration that induced perceptual equivalence with the 320ms test duration from within the same modality) as a function of 11 distracter durations averaged across observers ($n=7$). Blue data points represent conditions where test and reference durations were visual (and distracter durations were auditory) with the reverse applying to red data points. Solid curves represent the auditory (red curve) and visual (blue curve) output of a Bayesian cue-combination model (see Methods for details). The R^2 value represents a combined fit of both model outputs to the data points. Error bars represent the standard error of the mean.

It has recently been suggested that unidirectional distortions such as those seen in Figure 89 are attributable to differential effects of visual vs. auditory stimulation on the ‘pulse rate’ of a putative pacemaker mechanism (Chen et

al., 2009, Klink et al., 2011). However, an alternative explanation is offered by developments in the field of multisensory cue-combination (Alais et al., 2004, Battaglia et al., 2003a, Ernst et al., 2002, Roach et al., 2006b, Heron et al., 2004, Ley et al., 2009, Parise et al., 2012, Gepshtein et al., 2003). These studies argue that multiple sources of information are combined by neural mechanisms that allocate differential perceptual weight between the modalities. This allocation process is governed by the relative sensory reliabilities of each information source, with the more reliable source of information dominating the integrated (multisensory) percept. In the current context, a reliability-based integration mechanism would allocate greater perceptual weight to audition than vision, in light of the former's superior temporal resolution (Westheimer, 1999, Goodfellow, 1934, Grondin et al., 1991a). A further feature of this computational framework relates to the physical discrepancy between the two sources: when sensory systems deliver sufficiently different perceptual estimates (e.g. markedly different spatial locations (Gepshtein et al., 2005, Heron et al., 2004, Kording et al., 2007) or temporal rates (Roach et al., 2006b), these estimates remain perceptually segregated from one another, regardless of their relative sensory reliabilities.

The next step therefore was to model Figure 89's illusory effects using a Bayesian model similar to that previously proposed for multisensory conflict between auditory and visual temporal rate (for details see Roach et al., 2006). The model includes a Gaussian likelihood function centred on the physical duration of the sensory stimulus whose duration is being judged (320ms), with a spread corresponding to the uncertainty (just-noticeable-difference,

JND) of that estimate. This spread was defined by the standard deviation of the psychometric function for the unimodal judgment of duration at a baseline of 320msec (48.75 and 21.27ms for vision and audition, respectively) confirming the superior temporal resolution of audition relative to vision. This likelihood function is combined with a prior representing knowledge of the probable correspondence between the stimulus to be judged and the distracter stimulus, built up via lifelong exposure to audiovisual stimuli with a common sensory duration (Ernst, 2005). The prior was centred on the distracter duration and possessed an uncertainty which was allowed to float so as to best fit the entire data set. The likelihood and prior were combined to produce a posterior distribution whose mode was taken to represent the final perceptual duration (Mamassian et al., 2002). Model calculations were carried out in logarithmic duration space and multiple iterations homed in on the prior which resulted in the best-fitting least squares fit to the data. The model has two free parameters, the width of the prior and the height of a pedestal on which the prior was superimposed. In addition, it was found necessary to incorporate a small (7ms) vertical shift of the JNDs from the model which suggests that - across all conditions - bimodal test stimuli are perceived as fractionally longer than unimodal reference stimuli. The model output is plotted against the data in Figure 89.

The model provides an excellent fit ($R^2=0.91$) to the combined data set for visual (blue function) and auditory (red function) conditions. The only difference between auditory and visual versions of the model's output is the widths of their underlying visual and auditory likelihood functions. Thus, differences in the magnitude of the interaction effect (i.e. the relative

gradients of the red and blue model outputs) are well described by a computational framework where relative perceptual weighting is critically dependent on differences in the sensitivity (JND) of the visual and auditory systems to changes in stimulus duration (Westheimer, 1999, Goodfellow, 1934, Grondin et al., 1991a).

Figure 89's data shows that audiovisual sensory integration produces marked distortions of visual duration. Having ascertained the physical audiovisual duration discrepancies that maximise these distortions, it is possible to investigate the relationship between sensory integration-based mechanisms and those underpinning duration adaptation-induced illusions. Specifically, what are their relative positions within the nervous system's sensory processing hierarchy? If sensory integration precedes duration adaptation, exposure to illusory durations generated via multisensory integration will generate duration aftereffects comparable to those induced by real durations (Heron et al., 2012) (see Figure 87A). Conversely, if the mechanisms underpinning duration adaptation have their neural locus at an earlier processing stage, sensory integration-based illusions will leave pre- and post-adaptation duration estimates unaffected (Figure 87B).

11.4 Adaptation

11.4.1 Procedure

The second experiment involved a method of duration reproduction – use of a motor action (the press of a computer key) to reproduce the duration of a preceding sensory event. Before beginning the experiment, subjects were given extensive practice in reproducing the duration of a Gaussian blob presented on the screen for one of three different durations (160ms, 320ms or 640ms). Initially, observers were given immediate feedback on the screen informing them of the actual duration of the visual stimulus and the duration for which they held down the computer key. At the end of a block subjects were given the means of their reproduction times for each of the physical durations.

In the test phase subjects were presented with a Gaussian blob for one of three durations. 50% of these were 320ms and on average the other 50% was divided equally between 160ms and 640ms. The object of the 160ms and 640ms presentations was to prevent subjects simply adopting a repetitive motor action resulting in fixed keypress duration. Data from these short and long conditions was discarded. Subjects performed 30 reproductions per block and a minimum of 3 blocks. No feedback was given during the experiment.

The experiment utilised two ‘adaptation’ conditions in which subjects adapted to a bimodal stimulus consisting of a 320ms Gaussian blob accompanied by

an auditory burst of white noise of either 200ms or 510ms. These auditory values were chosen because they induce the maximum perceived distortion of a 320ms visual duration (see Figure 89), and thus the maximum perceived difference between visual adapting and test stimuli. The 'perceptual difference' conditions were compared with 'physical difference' conditions where observers adapted to separate unimodal blocks of 260ms and 375ms duration visual stimuli.

These two durations were used because they represent the average perceived duration of the visual stimulus in the adaptation experiment when accompanied by the relevant durations of white noise (200ms or 510ms) determined in Experiment 1 (see Figure 89).

Subjects were presented with the bimodal adapting stimuli one hundred times. This was followed by 4 top up stimuli before the subject made the first response. Following each response the subject was presented with another 4 top ups before they reproduced the next duration.

The extent of duration adaptation was quantified by subtracting the reproduced duration following adaptation to the longer (real or illusory) stimulus from that following adaptation to the shorter stimulus. Any 'rebound' type of duration adaptation effect would therefore manifest itself as a positive value.

11.4.2 Results

Figure 90A plots aftereffect magnitude following adaptation to *perceptual* (black bar) and *physical* (grey bar) differences between adapting and test stimuli. As expected, physical differences between adapting and test durations generate robust duration aftereffects comparable with those reported elsewhere in the literature (Heron et al., 2012). However, consistent with the scenario depicted in Figure 87B, perceptual differences fails to generate any significant aftereffects a result which suggests that duration adaptation mechanisms do not have access to illusory durations generated via multisensory integration.

Alternatively, the absence of an effect for the illusory stimuli may arise from the bimodal nature of the adapting stimulus and the unimodal nature of the test stimulus. For example, the auditory stimulus may have diverted attention away from its visual counterpart – a scenario known to ameliorate later-stage aftereffects (Chaudhuri, 1990, Montaser-Kouhsari et al., 2004). It is also possible that, a greater degree of categorical similarity between adapting and test stimuli (e.g. adapt bimodal, test bimodal) may be required to activate duration adaptation mechanisms, as is the case for other high-level aftereffects (Bestelmeyer et al., 2008, Little et al., 2011, Rotshtein et al., 2005). This possibility was investigated by conducting a control experiment identical to the previous experiment with the following exceptions: auditory and visual components of the bimodal adapting stimuli had physically identical durations (either 640 or 160ms (durations known to produce reliable duration aftereffects in both visual and auditory domains (Heron et al., 2012),

presented in separate blocks). Test stimuli were either auditory, bimodal (again, of matched physical duration) or visual (see Figure 93). If attention or categorical similarity is/are responsible for the absence of effect seen with illusory stimuli, aftereffects should only be elicited with matching adaptation and test stimuli. Figure 90B shows equivalent aftereffects are manifest with all three test stimuli demonstrating that unisensory auditory and visual durations can be distorted via multisensory duration adaptation. In two conditions, (test auditory and test bimodal) the effects are significant ($t(6)$ 5.8691, $p < 0.005$ and $t(6)$ 3.4724, $p < 0.05$). The large standard deviation of the results of the test visual condition renders it non-significant ($t(6)$ 1.783, $p = 0.1249$). However, this condition is nevertheless significantly different from the illusory adapt condition ($t(6)$ 2.4664, $p < 0.05$) and not significantly different from the visual adapt control condition ($t(6)$ 0.10242, $p = 0.9218$). A further possibility is that our illusory stimuli *are* available to the duration adaptation mechanism (i.e. sensory integration occurs *prior* to temporal recalibration) but their failure to generate duration aftereffects simply reflects the fact that *perceptual* adapt-test differences are ineffective drivers of the adaptation mechanism when compared to *physical* adapt-test differences. In other words, illusory durations may be perceptually distinct from their (duration-matched) physical counterparts. This scenario is not compatible with the threshold data from our initial experiment where the variation in visual thresholds mirrors the variation in PSE (c.f. the sigmoidal distribution of the blue data in Figures 90 and 91A). This suggests that observer's perception of these illusory durations conforms to one of the fundamental properties of real

durations: the proportional relationship between duration discrimination thresholds and mean estimated duration (Weber's law for duration).

