The role of sensory history and stimulus context in human time perception

Adaptive and integrative distortions of perceived duration

Corinne FULCHER

Submitted for the Degree of Doctor of Philosophy

Faculty of Life Sciences

University of Bradford

2017

Abstract

Corinne Fulcher

The Role of Sensory History and Stimulus Context in Human Time Perception

Adaptive and Integrative Distortions of Perceived Duration

Keywords: Time perception, Duration, Human, Adaptation, Integration, Subsecond, Vision, Audition, Psychophysics, Modelling

This thesis documents a series of experiments designed to investigate the mechanisms subserving sub-second duration processing in humans. Firstly, duration aftereffects were generated by adapting to consistent duration information. If duration aftereffects represent encoding by neurons selective for both stimulus duration and non-temporal stimulus features, adapt-test changes in these features should prevent duration aftereffect generation. Stimulus characteristics were chosen which selectively target differing stages of the visual processing hierarchy. The duration aftereffect showed robust interocular transfer and could be generated using a stimulus whose duration was defined by stimuli invisible to monocular mechanisms, ruling out a pre-cortical locus. The aftereffects transferred across luminance-defined visual orientation and facial identity. Conversely, the duration encoding mechanism was selective for changes in the contrast-defined envelope size of a Gabor and showed broad spatial selectivity which scaled proportionally with adapting stimulus size. These findings are consistent with a second stage visual spatial mechanism that pools input across proportionally smaller,

spatially abutting filters. A final series of experiments investigated the pattern of interaction between concurrently presented cross-modal durations. When duration discrepancies were small, multisensory judgements were biased towards the modality with higher precision. However, when duration discrepancies were large, perceived duration was compressed by both longer and shorter durations from the opposite modality, irrespective of unimodal temporal reliability. Taken together, these experiments provide support for a duration encoding mechanism that is tied to mid-level visual spatial processing. Following this localised encoding, supramodal mechanisms then dictate the combination of duration information across the senses.

Acknowledgements

I wish to firstly thank Dr James Heron, for guiding me through the intricacies of human psychophysics and for his wealth of advice and support throughout my postgraduate studies. In addition, I am indebted to my collaborators and colleagues Professor David Whitaker, Dr Neil Roach, Dr Andrew Logan and Professor Paul McGraw for their professional expertise and help, particularly in regards to mathematics and computational modelling.

I am incredibly grateful to my fellow postgraduate colleagues in G39, for your continued friendship and support, for steering me through the twists and turns of vision research, but also for giving up your time to participate in experiments. Thank you in particular to Dr Samantha Strong, who not only helped me to negotiate Matlab, SPSS and a host of neuroimaging literature, but who also became a fantastic friend over the last four years.

Lastly I must thank my wonderful husband Jacko, and my parents Jen and Richard, for giving me unconditional love and support and the courage to undertake a PhD in the first place. You never stopped believing in me, and for that I will be eternally grateful.

Table of Contents

Table of Figures

Chapter 1: Introduction

Time is ubiquitous. Consciously or unconsciously it forms part of our everyday lives, and our ability to accurately process and utilise temporal signals allows us to successfully navigate the world. Humans and animals have evolved to deal with timescales over ten orders of magnitude, and time forms one of the only perceptual metrics to transcend all other sensory modalities. The formation of accurate temporal estimates can be used to guide our actions or behaviour, for example when catching a ball, a person needs to estimate how long it will take the ball to reach them and also reach out to catch the ball at the correct moment. This trade-off between perception (watching the ball's trajectory) and action (moving to catch the ball) has been shown to be statistically optimal, suggesting that humans have evolved to be so precisely aware of timing (both of themselves and external stimuli) that our behaviours reflect the least possible task variability (Faisal and Wolpert 2009). We can also process how much time has elapsed during an event, which we often describe as the event's duration. This is involved in understanding and generating speech, which is a particularly important element of daily human interaction. We can differentiate 'black bird' from 'blackbird' by processing the duration of the pause between each syllable. In fact some communication is entirely temporal, for example Morse code relies entirely on our ability to discriminate between dots (short duration tones) and dashes (long duration tones) (see Figure 1.1).