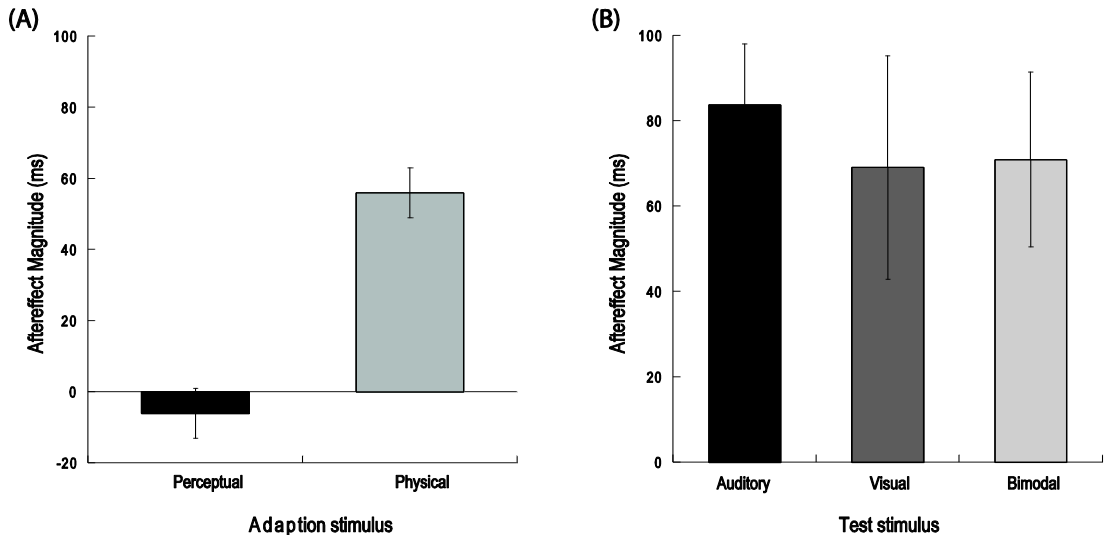


Figure 90 (A) Mean aftereffect magnitude (the arithmetic difference between reproduced durations following adaptation to relatively short and long duration stimuli – see Methods for details) following adaptation to perceptual differences between adapting and test durations (black bar). These perceptual differences were introduced via the MSI associated with stimulus configuration shown in Figure 88. These effects were compared with duration-matched physical adapt-test differences (grey bar) (n=7). (B) Data from a control experiment where mean aftereffect magnitude was compared for the three adapt-test conditions shown in Figure 87C. Observers reproduced visual, auditory and bimodal test stimuli following adaptation to the concurrent presentation of duration matched 160ms or 640ms duration auditory and visual stimuli (n=7). Errors bars indicate the standard error of the mean.

This relationship is illustrated in Figure 91B where multiplying Figure 90's visual PSE data by a single Weber fraction (averaged across distracter

durations) of 0.18 (SD = 0.01) allows threshold and PSE data sets to collapse on top of one another.

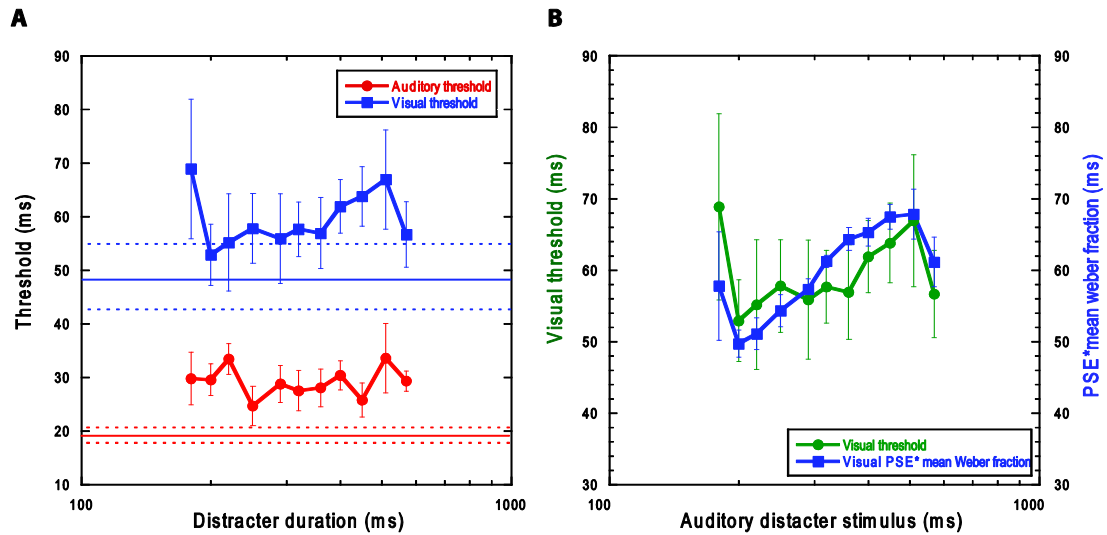


Figure 91 (A) Mean threshold data arising from visual (blue squares) or auditory (red squares) duration discrimination judgments made in the presence of distracter stimuli presented to the opposite modality (as depicted in Figures 87A and B respectively) . Horizontal blue and red lines represent unimodal visual and auditory (i.e. without distracter stimuli) thresholds, respectively. (B) Green data points represent the same unimodal visual threshold data shown in figure 90A, whereas blue data points represent the visual PSE data shown in Figure 88 (blue squares) multiplied by a single Weber fraction of 0.18 (see main text for details). For both plots, error bars and dashed horizontal lines indicate the standard error of the mean (n=7).

11.4.3 Discussion

The experiments in this chapter were designed to investigate the hierarchy of sensory processing by measuring the interdependency of two temporal distortions, one induced via sensory integration and one via adaptation. In our first experiment, we show that the integration of concurrently presented

visual and auditory duration information induces marked distortions of perceived duration: visual duration is expanded or contracted in the direction of the auditory stimulus yet - in relative terms - auditory judgments are impervious to the concurrent presentation of visual durations. These visual distortions are bandwidth-limited to approximately 0.2 log units either side of our central (320ms) duration (Figure 89). Subsequently, either physical or perceptual differences were introduced between adapting and test durations. Whilst physical adapt-test differences generated robust aftereffects (Figure 89A), test stimuli were perceived veridically when perceptual adapt-test differences were generated by Multisensory integration (MSI) mechanisms (Figure 89 and 90A). These findings reveal a sensory hierarchy where duration adaptation represents a neural operation that is executed prior to the integration of temporal information across the senses.

On first inspection, the sensory integration data appears to suggest that visual durations have no influence on the perception of auditory durations (Figure 89 – red data points), a finding in seeming agreement with recent reports (Chen et al., 2009, Klink et al., 2011). However, our model output (Figure 90 – red curve) shows that this interpretation is misleading: the model predicts small distortions of perceived auditory duration which can masquerade as an absence of interaction. That the interaction is so grossly asymmetrical across visual and auditory domains is entirely consistent with the extent to which the model is influenced by differential temporal sensitivity between the senses (Ernst, 2005). Without resorting to putative visual, auditory or audiovisual pacemaker-based explanations for this asymmetry, the model provides a straightforward, parsimonious account of multisensory

duration perception. The model output appears to share common properties with the integration of auditory and visual temporal rate information (Roach et al., 2006b), most notably the interaction between individual uncertainties in the two unimodal duration estimates and the width of the prior which controls the level of discrepancy across which the two signals are treated as having a common cause. It is worth noting that auditory distortions of perceived visual duration found in experiment 1 occur in spite of subjects being instructed to devote all their attention to the visual stimulus, suggesting that whilst these distortions may have been of a greater magnitude under conditions of divided attention, the marked interaction seen in Figure 90 forms a mandatory – attention independent – form of sensory integration (Vroomen et al., 2001a, Helbig et al., 2008b).

Unidirectional interaction between the current study's two temporal distortions has implications for the neural loci at which they are generated. Firstly, the integration of duration information across sensory systems appears to be a relatively late stage process, making extrastriate regions such as the superior temporal sulcus (audio-visual speech integration (Beauchamp et al., 2010, Nath et al., 2011, Pasalar et al., 2010)) and dorsal medial superior temporal area (visual-vestibular motion integration (Fetsch et al., 2012, Gu et al., 2008)) more credible neural sites than the primary sensory cortices. Conversely, duration adaptation effects displayed characteristics consistent with having arisen at a relatively early stage of sensory processing, a finding which provides support for time perception mechanisms underpinned by the duration selective neurons found in cat primary visual cortex (Duysens et al., 1996), cat auditory cortex (He et al.,

1997) and auditory midbrain of amphibians (Gooler et al., 1992, Leary et al., 2008), bats (Casseday et al., 2000, Mora et al., 2004, Faure et al., 2003, Casseday et al., 1994a), rats (Perez-Gonzalez et al., 2006), guinea pigs (Wang et al., 2006, Yin et al., 2008) and mice (Brand et al., 2000, Xia et al., 2000) (for a recent review see (Aubie et al., 2012)). A neural locus for adaptation within the primary sensory pathways would help to explain why related phenomena show narrow (Ayhan et al., 2009a), retinotopic spatial tuning, and a lack of dichoptic transfer (Bruno et al., 2010). A further prediction for 'early adaptation' is that illusory durations generated via adaptation should feed forward to influence later-stage multisensory integration. It is proposed to test for this in a follow up experiment where observers adapt to relatively long or short auditory durations before a test phase consisting of duration discrimination judgments between a bimodal test duration (concurrently presented 320ms visual and auditory durations) and a variable visual reference. If duration adaptation precedes multisensory integration, the former will induce distortions in the auditory component of the test stimulus. Multisensory integration would then transfer these distortions to the perceived duration of the visual test stimuli. Another interesting question for further work would be to address the extent to which *within*-modality perception is influenced by adaptation-based duration distortions. For example, if adaptation induced a sufficiently early expansion of perceived duration, we should see the perceived luminance/loudness of very brief stimuli increase in line with an equivalent increase in physical duration.

In summary, duration perception is a highly flexible process that is modified via the twin mechanisms of sensory integration and temporal adaptation.

Although these mechanisms produce superficially similar distortions of temporal perception, the experiments described in this chapter tease apart a clear separation in their processing order: It appears that, similar to visual spatial frequency or auditory pitch, event duration is a low-level stimulus attribute that undergoes early, adaption-based recalibration prior to the formation of multisensory perception.

Chapter 12

Conclusions

Our understanding of the mechanisms underlying “automatic” human time perception remains incomplete. This thesis is concerned with sensory timing and our stimuli have all been static. However this represents only a small part of the whole story. For instance, when confronted with moving or changing stimuli, we are required to compute the speed of movement/change in order to predict the future position/state of the object of regard. This implies a link between spatial and temporal processing and it seems likely that different or additional mechanisms are required for tasks of this nature.

Many other factors may influence our perception of a brief duration. These include attention, stimulus nature, memory, arousal and not least adaptation. All these factors will have their own underlying influences which contribute to our perception of temporal extent. In addition the point at which cognitive factors come into play is not clear. General opinion seems to be that this can occur anywhere between 500ms and 2,000ms.

The experiments described in this thesis contribute to our overall knowledge of the way the human nervous system processes duration and provides some pointers to the possible mechanisms involved in automatic sensory timing.

In chapter 8, a series of experiments are outlined, from which we may draw a number of conclusions. Firstly we are able to confirm the findings of Walker et al (1981), that duration adaptation aftereffects occur intra modally. This casts doubt on the long held central pacemaker/accumulator theory described in Chapter 2. In addition we find these repulsive aftereffects to be tuned in a similar way to spatial frequency adaptation effects. We have found that these aftereffects occur in a similar way over three different ranges within sub second timing. We have also produced a model involving overlapping duration sensitive channels which closely fits our data. The precise nature of these channels is not as yet clear, but likely candidates include coincidence detection mechanisms or readout neurons which detect the state of a neural network.

One way to further investigate the nature of duration channels might be by use of a simultaneous adaptation paradigm. The data produced in our experiments is very similar in nature to that found in adaptation experiments involving spatial frequency and orientation. In these experiments adaptation to stimuli of a particular spatial frequency or orientation produces a post adaptation bias *and* an increase in detection threshold centred on the adapted spatial frequency or orientation. The proposed experiment would have subjects adapt concurrently to stimuli which are shorter and longer in duration than the test duration. This would give us a measure of post adaptation discrimination but with no bias. Our channel-based model would predict that thresholds for duration discrimination tested at the adapted duration should rise due to a damping down of the neurons responsible for these durations. The threshold for the test duration positioned *between* the

adapted durations, on the other hand, should fall because the population response will be sharpened, due to the reduced response of the neurons on either side of the test duration (See Chapter 9).

In chapter 10 the specificity of duration adaptation aftereffects is investigated further. These duration adaptation aftereffects are not produced by brief single adaptors. The effect cannot be isolated for adapting stimuli of different spatial frequencies or orientations suggesting that its locus lies outside V1. We found duration adaptation to be non spatially specific, in other words it occurs throughout visual space suggesting that different mechanisms may be involved than those found when subjects adapt to moving gratings.

The fact that duration adaptation is found to be intra modal, suggests that it occurs before multimodal integration, a conclusion which is confirmed by the experiments described in Chapter 11. Here we also produce a model for cue combination involving Bayesian integration. Specifically, the model combines a likelihood function based on the unimodal uncertainty of the visual and auditory estimates, together with a prior produced by life long experience of audio/visual stimuli of equal duration. The mode of the resulting distribution is taken to be the perceived duration. This model provides an impressive account of the data.

This thesis also throws new light on the “oddball effect” in which an unexpected/different stimulus presented within a series of identical stimuli is perceived as being longer (Chapter 8). We find that the low level characteristics of the visual stimulus can have a profound effect on the perceived duration of an oddball stimulus. Indeed if a Gabor of 2cpd is used

as a standard stimulus” and 8cpd as the “oddball stimulus”, the effect is a *contraction* of perceived duration relative to preceding repeated stimuli. These experiments also find that the perceived duration of mid range spatial frequencies is greater than that of more peripheral frequencies. A surprising finding was that Gabors of 2cpd were perceived as being longer than our auditory reference stimuli. This finding is contrary to the established wisdom that auditory stimuli are perceived as being longer than visual stimuli.

This effect of spatial frequency on the perceived duration of oddballs is found over different ranges of sub second stimuli, but is not a proportion of base duration. The effect appears to be additive, which is consistent with variations in visual persistence across different spatial frequencies, although it may be that mid range spatial frequencies are perceived before their more extreme neighbours. Further experiments are needed in order discover whether a differential in onsets or offsets for different spatial frequencies exists. This would help us to uncover the locus of this additive effect.

To sum up, the mechanisms underlying automatic timing in humans represent a fundamental factor in our overall understanding of human sensory perception. The experiments described in this thesis contribute towards the growing body of research in this area and the suggested further experiments should advance our, as yet, limited understanding of this field. An understanding of the way automatic timing forms a coherent, yet effortless whole, and the brain mechanisms responsible remains the ultimate goal.

References

- AAEN-STOCKDALE, C., HOTCHKISS, J., HERON, J. & WHITAKER, D. 2011. Perceived time is spatial frequency dependent. *Vision Research*, 51, 1232-1238.
- ACKERMANN, H., GRABER, S., HERTRICH, I. & DAUM, I. 1997. Categorical speech perception in cerebellar disorders. *Brain and Language*, 60, 323-331.
- ALAIS, D. & BURR, D. 2004. The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 14, 257-262.
- ALAIS, D., MORRONE, C. & BURR, D. 2006. Separate attentional resources for vision and audition. *Proceedings of the Royal Society B-Biological Sciences*, 273, 1339-1345.
- ALEXANDER, D. M. & VAN LEEUWEN, C. 2010. Mapping of contextual modulation in the population response of primary visual cortex. *Cognitive Neurodynamics*, 4, 1-24.
- ALLAN, L. G. 1979. Perception of time. *Perception & Psychophysics*, 26, 340-354.
- ANGELUCCI, A. & BULLIER, J. 2003. Reaching beyond the classical receptive field of V1 neurons: horizontal or feedback axons? *Journal of Physiology-Paris*, 97, 141-154.
- AUBIE, B., BECKER, S. & FAURE, P. 2009. Computational models of millisecond level duration tuning in neural circuits. *Journal of Neuroscience*, 29, 9255-70.
- AUBIE, B., SAYEGH, R. & FAURE, P. A. 2012. Duration Tuning across Vertebrates. *Journal of Neuroscience*, 32, 6373-90.
- AYHAN, I., BRUNO, A., NISHIDA, S. & JOHNSTON, A. 2009. The spatial tuning of adaptation-based time compression. *Journal of Vision*, 9, 1-12.
- BARO, J. A., BRZEZICKI, L. J., LEHMKUHLE, S. & HUGHES, H. C. 1992. The perceived duration of gratings. *Perception*, 21, 161-166.
- BARTOLO, R. & MERCHANT, H. 2009. Learning and generalization of time production in humans: rules of transfer across modalities and interval durations. *Experimental Brain Research*, 197, 91-100.
- BATTAGLIA, P. W., JACOBS, R. A. & ASLIN, R. N. 2003. Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America a-Optics Image Science and Vision*, 20, 1391-1397.
- BEAR, M. C., B; PARADISO, M; 2001. *Neuroscience: Exploring the brain*, Baltimore MA, Lippincott williams & Wilkins.
- BEAUCHAMP, M. S., NATH, A. R. & PASALAR, S. 2010. fMRI-Guided Transcranial Magnetic Stimulation Reveals That the Superior Temporal Sulcus Is a Cortical Locus of the McGurk Effect. *Journal of Neuroscience*, 30, 2414-2417.
- BECKER, M. W. & RASMUSSEN, I. P. 2007. The rhythm aftereffect: Support for time sensitive neurons with broad overlapping tuning curves. *Brain and Cognition*, 64, 274-281.
- BEHAR, I. & BEVAN, W. 1961. Perceived duration of auditory and visual intervals cross modal comparison and interaction. *American Journal of Psychology*, 74, 17-26.
- BERTELSON, P., VROOMEN, J., DE GELDER, B. & DRIVER, J. 2000. The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics*, 62, 321-332.
- BESTELMEYER, P. E. G., JONES, B. C., DEBRUINE, L. M., LITTLE, A. C., PERRETT, D. I., SCHNEIDER, A., WELLING, L. L. M. & CONWAY, C. A. 2008. Sex-contingent face aftereffects depend on perceptual category rather than structural encoding. *Cognition*, 107, 353-365.
- BLAKE, R. & SEKULER, R. 2006. *Perception*, Fifth edition, New York: McGraw-Hill.

- BLAKEMORE, C. & CAMPBELL, F. W. 1969(a). On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology-London*, 203, 237-260.
- BLAKEMORE, C. & CAMPBELL, F. W. 1969(b). Adaptation to spatial stimuli. *Journal of Physiology (Cambridge)*, 200, 11P-13P.
- BLAKEMORE, C., NACHMIAS, J. & SUTTON, P. 1970. The perceived spatial frequency shift: Evidence for frequency-selective neurons in the human brain. *Journal of Physiology-London*, 210, 727-750.
- BLASDEL, G. G. 1992. Orientation selectivity, preference, and continuity in the monkey striate cortex. *Journal of Neuroscience*, 12, 3139-3161.
- BOBKO, D. J., SCHIFFMAN, H. R., CASTINO, R. J. & CHIAPPETTA, W. 1977(a). Contextual effects in duration experience. *American Journal of Psychology*, 90, 577-586.
- BOBKO, D. J., THOMPSON, J. G. & SCHIFFMAN, H. R. 1977(b). Perception of brief temporal intervals - power functions for auditory and visual stimulus intervals. *Perception*, 6, 703-709.
- BOUDREAU, J. C. & TSUCHITANI, C. 1968. Binaural interaction in the cat superior olive S segment. *Journal of Neurophysiology*, 31, 442-54.
- BOWLING, A. & LOVEGROVE, W. 1982. Iconic memory: fallacies persist. *Perception & Psychophysics*, 31, 194-198.
- BRAITENBERG, V. & BRAITENBERG, C. 1979. Geometry of orientation columns in the visual cortex. *Biological Cybernetics*, 33, 179-186.
- BRAND, A., URBAN, A. & GROTHE, B. 2000. Duration tuning in the mouse auditory midbrain. *Journal of Neurophysiology*, 84, 1790-1799.
- BRIGNER, W. L. 1986. Effect of perceived brightness on perceived time. *Perceptual and Motor Skills*, 63, 427-430.
- BRODY, C. D., HERNANDEZ, A., ZAINOS, A. & ROMO, R. 2003. Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. *Cerebral Cortex*, 13, 1196-1207.
- BROWN, S. W. 1995. Time, change and motion: The effects of stimulus movement on temporal perception. *Perception & Psychophysics*, 57, 105-116.
- BROWN, S. W. 1998. Influence of individual differences in temporal sensitivity on timing performance. *Perception*, 27, 609-625.
- BROWN, S. W. & BOLTZ, M. G. Attentional processes in time perception: Effects of mental workload and event structure. 70th Annual Meeting of the Eastern Psychological Association, 1999 Providence, Ri. 600-615.
- BROWN, S. W. & BOLTZ, M. G. 2002. Attentional processes in time perception: Effects of mental workload and event structure. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 600-615.
- BRUNO, A., AYHAN, I. & JOHNSTON, A. 2010. Retinotopic adaptation-based visual duration compression. *Journal of Vision*, 10, 1-18.
- BUETI, D., BAHRAMI, B. & WALSH, V. 2008(a). Sensory and association cortex in time perception. *Journal of Cognitive Neuroscience*, 20, 1054-1062.
- BUETI, D., VAN DONGEN, E. & WALSH, V. 2008(b). The Role of Superior Temporal Cortex in Auditory Timing. *PLoS One*, 3, Article No.: e2481.
- BUETI, D., WALSH, V., FRITH, C. & REES, G. 2008(c). Different brain circuits underlie motor and perceptual representations of temporal intervals. *Journal of Cognitive Neuroscience*, 20, 204-214.