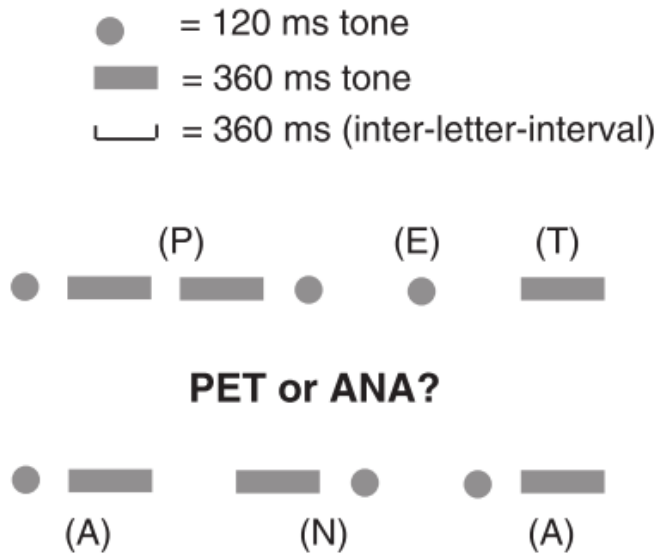

Figure 1.1: An example of morse code, a language built upon the temporal properties of auditory tones. In the above example, the 'dots' are short 120ms tones and the 'dashes' are longer 360ms tones. Differentiating between the letters PET and ANA requires an ability to process not only the tone durations, but also the gaps between the tones (the inter-letter-interval). Figure reprinted from (Hardy and Buonomano 2016) with permission from Elsevier.

Many forms of sensory information appear to have dedicated neural hardware allocated to their processing. Often these neurons are located in functionally specific and spatially localised brain regions, for example the processing of colour occurs in visual area V4 and visual motion is largely processed in area V5 (Zihl et al. 1983; Zeki et al. 1991). However, despite decades of research into human and non-human animal time perception evidence for a 'time specific' brain region(s) remains elusive. Whilst evidence from clinical populations has found that damage to certain areas of the brain can interfere with our performance in timing tasks (Pastor et al. 1992; Nichelli et al. 1996), impaired or loss of function in any one area of the brain does not appear to completely eradicate temporal processing in the way that damage to area V1 in the visual cortex might cause cortical blindness (Aldrich et al.

1987). This has led to suggestions that temporal information may be processed in an altogether different manner to other sensory information. Rather than using a dedicated system for timing (Figure 1.2a), processing may occur across linked networks of different brain regions (Figure 1.2b), or locally in numerous neural areas (Figure 1.2c).

Studies in the timing literature may be categorised by the type of temporal feature under investigation. These include, but are not limited to, studies of temporal rate / rhythm (e.g. Gebhard and Mowbray 1959; Becker and

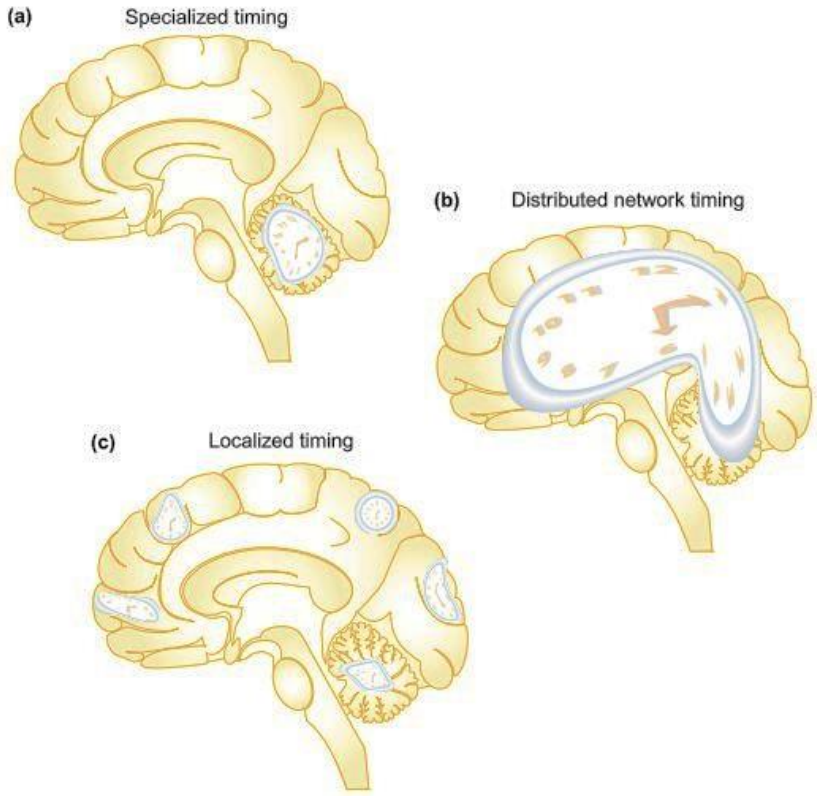

Figure 1.2: Three different theories for how time could be processed in the brain: **a)** an example of a dedicated system, represented here by the cerebellum, **b)** a distributed network across all regions of the brain, and **c)** localised timing where each area of the brain performs its own temporal processing either with a specific mechanism or through intrinsic processes. Figure reprinted from (Ivry and Spencer 2004a) with permission from Elsevier.

Rasmussen 2007), simultaneity / temporal order (e.g. Vroomen et al. 2004; Van de Burg et al. 2015) or event duration (Walker and Scott 1981; Grondin 1993; Penton-Voak et al. 1996).

Timing studies can also be divided into several categories based on the timescales that are required for various processes or actions (see Figure 1.3). Cycles of time refer to repeating cycles of activity over a given duration, such as the circadian rhythms. These are biological processes that are linked to the sleep/wake cycle and regulation of core body temperature (Zeiler and Hoyert 1989; Krauchi 2002). They operate over approximately 24 hours and are regulated by the 'circadian clock' found in the suprachiasmatic nucleus of the hypothalamus. Here a series of positive and negative feedback loops function in an oscillatory manner to maintain the rhythm (Jagota et al. 2000). The 'clock' of the suprachiasmatic nucleus (SCN) is synchronised with solar time through photosensitive retinal ganglion cells which directly project to the SCN, and relay information about light and dark cycles (Berson et al. 2002). A working visual system may also be necessary for this coordination, as studies into the circadian rhythms of totally blind subjects found that most participants had abnormal or 'free-running' rhythms (Miles et al. 1977; Sack et al. 1992). The presence of circadian rhythms is found in all organisms ranging from bacteria and plants to animals and humans. They allow the organism to adapt to the environment in which it lives, for example in humans the synchronisation of cell proliferation in epithelial tissue is linked to our circadian rhythm (Bjarnason et al. 1999).

Supra-second timing is concerned with seconds, minutes and hours, and is involved in decision making and conscious estimates of time. Sub-second

timing refers to durations in the millisecond range and as mentioned earlier has a role in speech recognition (Shannon et al. 1995), fine motor control, enabling for example, interception of a moving target (Port et al. 2001) and learning musical rhythms (Tillmann et al. 2011).

It has been queried whether our ability to time in the seconds and minutes range could be achieved by simple strategies such as counting, and certainly children as young as 6 years old have been known to use a spontaneous

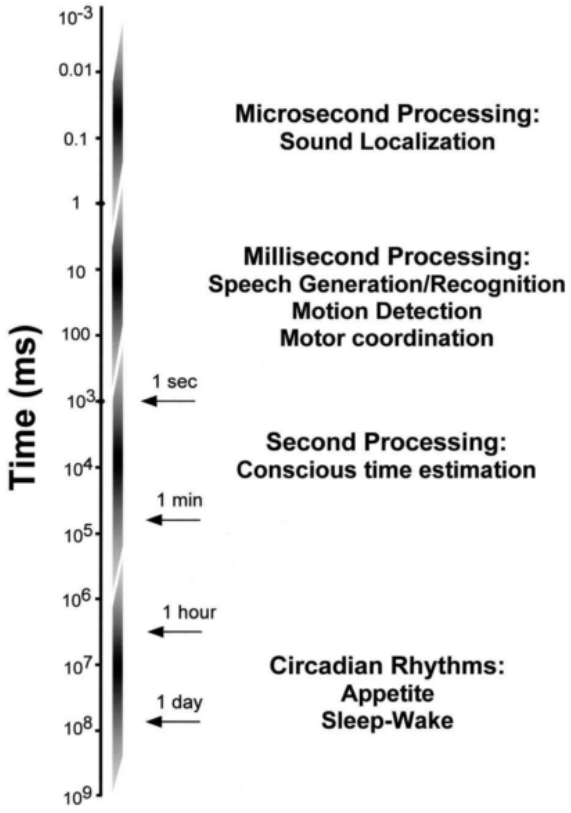

Figure 1.3: Divisions of temporal processing over ten orders of magnitude, ranging from microseconds to circadian rhythms. Figure reprinted from (Buonomano and Karmarkar 2002) with permission from SAGE Publications.

counting strategy to estimate durations (Espinosa-Fernandez et al. 2004). However behavioural evidence suggests that counting and timing may be two separate neural mechanisms (Hinton et al. 2004). Furthermore since

counting only appears to significantly improve performance for durations greater than one second (Grondin et al. 1999; Grondin et al. 2004), it cannot account for our ability to time durations in the millisecond range.