- BUHUSI, C. V. & MECK, W. H. 2002. Differential effects of methamphetamine and haloperidol on the control of an internal clock. *Behavioral Neuroscience*, 116, 291-297.
- BUHUSI, C. V. & MECK, W. H. 2005. What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, 6, 755-765.
- BUONOMANO, D. V. 2000. Decoding temporal information: A model based on short-term synaptic plasticity. *Journal of Neuroscience*, 20, 1129-1141.
- BUONOMANO, D. V. & KARMARKAR, U. R. 2002. How do we tell time? *Neuroscientist*, 8, 42-51.
- BUONOMANO, D. V. & LAJE, R. 2010. Population clocks: motor timing with neural dynamics. *Trends in Cognitive Sciences*, 14, 520-527.
- BUONOMANO, D. V. & MAUK, M. D. 1994. Neural-network model of the cerebellum-temporal discrimination and the timing of motor responses. *Neural Computation*, 6, 38-55.
- BUONOMANO, D. V. & MERZENICH, M. M. 1995. Temporal information transformed into a spatial code by a neural network with realistic properties. *Science*, 267, 1028-1030.
- BURR, D. & ALAIS, D. 2006. Combining visual and auditory information. *Visual Perception, Pt 2: Fundamentals of Awareness: Multi-Sensory Integration and High-Order Perception*, 155, 243-258.
- BURR, D., BANKS, M. S. & MORRONE, M. C. 2009. Auditory dominance over vision in the perception of interval duration. *Experimental Brain Research*, 198, 49-57.
- BURR, D., TOZZI, A. & MORRONE, M. 2007. Neural mechanisms for timing visual events are spatially selective in real-world coordinates. *Nature Neuroscience*, 10, 423-425.
- BUSHARA, K. O., HANAKAWA, T., IMMISCH, I., TOMA, K., KANSAKU, K. & HALLETT, M. 2003. Neural correlates of cross-modal binding. *Nature Neuroscience*, 6, 190-195.
- CALVERT, G. A., BRAMMER, M. J., BULLMORE, E. T., CAMPBELL, R., IVERSEN, S. D. & DAVID, A. S. 1999. Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport*, 10, 2619-2623.
- CALVERT, G. A., BULLMORE, E. T., BRAMMER, M. J., CAMPBELL, R., WILLIAMS, S. C. R., MCGUIRE, P. K., WOODRUFF, P. W. R., IVERSON, S. D. & DAVID, A. S. 1997. Activation of auditory cortex during silent lipreading. *Science*, 276, 593-596.
- CALVERT, G. A. & THESEN, T. 2004. Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology-Paris*, 98, 191-205.
- CAMPBELL, F. W. & ROBSON, J. G. 1968. Application of Fourier analysis to the visibility of gratings. *The Journal of physiology*, 197, 551-66.
- CAMPBELL, L. A. & BRYANT, R. A. 2007. How time flies: A study of novice skydivers. *Behaviour Research and Therapy*, 45, 1389-1392.
- CAPPE, C., ROUILLER, E. M. & BARONE, P. 2009. Multisensory anatomical pathways. *Hearing Research*, 258, 28-36.
- CASSEDAY, J. H., EHRLICH, D. & COVEY, E. 1994. Neural Tuning for Sound Duration - Role of Inhibitory Mechanisms in the Inferior Colliculus. *Science*, 264, 847-850.
- CASSEDAY, J. H., EHRLICH, D. & COVEY, E. 2000. Neural measurement of sound duration: Control by excitatory-inhibitory interactions in the inferior colliculus. *Journal of Neurophysiology*, 84, 1475-1487.
- CEVIK, M. O. 2003. Effects of methamphetamine on duration discrimination. *Behavioral Neuroscience*, 117, 774-784.
- CHAUDHURI, A. 1990. Modulation of the Motion Aftereffect by Selective Attention. *Nature*, 344, 60-62.

- CHEN, K. M. & YEH, S. L. 2009. Asymmetric cross-modal effects in time perception. *Acta Psychologica*, 130, 225-234.
- CHIK, D., BORISYUK, R. & KAZANOVICH, Y. 2009. Selective attention model with spiking elements. *Neural Networks*, 22, 890-900.
- CHKLOVSKII, D. B. & KOULAKOV, A. A. 2004. Maps in the brain: What can we learn from them? *Annual Review of Neuroscience*, 27, 369-392.
- CHURCH, R. M. & BROADBENT, H. A. 1990. Alternative representation of time, number, and rate. *Cognition*, 37, 55-81.
- CHURCH, R. M., MECK, W. H. & GIBBON, J. 1994. Application of scalar timing theory to individual trials. *Journal of Experimental Psychology-Animal Behavior Processes*, 20, 135-155.
- CHURCH, R. M., MILLER, K. D., MECK, W. H. & GIBBON, J. 1991. Symmetrical and asymmetrical sources of variance in temporal generalization. *Animal Learning & Behavior*, 19, 207-214.
- CORNSWEET, T. N. 1962. The staircase-method in psychophysics. *The American Journal of Psychology*, 75, 485-91.
- COULL, J. T., VIDAL, F., NAZARIAN, B. & MACAR, F. 2004. Functional anatomy of the attentional modulation of time estimation. *Science*, 303, 1506-1508.
- CREELMAN, C. D. 1962. Human discrimination of auditory duration. *The Journal of the Acoustical Society of America*, 34, 582-593.
- DARTNALL, H. J. A., BOWMAKER, J. K. & MOLLON, J. D. 1983. Human visual pigments-microspectrophotometric results from the eyes of 7 persons. *Proceedings of the Royal Society B-Biological Sciences*, 220, 115-130.
- DEVALOIS, R. & DEVALOIS, K. 1990. *Spatial vision*, Oxford, UK: Oxford University Press.
- DEVALOIS, R. L., ALBRECHT, D. G. & THORELL, L. G. 1982. Spatial-frequency selectivity of cells in macaque visual-cortex. *Vision Research*, 22, 545-559.
- DOESBURG, S. M., ROGGEVEEN, A. B., KITAJO, K. & WARD, L. M. 2008. Large-scale gamma-band phase synchronization and selective attention. *Cerebral Cortex*, 18, 386-396.
- DRIVER, J. & NOESSELT, T. 2008. Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron*, 57, 11-23.
- DROIT-VOLET, S. & GIL, S. 2009. The time-emotion paradox. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 364, 1943-1953.
- DUHAMEL, J. R., BREMMER, F., BENHAMED, S. & GRAF, W. 1997. Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature*, 389, 845-848.
- DUHAMEL, J. R., COLBY, C. L. & GOLDBERG, M. E. 1992. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, 255, 90-92.
- DUYSENS, J., SCHAAFSMA, S. J. & ORBAN, G. A. 1996. Cortical off response tuning for stimulus duration. *Vision Research*, 36, 3243-3251.
- EAGLEMAN, D. M. 2008. Human time perception and its illusions. *Current Opinion in Neurobiology*, 18, 131-136.
- EAGLEMAN, D. M. & PARIYADATH, V. 2009. Is subjective duration a signature of coding efficiency? *Philosophical Transactions of the Royal Society B-Biological Sciences*, 364, 1841-1851.
- EDWARDS, S. B., GINSBURGH, C. L., HENKEL, C. K. & STEIN, B. E. 1979. Sources of sub-cortical projections to the superior colliculus in the cat. *Journal of Comparative Neurology*, 184, 309-329.
- ERNST, M. O. 2005. A Bayesian view on multimodal cue integration. In: KNOBLICH, G., THORNTON, I., GROSJEAN, M. & SHIFFRAN, M. (eds.) *Human body perception from the inside out*. New York, NY: Oxford University Press.