At the shortest end of the scale is microsecond processing, which has a role in sound localisation (Jeffress 1948).The sound originating from an external event may reach one ear before reaching the fellow ear, depending on its location in space. This results in microsecond interaural differences arising in the axonal conduction of the signals, which are then detected by binaural neurons in the superior olivery nucleus of the brainstem. Humans have shown remarkable accuracy for the spatial localisation of sounds with thresholds as low as 10μs (Klumpp and Eady 1956).

Therefore, relative to what we know the mechanisms involved in processing very long durations (e.g. circadian rhythms) and very short durations (e.g. microseconds), our understanding of the temporal processes involved in supra-second and sub-second timing is still unclear.

Incredibly our perception of time can deviate markedly from reality, seemingly with relatively little consequence to our wellbeing. As described anecdotally, our attention to the passage of time can result in the speeding up (“time flies when you’re having fun”) or slowing down (“a watched pot never boils”) of subjective time, which has also formed the basis of a large body of research within the timing literature (e.g. Cahoon and Edmonds 1980; Brown 1985; Coull et al. 2004). Furthermore, perceived duration can vary with sensory modality (Goldstone et al. 1959; Goldstone and Lhamon 1974; Wearden et al. 1998), non-temporal magnitude (Goldstone et al. 1978;

Brigner 1986; Xuan et al. 2007), recent stimulus history (Walker et al. 1981; Johnston et al. 2006; Burr et al. 2007; Heron et al. 2012) or when two duration signals are placed in conflict (Klink et al. 2011; Bausenhart et al. 2013).

This thesis consists of a series of psychophysical investigations studying distortions of perceived duration, chiefly in the sub-second range. To begin, Chapter 2 will review the dominant models in the field of time perception, and examine some of the psychophysical and neurophysiological findings which have contributed towards our current understanding of human sub-second temporal processing. Chapter 3 will provide an overview of the psychophysical methods used in timing studies, before describing the methods and calibration techniques used in this thesis. Chapters 4 – 7 will then detail a series of novel experiments which contribute to our understanding of the neural mechanisms underpinning human duration processing. The implications of these findings and conclusions drawn from this research will be discussed in Chapter 8.

Chapter 2: Literature Review

2.1 Introduction

The mechanisms with which we process temporal information remain elusive despite over 60 years of research. Historically, models of timing have centred on a dedicated system that operates in a similar manner to a clock. The success of these models originates from their relatively simple design and their ability to explain a large body of psychophysical data from decades of research. However, we have yet to find robust, neurobiological evidence of any such structure, and thus in recent years scientists have turned their attention to other models of temporal processing.

The following review examines the pioneering work on 'internal clock' models and the principal experimental findings which led to their development. The main competitors to this model are then discussed, followed by an overview of the psychophysical and neurophysiological evidence contributing towards our understanding of human duration processing.

2.2 Classical models of temporal processing

2.2.1 Scalar timing

Traditional models of temporal processing are dominated by variations on 'internal clock' models. These theoretical frameworks are said to be 'dedicated' (i.e. tasked with temporal processing alone), and 'central' (i.e. at a centralised neural location) in nature and measure time in a linear or metric fashion, akin to the ticking of a stopwatch. Also known as "pacemaker-

accumulator" models, they describe a modular mechanism with distinct processing stages, and were first conceptualised in the 1960's (Creelman 1962; Treisman 1963: see Figure 2.1).

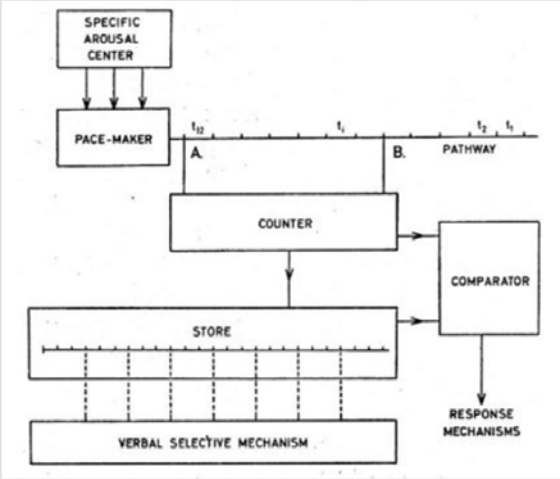

Figure 2.1: Schematic of the internal clock model proposed by Treisman (1963). During an interval to be timed (e.g. from A to B in the above example), the pacemaker emits pulses at a constant rate. These are then counted, and compared to previously stored values in memory (aided by previously learnt 'verbal labels', e.g. "10 seconds"). Depending on the outcome of this comparison, an appropriate response can be made. Figure reprinted with permission from the American Psychological Association.