- ERNST, M. O. & BANKS, M. S. 2002. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415, 429-433.
- ERNST, M. O. & BULTHOFF, H. H. 2004. Merging the senses into a robust percept. *Trends in Cognitive Sciences*, 8, 162-169.
- FALCHIER, A., CLAVAGNIER, S., BARONE, P. & KENNEDY, H. 2002. Anatomical evidence of Multimodal integration in primate striate cortex. *Journal of Neuroscience*, 22, 5749-5759.
- FAURE, P. A., FREMOUW, T., CASSEDAY, J. H. & COVEY, E. 2003. Temporal masking reveals properties of sound-evoked inhibition in duration-tuned neurons of the inferior colliculus. *Journal of Neuroscience*, 23, 3052-3065.
- FELL, J., FERNANDEZ, G., KLAVER, P., ELGER, C. E. & FRIES, P. 2003. Is synchronized neuronal gamma activity relevant for selective attention? *Brain Research Reviews*, 42, 265-272.
- FENDRICH, R. & CORBALLIS, P. M. 2001. The temporal cross-capture of audition and vision. *Perception & Psychophysics*, 63, 719-725.
- FERRANDEZ, A. M., HUGUEVILLE, L., LEHERICY, S., POLINE, J. B., MARSAULT, C. & POUTHAS, V. 2003. Basal ganglia and supplementary motor area subtend duration perception: an fMRI study. *Neuroimage*, 19, 1532-1544.
- FETSCH, C. R., POUGET, A., DEANGELIS, G. C. & ANGELAKI, D. E. 2012. Neural correlates of reliability-based cue weighting during multisensory integration. *Nature Neuroscience*, 15, 146-185.
- FLETCHER, H. F. 1940. Auditory patterns. *Rev.Mod.Phys.*, 12, 47-65.
- FOX, J. J. & SCHROEDER, C. E. 2005. The case for feedforward multisensory convergence during early cortical processing. *Neuroreport*, 16, 419-423.
- FRAISSE, P. 1984. Perception and estimation of time. *Annual Review of Psychology*, 35, 1-36.
- GALAZYUK, A. V. & FENG, A. S. 1997. Encoding of sound duration by neurons in the auditory cortex of the little brown bat, *Myotis lucifugus*. *Journal of Comparative Physiology a-Sensory Neural and Behavioral Physiology*, 180, 301-311.
- GALLETI, C., BATTAGLINI, P. P. & FATTORI, P. 1995. Eye position influence on the parieto-occipital area PO (V6) of the macaque monkey. *European Journal of Neuroscience*, 7, 2486-2501.
- GENOVESIO, A., TSUJIMOTO, S. & WISE, S. P. 2006. Neuronal activity related to elapsed time in prefrontal cortex. *Journal of Neurophysiology*, 95, 3281-3285.
- GENOVESIO, A., TSUJIMOTO, S. & WISE, S. P. 2009. Feature- and Order-Based Timing Representations in the Frontal Cortex. *Neuron*, 63, 254-266.
- GEORGESON, M. A. & SULLIVAN, G. D. 1975. Contrast constancy - deblurring in human vision by spatial frequency channels. *Journal of Physiology-London*, 252, 627-656.
- GEPSHTEIN, S. & BANKS, M. S. 2003. Viewing Geometry Determines How Vision and Haptics Combine in Size Perception. *Current Biology*, 13, 483-488.
- GEPSHTEIN, S., BURGE, J., ERNST, M. O. & BANKS, M. S. 2005. The combination of vision and touch depends on spatial proximity. *Journal of Vision*, 5, 1013-1023.
- GESCHIEDER, G. 1985. *Psychophysics Method Theory and Application*. Second edition, Lawrence Erlbaum Associates: Hillsdale, New Jersey.
- GETTY, D. J. 1975. Discrimination of short temporal intervals - comparison of 2 models. *Perception & Psychophysics*, 18, 1-8.
- GETTY, D. J. 1976. Counting processes in human timing. *Perception & Psychophysics*, 20, 191-197.
- GHAZANFAR, A. A. & SCHROEDER, C. E. 2006. Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, 10, 278-285.

- GIARD, M. H. & PERONNET, F. 1999. Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, 11, 473-490.
- GIBBON, J. 1977. Scalar expectancy - Theory and Webers law in animal timing. *Psychological Review*, 84, 279-325.
- GIBBON, J., CHURCH, R. M. & MECK, W. H. 1984. Scalar timing in memory. *Annals of the New York Academy of Sciences*, 423, 52-77.
- GIBBON, J., MALAPANI, C., DALE, C. L. & GALLISTEL, C. R. 1997. Toward a neurobiology of temporal cognition: Advances and challenges. *Current Opinion in Neurobiology*, 7, 170-184.
- GIBSON, J. J. & RADNER, M. 1937. Adaptation after-effect and contrast in the perception of tilted lines I. Quantitative studies. *Journal of Experimental Psychology-General*, 20, 453-467.
- GOLDSTEIN, A. G. 1958. On the after-effects of the waterfall and spiral illusions. *The American journal of psychology*, 71, 608-609.
- GOLDSTONE, S. 1968. Production and reproduction of duration: intersensory comparisons. *Perceptual and Motor Skills*, 26, 755-60.
- GOLDSTONE, S. & LHAMON, W. T. 1974. Studies of auditory - visual differences in human time judgment 1. Sounds are judged longer than lights. *Perceptual and Motor Skills*, 39, 63-82.
- GOLDSTONE, S. & LHAMON, W. T. 1976. Signal pulse rate and judged duration. *Perceptual and Motor Skills*, 42, 655-661.
- GOODFELLOW, L. D. 1934. An empirical comparison of audition, vision and touch in the discrimination of short intervals of time. *American Journal of Psychology*, 46, 243-258.
- GOOLER, D. M. & FENG, A. S. 1992. Temporal Coding in the Frog Auditory Midbrain - the Influence of Duration and Rise-Fall Time on the Processing of Complex Amplitude-Modulated Stimuli. *Journal of Neurophysiology*, 67, 1-22.
- GRANTHAM, D. W. 1995. Spatial hearing and related phenomena. In: MOORE, B. C. J. (ed.) *Hearing*. London: Academic Press.
- GRILL-SPECTOR, K., HENSON, R. & MARTIN, A. 2006. Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10, 14-23.
- GRILLNER, S., WALLEN, P., BRODIN, L. & LANSNER, A. 1991. Neuronal network generating locomotor behaviour in lamprey - circuitry, transmitters, membrane-properties, and simulation. *Annual Review of Neuroscience*, 14, 169-199.
- GRONDIN, S. 2005. Overloading temporal memory. *Journal of Experimental Psychology-Human Perception and Performance*, 31, 869-879.
- GRONDIN, S., MEILLEUR-WELLS, G., OUELLETTE, C. & MACAR, F. 1998. Sensory effects on judgments of short time-intervals. *Psychological Research-Psychologische Forschung*, 61, 261-268.
- GRONDIN, S. & ROUSSEAU, R. 1991. Judging the relative duration of multimodal short empty time intervals. *Perception & Psychophysics*, 49, 245-256.
- GU, Y., ANGELAKI, D. E. & DEANGELIS, G. C. 2008. Neural correlates of multisensory cue integration in macaque MSTd. *Nature Neuroscience*, 11, 1201-1210.
- HACKETT, T. A. & KAAS, J. H. 2003. Auditory Processing in the Primate Brain. In: NELSON, M. G. R. J. (ed.) *Handbook of Psychology Vol. 3: Biological Psychology*. New York: John Wiley & Sons, Inc.
- HACKNEY, C. M. 1987. Anatomical features of the auditory pathway from cochlea to cortex. *British Medical Bulletin*, 43, 780-801.

- HAMMOND, P. 1978. Directional tuning of complex cells in area 17 of the feline visual-cortex. *Journal of Physiology-London*, 285, 479-491.
- HARRINGTON, D. L., BOYD, L. A., MAYER, A. R., SHELTRAW, D. M., LEE, R. R., HUANG, M. & RAO, S. M. 2004. Neural representation of interval encoding and decision making. *Brain Res Cogn Brain Res*, 21, 193-205.
- HARTCHER-O'BRIEN, A. E. A. 2011. Temporal Ventriloquism in a Purely Temporal Context. *Journal of Experimental Psychology*, 1-13.
- HARVEY, L. O. 1986. Efficient estimation of sensory thresholds. *Behavior Research Methods Instruments & Computers*, 18, 623-632.
- HAZELTINE, E., HELMUTH, L. L. & IVRY, R. B. 1997. Neural mechanisms of timing. *Trends in Cognitive Sciences*, 1, 163-169.
- HE, J. F., HASHIKAWA, T., OJIMA, H. & KINOCHI, Y. 1997. Temporal integration and duration tuning in the dorsal zone of cat auditory cortex. *Journal of Neuroscience*, 17, 2615-2625.
- HELBIG, H. & ERNST, M. 2008. Visual-haptic cue weighting is independent of modality-specific attention. *Journal of Vision*, 8.
- HELBIG, H. B., ERNST, M. O., RICCIARDI, E., PIETRINI, P., THIELSCHER, A., MAYER, K. M., SCHULTZ, J. & NOPPENY, U. 2012. The neural mechanisms of reliability weighted integration of shape information from vision and touch. *NeuroImage*, 60, 1063-1072.
- HELLSTROM, A. & ALMKVIST, O. 1997. Tone duration discrimination in demented, memory-impaired, and healthy elderly. *Dementia and Geriatric Cognitive Disorders*, 8, 49-54.
- HENSON, R. N. A. & RUGG, M. D. 2003. Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia*, 41, 263-270.
- HERON, J., AEN-STOCKDALE, C., HOTCHKISS, J., ROACH, N. W., MCGRAW, P. V. & WHITAKER, D. 2012. Duration channels mediate human time perception. *Proceedings of the Royal Society B-Biological Sciences*, 279, 690-698.
- HERON, J., WHITAKER, D. & MCGRAW, P. V. 2004. Sensory uncertainty governs the extent of audio-visual interaction. *Vision Research*, 44, 2875-84.
- HIKOSAKA, O., MIYAUCHI, S. & SHIMOJO, S. 1993. Focal visual attention produces illusory temporal-order and motion sensation. *Vision Research*, 33, 1219-1240.
- HINTON, S. C. & MECK, W. H. 1997. The 'internal clocks' of circadian and interval timing. *Endeavour*, 21, 3-8.
- HORTON, J. C. & HUBEL, D. H. 1981. Regular patchy distribution of cytochrome oxidase staining in primary visual cortex of macaque monkey. *Nature*, 292, 762-764.
- HUBEL, D. H. & WIESEL, T. N. 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160, 106-54.
- striate cortex. *Journal of physiology*, 195, 215-243.
- HUBEL, D. H. & WIESEL, T. N. 1974. Sequence regularity and geometry of orientation columns in monkey striate cortex. *Journal of Comparative Neurology*, 158, 267-294.
- HUGHES, J., LISHMAN, J. R. & PARKER, D. M. 1992. Apparent duration and spatial structure. *Perception & Psychophysics*, 52, 222-230.
- IRVINE, D. R. F. & MALMIERCA, M. S. 2005. Auditory spectral processing: An overview. *Auditory Spectral Processing*, 70, 1-6.
- IVRY, R. B. 1996. The representation of temporal information in perception and motor control. *Current Opinion in Neurobiology*, 6, 851-857.
- 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 and Performance*, 21, 3-18.

- IVRY, R. B. & SCHLERF, J. E. 2008. Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences*, 12, 273-280.
- IVRY, R. B. & SPENCER, R. M. C. 2004. The neural representation of time. *Current Opinion in Neurobiology*, 14, 225-232.
- JACOBS, R. A. 1999. Optimal integration of texture and motion cues to depth. *Vision Research*, 39, 3621-3629.
- JAMIESON, D. G. & PETRUSIC, W. M. 1975. Presentation order effects in duration discrimination. *Perception & Psychophysics*, 17, 197-202.
- JANTZEN, K. J., STEINBERG, F. L. & KELSO, J. A. S. 2005. Functional MRI reveals the existence of modality and coordination-dependent timing networks. *Neuroimage*, 25, 1031-1042.
- JAZAYERI, M. & MOVSHON, J. A. 2006. Optimal representation of sensory information by neural populations. *Nature Neuroscience*, 9, 690-696.
- JAZAYERI, M. & SHADLEN, M. N. 2010. Temporal context calibrates interval timing. *Nature Neuroscience*, 13, 1020-1026.
- JIANG, W. & STEIN, B. E. 2003. Cortex controls multisensory depression in superior colliculus. *Journal of Neurophysiology*, 90, 2123-2135.
- JOHNSTON, A., ARNOLD, D. H. & NISHIDA, S. 2006. Spatially localized distortions of event time. *Current Biology*, 16, 472-479.
- JOUSMAKI, V. & HARI, R. 1998. Parchment-skin illusion: sound-biased touch. *Current Biology*, 8, R190-R190.
- KADUNCE, D. C., VAUGHAN, J. W., WALLACE, M. T. & STEIN, B. E. 2001. The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. *Experimental Brain Research*, 139, 303-310.
- KANAI, R., PAFFEN, C., HOGENDOORN, H. & VERSTRATEN, F. 2006. Time dilation in dynamic visual display. *Journal of Vision*, 6, 1421-1430.
- KANEKO, S. & MURAKAMI, I. 2009. Perceived duration of visual motion increases with speed. *Journal of Vision*, 9, 1-12.
- KARMARKAR, U. R. & BUONOMANO, D. V. 2007. Timing in the absence of clocks: Encoding time in neural network states. *Neuron*, 53, 427-438.
- KASHINO, M. & NISHIDA, S. 1998. Adaptation in the processing of interaural time differences revealed by the auditory localization aftereffect. *Journal of the Acoustical Society of America*, 103, 3597-3604.
- 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, 173-191.
- KLEIN, S. A. 2001. Measuring, estimating, and understanding the psychometric function: A commentary. *Perception & Psychophysics*, 63, 1421-1455.
- KLINK, P. C., MONTIJN, J. S. & VAN WEZEL, R. J. A. 2011. Crossmodal duration perception involves perceptual grouping, temporal ventriloquism, and variable internal clock rates. *Attention Perception & Psychophysics*, 73, 219-236.
- KOHN, A. 2007. Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology*, 97, 3155-3164.
- KORDING, K. P., BEIERHOLM, U., MA, W. J., QUARTZ, S., TENENBAUM, J. B. & SHAMS, L. 2007. Causal Inference in Multisensory Perception. *PLoS ONE*, 2, 943.
- KOVACS, G., CZIRAKI, C., VIDNYANSZKY, Z., SCHWEINBERGER, S. R. & GREENLEE, M. W. 2008. Position-specific and position-invariant face aftereffects reflect the adaptation of different cortical areas. *Neuroimage*, 43, 156-164.

- KOVACS, G., ZIMMER, M., HARZA, I. & VIDNYANSZKY, Z. 2007. Adaptation duration affects the spatial selectivity of facial aftereffects. *Vision Research*, 47, 3141-3149.
- 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, 291-305.
- LAURIENTI, P. J., BURDETTE, J. H., WALLACE, M. T., YEN, Y. F., FIELD, A. S. & STEIN, B. E. 2002. Deactivation of sensory-specific cortex by cross-modal stimuli. *Journal of Cognitive Neuroscience*, 14, 420-429.
- LEARY, C. J., EDWARDS, C. J. & ROSE, G. J. 2008. Midbrain auditory neurons integrate excitation and inhibition to generate duration selectivity: An in vivo whole-cell patch study in anurans. *Journal of Neuroscience*, 28, 5481-5493.
- LEBEDEV, M. A., O'DOHERTY, J. E. & NICOLELIS, M. A. L. 2008. Decoding of temporal intervals from cortical ensemble activity. *Journal of Neurophysiology*, 99, 166-186.
- LENNIE, P. 1998. Single units and visual cortical organization. *Perception*, 27, 889-935.
- LEON, M. I. & SHADLEN, M. N. 2003. Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron*, 38, 317-327.
- LEVAY, S., HUBEL, D. H. & WIESEL, T. N. 1975. Pattern of ocular dominance columns in macaque visual cortex revealed by a reduced silver stain. *Journal of Comparative Neurology*, 159, 559-575.
- LEVINSON, E. & SEKULER, R. 1976. Adaptation alters perceived direction of motion. *Vision Research*, 16, 779-781.
- LEVITT, H. 1971. Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49, 467.
- LEWIS, P. A. & MIAL, R. C. 2003. Distinct systems for automatic and cognitively controlled time measurement: evidence from neuroimaging. *Current Opinion in Neurobiology*, 13, 250-255.
- LEY, I., HAGGARD, P. & YARROW, K. 2009. Optimal Integration of Auditory and Vibrotactile Information for Judgments of Temporal Order. *Journal of Experimental Psychology-Human Perception and Performance*, 35, 1005-1019.
- LITTLE, A. C., DEBRUINE, L. M. & JONES, B. C. 2011. Category-contingent face adaptation for novel colour categories: Contingent effects are seen only after social or meaningful labelling. *Cognition*, 118, 116-122.
- LIVINGSTONE, M. S. & HUBEL, D. H. 1984. Specificity of intrinsic connections in primate visual cortex. *Journal of Neuroscience*, 4, 2830-2835.
- LIVINGSTONE, M. S. & HUBEL, D. H. 1988. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, 240, 740-749.
- LONG, G. M. & BEATON, R. J. 1980. The effects of spatial-frequency and target type on perceived duration. *Perception & Psychophysics*, 28, 413-421.
- LONG, G. M. & SAKITT, B. 1981. Differences between flicker and non-flicker persistence tasks-the effects of luminance and the number of cycles in a grating target. *Vision Research*, 21, 1387-1393.
- LUNENBURGER, L., KLEISER, R., STUPHORN, V., MILLER, L. E. & HOFFMANN, K. P. 2001. A possible role of the superior colliculus in eye-hand coordination. In: CASANOVA, C. & PTITO, M. (eds.) *Vision: From Neurons to Cognition*. Elsevier.
- MACALUSO, E., FRITH, C. D. & DRIVER, J. 2000. Modulation of human visual cortex by crossmodal spatial attention. *Science*, 289, 1206-1208.
- MACALUSO, E., GEORGE, N., DOLAN, R., SPENCE, C. & DRIVER, J. 2004. Spatial and temporal factors during processing of audiovisual speech: a PET study. *Neuroimage*, 21, 725-732.

- MACAR, F., GRONDIN, S. & CASINI, L. 1994. Controlled attention sharing influences time estimation. *Memory & Cognition*, 22, 673-686.
- MAMASSIAN, P., LANDY, M. & MALONEY, L. T. 2002. Bayesian modelling of visual perception. In: RAO, R. P. N., OLSHAUSEN, B. A. & LEWICKI, M. S. (eds.) *Probabilistic Models of the Brain: Perception and Neural Function*. Cambridge, MA: MIT Press.
- MARCHANT, J. L., RUFF, C. C. & DRIVER, J. 2012. Audiovisual synchrony enhances BOLD responses in a brain network including multisensory STS while also enhancing target-detection performance for both modalities. *Human Brain Mapping*, 33, 1212-1224.
- MATELL, M. S. & MECK, W. H. 2004. Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Cognitive Brain Research*, 21, 139-170.
- MATHER, G. 1980. The movement aftereffect and a distribution-shift model for coding the direction of visual movement. *Perception*, 9, 379-392.
- MAUK, M. D. & BUONOMANO, D. V. 2004. The neural basis of temporal processing. *Annual Review of Neuroscience*, 27, 307-340.
- MAUK, M. D. & DONEGAN, N. H. 1997. A model of Pavlovian eyelid conditioning based on the synaptic organization of the cerebellum. *Learning & Memory*, 4, 130-158.
- MCADAMS, C. J. & MAUNSELL, J. H. R. 1999. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience*, 19, 431-441.
- MCDONALD, J. J., TEDER-SALEJARVI, W. A., DI RUSSO, F. & HILLYARD, S. A. 2003. Neural substrates of perceptual enhancement by cross-modal spatial attention. *Journal of Cognitive Neuroscience*, 15, 10-19.
- MCGURK, H. & MACDONALD, J. 1976. Hearing lips and seeing voices. *Nature*, 264, 746-748.
- MECK, W. H. 1996. Neuropharmacology of timing and time perception. *Cognitive Brain Research*, 3, 227-242.
- MECK, W. H., PENNEY, T. B. & POUTHAS, V. 2008. Cortico-striatal representation of time in animals and humans. *Current Opinion in Neurobiology*, 18, 145-152.
- MEEGAN, D. V., ASLIN, R. N. & JACOBS, R. A. 2000. Motor timing learned without motor training. *Nature Neuroscience*, 3, 860-862.
- MEESE, T. 2002. Spatial Vision. In: ROBERTS, D. (ed.) *Signals and perception: the fundamentals of human sensation*. New York: Pargrave Macmillan.
- MERCHANT, H., ZARCO, W., PEREZ, O., PRADO, L. & BARTOLO, R. 2011. Measuring time with different neural chronometers during a synchronization-continuation task. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 19784-19789.
- MERCHANT, H., ZARCO, W. & PRADO, L. 2008. Do we have a common mechanism for measuring time in the hundreds of millisecond range? Evidence from multiple-interval timing tasks. *Journal of Neurophysiology*, 99, 939-949.
- MEREDITH, M. A. 2002. On the neuronal basis for multisensory convergence: a brief overview. *Cognitive Brain Research*, 14, 31-40.
- MEREDITH, M. A. & STEIN, B. E. 1986. Visual, auditory and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, 56, 640-662.
- MEYER, G. E. & MAGUIRE, W. M. 1977. Spatial-frequency and mediation of short-term visual storage. *Science*, 198, 524-525.
- MIALL, C. 1996. Models of neural timing. In: PASTOR, M. A. A., J. (ed.) *Time, Internal clocks and movement*. Elsevier Science
- MISHKIN, M., UNGERLEIDER, L. G. & MACKO, K. A. 1983. Object vision and spatial vision - 2 cortical pathways. *Trends in Neurosciences*, 6, 414-417.