Perhaps the most popular instantiation of these models is the "information processing" variant proposed by Gibbon and colleagues (Gibbon and Church 1984; Gibbon et al. 1984) and arose from "Scalar Expectancy Theory". This approach led to an exploration of what is now known to be a defining feature of both human and non-human time perception: the "scalar property" (Gibbon, 1971, Gibbon, 1977). This is a form of Weber's law for timing which describes the relationship between the perception of temporal extent and the standard deviation of these judgments. Specifically, linear increases in physical event duration produce similarly linear increases in perceived

duration. Critically, the variability of these duration estimates also increases in an approximately linear fashion. In other words, sensory variance is typically proportional to the magnitude of the estimated duration.

The scalar property is often compared across experimental conditions, sensory modalities and even species by calculating dimensionless 'Weber

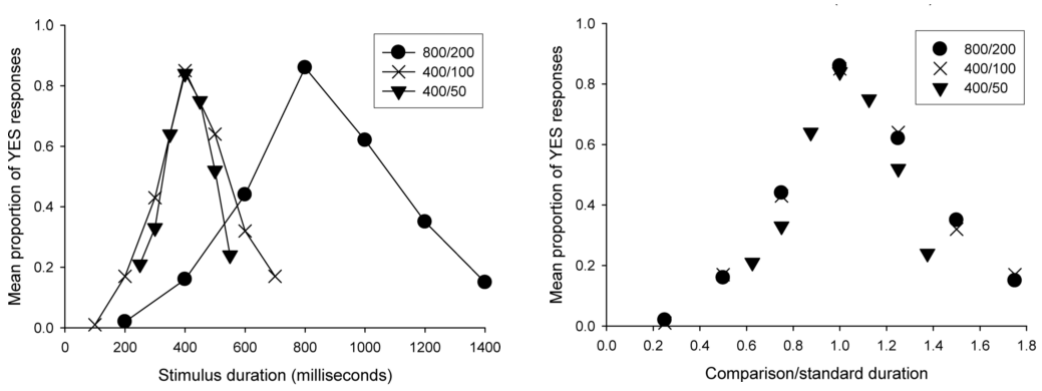

Figure 2.2: Hypothetical data sets from three 'temporal generalisation' tasks where observers decide whether the presented stimulus matches their memory of a reference duration (yes/no?). Typical results from three reference durations are plotted on the left in absolute terms, and demonstrate that the peak "yes" responses coincides with the reference duration (i.e. observer's perception of duration is, on average, approximately veridical). The spread of responses increases around the peak which reflects increasing response variability as a function of reference duration. On the right the same data are plotted on a relative scale, demonstrating commonality of each distribution's shape when its width is expressed in proportional terms. Figure reprinted from (Wearden and Lejeune 2008a) with permission from Taylor & Francis.

fractions' (the duration estimate variance divided by physical event duration). More commonly the timing literature quotes the 'coefficient of variation' (CV): the standard deviation of duration estimates divided by physical duration.

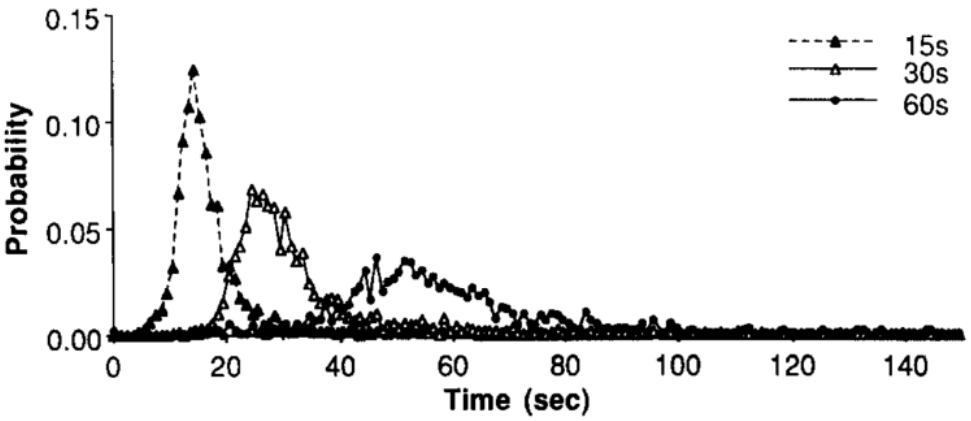

Figure 2.3: Mean response distributions for three groups of 10 rats, as a function of time. Each group was trained with a different interstimulus interval (ISI): 15 seconds, 30 seconds or 60 seconds. The spread of responses increases with the magnitude of the ISI, demonstrating the scalar property. Figure reprinted from (Gibbon 1992) with permission from Elsevier.