- MIYAZAKI, M., NOZAKI, D. & NAKAJIMA, Y. 2005. Testing Bayesian models of human coincidence timing. *Journal of Neurophysiology*, 94, 395-399.
- MONTASER-KOUHSARI, L. & RAJIMEHR, R. 2004. Attentional modulation of adaptation to illusory lines. *Journal of Vision*, 4, 434-444.
- MOORE, B. C. J. 2005. Basic Auditory Processes. In: GOLDSTEIN, E. B. (ed.) *Blackwell Handbook of Sensation and Perception*. Malden, MA: Blackwell Publishing Ltd.
- MORA, E. C. & KOSSL, M. 2004. Ambiguities in sound-duration selectivity by neurons in the inferior colliculus of the bat *Molossus molossus* from Cuba. *Journal of Neurophysiology*, 91, 2215-2226.
- MORAN, J. & DESIMONE, R. 1985. Selective attention gates visual processing in the extrastriate cortex. *Science*, 229, 782-784.
- MORRONE, M. C., ROSS, J. & BURR, D. 2005. Saccadic eye movements cause compression of time as well as space. *Nature Neuroscience*, 8, 950-954.
- MUELLER, K. M., WILKE, M. & LEOPOLD, D. A. 2009. Visual adaptation to convexity in macaque area V4. *Neuroscience*, 161, 655-662.
- NACHMIAS, J. 2006. The role of virtual standards in visual discrimination. *Vision Research*, 46, 2456-2464.
- 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, 1559-1570.
- NAKAYAMA, K. & MACKEBEN, M. 1989. Sustained and transient components of focal visual attention. *Vision Research*, 29, 1631-1647.
- NATH, A. R. & BEAUCHAMP, M. S. 2011. Dynamic Changes in Superior Temporal Sulcus Connectivity during Perception of Noisy Audiovisual Speech. *Journal of Neuroscience*, 31, 1704-1714.
- NEW, J. J. & SCHOLL, B. J. 2009. Subjective time dilation: Spatially local, object-based, or a global visual experience? *Journal of Vision*, 9, 1-11.
- NISLY, S. J. & WASSERMAN, G. S. 1989. Intensity dependence of perceived duration - Data, theories and neural integration. *Psychological Bulletin*, 106, 483-496.
- NOESSELT, T., RIEGER, J. W., SCHOENFELD, M. A., KANOWSKI, M., HINRICHS, H., HEINZE, H.-J. & DRIVER, J. 2007. Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *Journal of Neuroscience*, 27, 11431-11441.
- NORTON, T., CORLISS, D. & BAILEY, J. 2002. *The Psychophysical Measurement of Visual Function*, Albuquerque: Richmond Products Inc.
- PALMER, S. 1999. *Vision Science*, Cambridge MA: MIT Press.
- PARISE, C. V., SPENCE, C. & ERNST, M. O. 2012. When Correlation Implies Causation in Multisensory Integration. *Current Biology*, 22, 46-49.
- PARIYADATH, V. & EAGLEMAN, D. 2007. The effect of predictability on subjective duration. *PLoS One*, 2, e1264.
- PARIYADATH, V. & EAGLEMAN, D. M. 2008. Brief subjective durations contract with repetition. *Journal of Vision*, 8, 1-6.
- PASALAR, S., RO, T. & BEAUCHAMP, M. S. 2010. TMS of posterior parietal cortex disrupts visual tactile multisensory integration. *European Journal of Neuroscience*, 31, 1783-1790.
- PATTERSON, R. D. 1976. Auditory filter shapes derived with noise stimuli. *Journal of the Acoustical Society of America*, 59, 640-654.

- PEREZ-GONZALEZ, D., MALMIERCA, M. S., MOORE, J. M., HERNANDEZ, O. & COVEY, E. 2006. Duration selective neurons in the inferior colliculus of the rat: Topographic distribution and relation of duration sensitivity to other response properties. *Journal of Neurophysiology*, 95, 823-836.
- POTTER, H. D. 1965. Patterns of acoustically evoked discharges of neurons in the mesencephalon of the bullfrog. *Journal of Neurophysiology*, 28, 1155-1184.
- POUTHAS, V. & PERBAL, S. 2004. Time perception depends on accurate clock mechanisms as well as unimpaired attention and memory processes. *Acta Neurobiologiae Experimentalis*, 64, 367-385.
- PRIEBE, N. J., LISBERGER, S. G. & MOVSHON, J. A. 2006. Tuning for spatiotemporal frequency and speed in directionally selective neurons of macaque striate cortex. *Journal of Neuroscience*, 26, 2941-2950.
- PRINZMETAL, W., AMIRI, H., ALLEN, K. & EDWARDS, T. 1998. Phenomenology of attention: 1. Color, location, orientation, and spatial frequency. *Journal of Experimental Psychology-Human Perception and Performance*, 24, 261-282.
- RAKITIN, B. C., GIBBON, J., PENNEY, T. B., MALAPANI, C., HINTON, S. C. & MECK, W. H. 1998. Scalar expectancy theory and peak-interval timing in humans. *Journal of Experimental Psychology-Animal Behavior Processes*, 24, 15-33.
- RAMMSAYER, T. H. 1999. Neuropharmacological evidence for different timing mechanisms in humans. *Quarterly Journal of Experimental Psychology Section B - Comparative and Physiological Psychology*, 52, 273-286.
- RAMMSAYER, T. H., HENNIG, J., HAAG, A. & LANGE, N. 2001. Effects of noradrenergic activity on temporal information processing in humans. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology*, 54, 247-258.
- RAMMSAYER, T. H. & SKRANDIES, W. 1998. Stimulus characteristics and temporal information processing: Psychophysical and electrophysiological data. *Journal of Psychophysiology*, 12, 1-12.
- REGAN, D. 1989. *Human brain electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine.*, New York: Elsevier.
- REGAN, D. & TANSLEY, B. W. 1979. Selective adaptation to frequency-modulated tones: evidence for an information - processing channel selective to frequency changes. *Journal of the Acoustical Society of America*, 65, 1249-1257.
- REMLINGTON, R. W., JOHNSTON, J. C. & YANTIS, S. 1992. Involuntary attentional capture by abrupt onsets. *Perception & Psychophysics*, 51, 279-290.
- ROACH, N., HERON, J. & MCGRAW, P. 2006. Resolving multisensory conflict: a strategy for balancing the costs and benefits of audio-visual integration. *Proceedings of The Royal Society B-Biological Sciences*, 273, 2159-2168.
- ROACH, N. & MCGRAW, P. 2009. Dynamics of spatial distortions reveal multiple time scales of motion adaptation. *Journal of Neurophysiology*, 102, 3619-3626.
- ROCKLAND, K. S. & OJIMA, H. 2003. Multisensory convergence in calcarine visual areas in macaque monkey. *International Journal of Psychophysiology*, 50, 19-26.
- ROSE, D. & SUMMERS, J. 1995. Duration illusions in a train of visual-stimuli. *Perception*, 24, 1177-1187.
- ROTSCHTEIN, P., HENSON, R. N. A., TREVES, A., DRIVER, J. & DOLAN, R. J. 2005. Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature Neuroscience*, 8, 107-113.
- ROUSSEAU, R. & KRISTOFF, A. 1973. Discrimination of bimodal temporal gaps. *Bulletin of the Psychonomic Society*, 1, 115-116.

- ROUSSEAU, R., POIRIER, J. & LEMYRE, L. 1983. Duration discrimination of empty time intervals marked by intermodal pulses. *Perception & Psychophysics*, 34, 541-548.
- ROUX, S., COULMANCE, M. & RIEHLE, A. 2003. Context-related representation of timing processes in monkey motor cortex. *European Journal of Neuroscience*, 18, 1011-1016.
- RUBIA, K. & SMITH, A. 2004. The neural correlates of cognitive time management: a review. *Acta Neurobiologiae Experimentalis*, 64, 329-340.
- RUSAK, B. & ZUCKER, I. 1979. Neural regulation of circadian rhythms. *Physiological Reviews*, 59, 449-526.
- SACHS, M. B., NACHMIAS, J. & ROBSON, J. G. 1971. Spatial-frequency channels in human vision. *Journal of the Optical Society of America*, 61, 1176-1186.
- SAYEGH, R., AUBIE, B. & FAURE, P. A. 2011. Duration tuning in the auditory midbrain of echolocating and non-echolocating vertebrates. *Journal of Comparative Physiology a-Neuroethology Sensory Neural and Behavioral Physiology*, 197, 571-583.
- SCHIFFMAN, H. R. & BOBKO, D. J. 1977. Role of number and familiarity of stimuli in perception of brief temporal intervals. *American Journal of Psychology*, 90, 85-93.
- SCHIRMER, A. 2004. Timing speech: a review of lesion and neuroimaging findings. *Cognitive Brain Research*, 21, 269-287.
- SCHRATER, P. R. & SIMONCELLI, E. P. 1998. Local velocity representation: evidence from motion adaptation. *Vision Research*, 38, 3899-3912.
- SCHWARTZ, O., HSU, A. & DAYAN, P. 2007. Space and time in visual context (Vol 8 pg 522, 2007). *Nature Reviews Neuroscience*, 8, 883-883.
- SEIFRIED, T. & ULRICH, R. 2010. Does the asymmetry effect inflate the temporal expansion of odd stimuli? *Psychological Research-Psychologische Forschung*, 74, 90-98.
- SERIES, P., STOCKER, A. A. & SIMONCELLI, E. P. 2009. Is the homunculus "aware" of sensory adaptation? *Neural Computation*, 21, 3271-3304.
- SHAMS, L., KAMITANI, Y., & SHIMOJO, S. 2000. Illusory flashing visual percept induced by sound. *Investigative Ophthalmology & Visual Science* 41 (4) S229-S229
- SHAMS, L., KAMITANI, Y., THOMPSON, S. & SHIMOJO, S. 2001. Sound alters visual evoked potentials in humans. *Neuroreport*, 12, 3849-3852.
- SHINOMOTO, S., OMI, T., MITA, A., MUSHIAKE, H., SHIMA, K., MATSUZAKA, Y. & TANJI, J. 2011. Deciphering elapsed time and predicting action timing from neuronal population signals. *Frontiers In Computational Neuroscience*, 5, 29.
- SPENCE, C. & SQUIRE, S. 2003. Multisensory integration: Maintaining the perception of synchrony. *Current Biology*, 13, R519-R521.
- SPENCER, R. M. C. & IVRY, R. B. 2005. Comparison of patients with Parkinson's disease or cerebellar lesions in the production of periodic movements involving event-based or emergent timing. *Brain and Cognition*, 58, 84-93.
- STEIN, B. E. & STANFORD, T. R. 2008. Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, 9, 255-266.
- STETSON, C., FIESTA, M. P. & EAGLEMAN, D. M. 2007. Does time really slow down during a frightening event? *Plos One*, 2, e 1295.
- STRASBURGER, H. 2001. Converting between measures of slope of the psychometric function. *Perception & Psychophysics*, 63, 1348-1355.
- SUN, J. & PERONA, P. 1998. Where is the sun? *Nature Neuroscience*, 1, 183-184.
- SUZUKI, S. 2001. Attention-dependent brief adaptation to contour orientation: a high-level aftereffect for convexity? *Vision Research*, 41, 3883-3902.
- SVAETICHIN, G. 1956. Spectral response curves from single cones. *Acta physiologica Scandinavica. Supplementum*, 39, 17-46.