Conformity to the scalar property has several requirements (Gibbon 1977). Firstly, if an observer makes multiple duration judgements about a series of reference stimuli, the mean estimate for each reference should increase linearly with the reference duration. Secondly the standard deviation of the estimates should increase proportionally with the reference duration, and thus the ratio of one to the other (e.g. CV and/or Weber fraction) should remain constant. Finally, if measures of timed behaviour from several different absolute judgements are plotted on the same relative scale (i.e. as a function of the relative difference between judged durations), they should superimpose (Gibbon 1977) (see Figure 2.2).

The scalar property originated not from the study of timing *per se*, but from investigations of animal behaviour. Gibbon and colleagues sought to explain the temporally sensitive behaviour of animals witnessed during associative

learning (Gibbon 1971; Gibbon 1972; Gibbon 1977) or “conditioning”: an animal’s ability to learn the temporal interval between two events (separated by an interstimulus interval or “ISI”) following repeat exposure. Animals learn to execute their response in line with the expected onset time of the second stimulus (which may be a positive or negative experience depending on the type of reinforcement used). It was therefore hypothesised that they must have access to some internal measure of elapsed time (Gibbon 1977). Gibbon noted that when the probability of an animal responding was plotted as a function of time, the peak of the function corresponded to the physical ISI between the two stimuli (see Figure 2.3). More importantly, the spread of the animal’s responses conformed to the scalar property: the variability of responding increased proportionally with ISI (Gibbon 1971; Gibbon 1972). It was this observation that led to the formalisation of Scalar Expectancy Theory (SET), describing the ‘scalar’ spread of an animal’s responses at the ‘expected’ time of reward.

2.2.2 Internal clock models

On the basis of Scalar Expectancy Theory (SET), Gibbon and colleagues then proposed their own dedicated, modular clock model known as the “information-processing” model of timing, so named because it reflects the three stages of information processing involved (e.g. ‘clock’, ‘memory’ and ‘decision’).This model has been used to explain both human and animal timing behaviour (Gibbon and Church 1984; Gibbon et al. 1984).

The model's key concepts are similar to those evoked in the earlier internal clock models of Creelman and Treisman (Creelman 1962; Treisman 1963). Of particular importance is that the components of these clock models were designed to lead us through the various stages required for an animal or human to decide when to respond (i.e. "has sufficient time elapsed for a response, yes/no?"). As such, they do not translate smoothly into an explanation of how humans might make comparative decisions, for example in judging the relative lengths of two consecutive durations (i.e. "was the first stimulus longer or shorter than the second?"). Despite this, the move towards applying these models to human timing grew in popularity throughout the late 1980's and early 1990's (Zeiler et al. 1987; Wearden and McShane 1988; Allan and Gibbon 1991; Wearden 1991a; Wearden 1991b). Variations on internal clock models have now dominated the field of timing behaviour ever since.

For ease of explanation, the following description of a traditional SET-based clock model will focus on how an animal might process duration in order to prepare a motor response (as originally described by Gibbon et al. 1984), although this could equally be applied to human perceptual decisions requiring a "yes/no" response.

The clock stage contains a 'pacemaker' that emits pulses at a constant rate, and an 'accumulator' which counts and stores the pulses (see Figure 2.4). If each pulse represents a unit of time, the sum of these pulses will correspond to a particular duration. Between the pacemaker and accumulator is a switch, which controls the flow of pulses and may be mediated by attention (Gibbon and Church 1984; Treisman et al. 1990; Zakay and Block 1995). At the onset

of a timed stimulus this switch closes, completing the circuit between the pacemaker and accumulator and allowing pulses to accrue within the