- SWINDALE, N. V. 1996. The development of topography in the visual cortex: A review of models. *Network-Computation in Neural Systems*, 7, 161-247.
- TALSMA, D., SENKOWSKI, D., SOTO-FARACO, S. & WOLDORFF, M. G. 2010. The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, 14, 400-410.
- TALSMA, D. & WOLDORFF, M. G. 2005. Selective attention and multisensory integration: Multiple phases of effects on the evoked brain activity. *Journal of Cognitive Neuroscience*, 17, 1098-1114.
- TAYLOR, M. M., FORBES, S. M. & CREELMAN, C. D. 1983. PEST reduces bias in forced choice psychophysics. *Journal of the Acoustical Society of America*, 74, 1367-1374.
- TEDER-SALEJARVI, W. A., MCDONALD, J. J., DI RUSSO, F. & HILLYARD, S. A. 2002. An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Brain research. Cognitive brain research*, 14, 106-14.
- TERAO, M., WATANABE, J., YAGI, A. & NISHIDA, S. 2008. Reduction of stimulus visibility compresses apparent time intervals. *Nature Neuroscience*, 11, 541-542.
- THOMAS, E. A. C. & WEAVER, W. B. 1975. Cognitive processing and time perception. *Perception & Psychophysics*, 17, 363-367.
- TOLHURST, D. J. & THOMPSON, I. D. 1981. On the variety of spatio-frequency selectivities shown by neurons in Area 17 of the cat. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 213, 183-199.
- TOOTELL, R. B. H., SILVERMAN, M. S., SWITKES, E. & DEVALOIS, R. L. 1982. Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science*, 218, 902-904.
- TOVEE, M. J. 1996. *An introduction to the visual system*, Cambridge University Press.
- TREISMAN, M. 1963. Temporal discrimination and the indifference interval-implications for a model of the "internal clock". *Psychological Monographs*, 77, 1-31.
- TSE, P. U., INTRILIGATOR, J., RIVEST, J. & CAVANAGH, P. 2004. Attention and the subjective expansion of time. *Perception & Psychophysics*, 66, 1171-1189.
- ULRICH, R., NITSCHKE, J. & RAMMSAYER, T. 2006(a). Crossmodal temporal discrimination: Assessing the predictions of a general pacemaker-counter model. *Perception & Psychophysics*, 68, 1140-1152.
- ULRICH, R., NITSCHKE, J. & RAMMSAYER, T. 2006(b). Perceived duration of expected and unexpected stimuli. *Psychological Research-Psychologische Forschung*, 70, 77-87.
- ULRICH, R. & VORBERG, D. 2009. Estimating the difference limen in 2AFC tasks: Pitfalls and improved estimators. *Attention, Perception and Psychophysics*, 71, 1219-27.
- VAN EE, R., VAN BOXTEL, J. J. A., PARKER, A. L. & ALAIS, D. 2009. Multisensory Congruency as a Mechanism for Attentional Control over Perceptual Selection. *Journal of Neuroscience*, 29, 11641-11649.
- VAN ESSEN, D. C., ANDERSON, C. H. & FELLEMAN, D. J. 1992. Information processing in the primate visual system: an integrated perspective. *Science*, 255, 419-423.
- VAN WASSENHOVE, V., BUONOMANO, D. V., SHIMOJO, S. & SHAMS, L. 2008. Distortions of subjective time perception within and across senses. *PLoS One*, 3, e1437.
- VATAKIS, A. & SPENCE, C. 2006. Audiovisual synchrony perception for music, speech, and object actions. *Brain Research*, 1111, 134-142.
- VATAKIS, A. & SPENCE, C. 2007. Crossmodal binding: Evaluating the "unity assumption" using audiovisual speech stimuli. *Perception & Psychophysics*, 69, 744-756.
- VROOMEN, J., BERTELSON, P. & DE GELDER, B. 2001. The ventriloquist effect does not depend on the direction of automatic visual attention. *Perception & Psychophysics*, 63, 651-9.

- WALKER, J. T., IRION, A. L. & GORDON, D. G. 1981(a). Simple and contingent aftereffects of perceived duration in vision and audition. *Perception & Psychophysics*, 29, 475-486.
- WALKER, J. T. & SCOTT, K. J. 1981(b). Auditory-visual conflicts in the perceived duration of lights tones and gaps. *Journal of Experimental Psychology-Human Perception and Performance*, 7, 1327-1339.
- WALLACE, M. T. & STEIN, B. E. 1997. Development of multisensory neurons and multisensory integration in cat superior colliculus. *Journal of Neuroscience*, 17, 2429-2444.
- WANG, J., VAN WIJHE, R., CHEN, Z. & YIN, S. K. 2006. Is duration tuning a transient process in the inferior colliculus of guinea pigs? *Brain Research*, 1114, 63-74.
- WARM, J. S., STUTZ, R. M. & VASSOLO, P. A. 1975. Intermodal transfer in temporal discrimination. *Perception & Psychophysics*, 18, 281-286.
- WATSON, A. B. & PELLI, D. G. 1983. Quest- A Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33, 113-120.
- 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-Comparative and Physiological Psychology*, 51, 97-120.
- WEARDEN, J. H. & LEJEUNE, H. 2008. Scalar properties in human timing: Conformity and violations. *Quarterly Journal of Experimental Psychology*, 61, 569-587.
- WEARDEN, J. H., NORTON, R., MARTIN, S. & MONTFORD-BEBB, O. 2007. Internal clock processes and the filled-duration illusion. *Journal of Experimental Psychology-Human Perception and Performance*, 33, 716-729.
- WEARDEN, J. H., TODD, N. P. M. & JONES, L. A. 2006. When do auditory/visual differences in duration judgements occur? *Quarterly Journal of Experimental Psychology*, 59, 1709-1724.
- WELCH, R. B., DUTTONHURT, L. D. & WARREN, D. H. 1986. Contributions of audition and vision to temporal rate perception. *Perception & Psychophysics*, 39, 294-300.
- WELCH, R. B. & WARREN, D. H. 1980(a). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88, 638-667.
- WELCH, R. B., WARREN, D. H., WITH, R. & WAIT, J. S. 1980(b). Visual capture - The effects of compellingness and the assumption of unity. *Bulletin of the Psychonomic Society*, 16, 168-168.
- WERBLIN, F. S. & DOWLING, J. E. 1969. Organization of the retina of the mudpuppy, *Necturus maculosus*. Intercellular recording. *Journal of Neurophysiology* 32, 339-355.
- WERNER, S. & NOPPENY, U. 2010. Superadditive Responses in Superior Temporal Sulcus Predict Audiovisual Benefits in Object Categorization. *Cerebral Cortex*, 20, 1829-1842.
- WESTHEIMER, G. 1984. Spatial vision. *Annual Review of Psychology*, 35, 201-226.
- WESTHEIMER, G. 1999. Discrimination of short time intervals by the human observer. *Experimental Brain Research*, 129, 121-126.
- WITTMANN, M. 2009. The inner experience of time. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 364, 1955-1967.
- WOLF, F. & GEISEL, T. 1998. Spontaneous pinwheel annihilation during visual development. *Nature*, 395, 73-78.
- WONG-RILEY, M. 1979. Changes in the visual system of monocularly sutured or enucleated cats demonstrable with cytochrome oxidase histochemistry. *Brain Research*, 171, 11-28.
- WOODROW, H. 1928. Behavior with respect to short temporal stimulus forms. *Journal of Experimental Psychology*, 11, 167-193.

- WRIGHT, T. M., PELPHREY, K. A., ALLISON, T., MCKEOWN, M. J. & MCCARTHY, G. 2003. Polysensory interactions along lateral temporal regions evoked by audiovisual speech. *Cerebral Cortex*, 13, 1034-1043.
- XIA, Y. F., QI, Z. H. & SHEN, J. X. 2000. Neural representation of sound duration in the inferior colliculus of the mouse. *Acta Oto-Laryngologica*, 120, 638-643.
- YIN, S. K., CHEN, Z. N., YU, D. Z., FENG, Y. M. & WANG, J. 2008. Local inhibition shapes duration tuning in the inferior colliculus of guinea pigs. *Hearing Research*, 237, 32-48.
- YUMOTO, N., LU, X., HENRY, T. R., MIYACHI, S., NAMBU, A., FUKAI, T. & TAKADA, M. 2011. A neural correlate of the processing of multi-second time intervals in primate prefrontal cortex. *PloS One*, 6, e19168.
- ZAKAY, D. 1993. Time estimation methods-Do they influence prospective duration estimates? *Perception*, 22, 91-101.
- ZAKAY, D. 1998. Attention allocation policy influences prospective timing. *Psychonomic Bulletin & Review*, 5, 114-118.
- ZAKAY, D. & BLOCK, R. A. 1997. Temporal cognition. *Current Directions in Psychological Science*, 6, 12-16.
- ZAMPINI, M., SHORE, D. I. & SPENCE, C. 2003. Audiovisual temporal order judgments. *Experimental Brain Research*, 152, 198-210.
- ZEKI, S. M. 1978. Functional specialisation in the visual cortex of the rhesus monkey. *Nature*, 274, 423-428.
- ZELAZNIK, H. N., SPENCER, R. M. C., IVRY, R. B., BARIA, A., BLOOM, M., DOLANSKY, L., JUSTICE, S., PATTERSON, K. & WHETTER, E. 2005. Timing variability in circle drawing and tapping: Probing the relationship between event and emergent timing. *Journal of Motor Behavior*, 37, 395-403.
- ZWICKER, E. 1961. Subdivision of the audible frequency range into critical bands. *Journal of the Acoustical Society of America*, 33, 248.