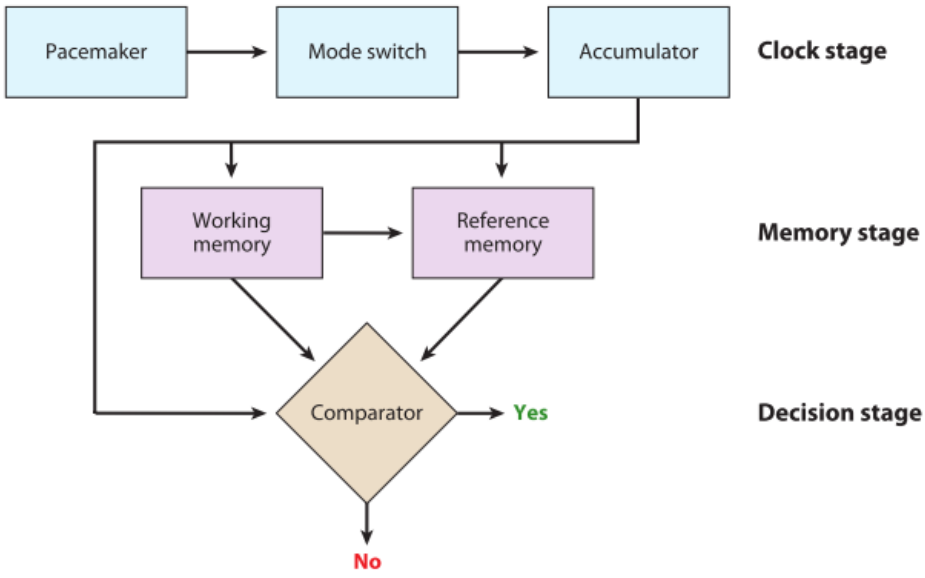

Figure 2.4: Schematic of the basic components of an internal clock (or information processing) model, with clock, memory and decision stages. Figure reprinted from (Allman et al. 2014) with permission from Annual Reviews.

accumulator (where each pulse passing the switch increases the value in the accumulator by one). At stimulus offset, the switch opens again, thereby terminating the flow of pulses. The number of pulses in the accumulator is held in working memory, and then compared to previously stored durations in reference memory. If a match is found, the animal (or human) can make a decision about whether to respond.

It is proposed that the pacemaker behaves as a Poisson timer: although each pulse is generated at a random onset time, collectively, the pulses will tend to be emitted at a consistent average rate over time. An accurate

Poisson timer would generally produce non-scalar variance, and therefore the source of the scalar property is often attributed to the memory and decision stages (Gibbon et al. 1984; Gibbon 1991; Gibbon 1992). Here, the scalar property arises not because of imprecision when encoding the duration initially, but because storing and then retrieving a previously encountered duration cannot be achieved accurately. A schematic outlining how this might arise is given in Figure 2.5, and a detailed explanation (as per the original information processing account) is given below.

During the initial phase of associative learning the animal is repetitively experiencing the same duration, D_1 on each trial (Figure 2.5a). If we assume that the number of pulses corresponding to this reinforced duration is equal to some value n_1, it is proposed that the scalar property arises as the result of copying this value from the accumulator into a more permanent representation in reference memory. During this process, the value n_1 is hypothetically multiplied by a "memory translation constant" k^*, (a value which is drawn from a normal distribution). Thus, the representation of n_1 in reference memory will also be normally distributed, which we can call n_{rep} (Figure 2.5b). Due to the multiplicative effect of k^*, this distribution will have a broader spread if D_1 is a longer duration (e.g. 800ms) compared to if it was a shorter duration (e.g. 200ms) (Gibbon and Church 1984; Gibbon et al. 1984) (Figure 2.5c). Later, the animal may be presented with a new duration D_2, and must decide "is D_2 equal to D_1?" The new value in the accumulator n_2 is therefore compared with a random sample taken from the distribution of n_{rep} (Figure 2.5d). Since the spread of this distribution is proportional to the magnitude of n_1 (and hence D_1), there will be a greater range of possible

values if D_1 was a relatively long duration than if it was a relatively short duration (e.g. Figure 2.5c). Therefore when a comparison between the n_{rep}

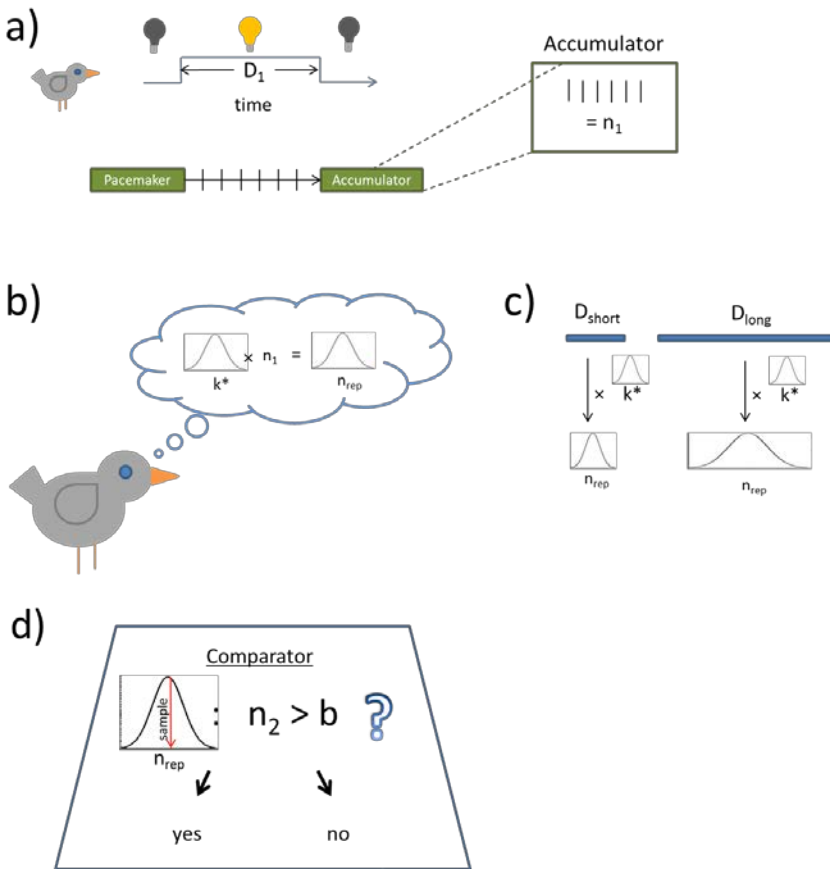

Figure 2.5: Schematic demonstrating how the scalar property might arise during the encoding of durations in reference memory. **a)** During associative learning the animal is repeatedly shown a stimulus of duration D_1, which corresponds to a number of pulses entering the accumulator. At the end of the trial, this number of pulses is given by the value n_1. **b)** To store the value of n_1 in memory, this is multiplied by a memory translation constant k*(drawn from a normal distribution). This leads to a representation of n_1 in memory which is also normally distributed, shown in this schematic as n_{rep}. **c)** Due to the multiplicative effect of k*, the neural representation of longer durations will have a larger spread than that of shorter durations. **d)** Later when a second duration is presented (corresponding to an accumulator value of n_2), the animal must decide if the ratio of this value to a random sample taken from n_{rep} is sufficient to meet a threshold value *b*. The decision to respond is then based on this outcome. If n_{rep} has a larger spread, a greater number of possible samples exist, and therefore the variability of this decision to respond increases (resulting in the scalar property).

sample and n_2 is made during the decision stage, this comparison will yield a more variable result if D_1 was a relatively long duration. Given that it is this ratio that is compared to some predetermined threshold value (perhaps determined by the animal's individual 'criterion' for deciding "yes") in order to decide if a behavioural response should be made, this manifests as increasingly variable responses with increasingly longer durations – resulting in the scalar property (Church and Gibbon 1982; Gibbon et al. 1984).

In addition to the memory stage, variance could arise from the clock or decision stages of 'internal clock' models. Timing tasks generally involve all three stages, making it difficult to isolate and manipulate the variance occurring at each individual stage. Furthermore, the memory and decision processes described by early clock models detail how durations might be stored and compared by animals under very narrow experimental conditions (e.g. learning the temporal relationship between stimulus and reward). How these processes would translate to humans, who process a potentially infinite range of ecologically relevant durations, remains unclear.

In order to make temporal judgements in everyday life we presumably require a 'database of stored durations', whose representations might be formed through 'real-world' experiences rather than as a result of repetitive learning in a lab, from which we can then make comparisons to a recent temporal experience. Whether the capacity of human memory is large enough to store such a potentially large database of durations, how these durations are neurally represented, and what decision processes govern the comparison between these durations pose important questions for internal clock models.

In addition, a persistent criticism of internal clock models is that non-specific references to memory constants, ratio comparisons and individual threshold criterions in the original models have led to numerous reinterpretations, such that a seemingly infinite variety of experimental findings can be accommodated by adding components to versions of the model (Staddon and Higa 1999; Wearden 1999; Van Rijn et al. 2014). Attempts to explain the abundance of temporal associations, illusions and perceptual distortions in the time perception literature (see Section 2.4) pose many problems for internal clock models. In addition to having a relatively large number of 'degrees of freedom', perhaps the biggest challenge facing these models is a lack of convincing neurobiological support (see Section 2.5). Despite decades of research the literature has no consensus as to the neural basis of the model components shown in Figure 2.4.

2.3 Alternative models of temporal processing

In addition to the dominant internal clock model discussed in Section 2.2.2, a growing number of alternative models have also been proposed for the processing of event duration. A detailed discussion of all of these models goes beyond the scope of this review, and therefore instead we will examine the dominant competitors.

2.3.1 The striatal beat frequency model

The original "beat frequency" model proposed that groups of oscillating neurons ("oscillators") could code duration through patterns of coincident

firing activity (Miall 1989). Oscillatory neurons are those whose firing activity occurs in a rhythmic, cyclical manner as a function of time (e.g. see Figure 2.7 upper panel – the firing activity of three different 'oscillators' each with a different frequency is shown). Assuming that each oscillator within the group operates at a slightly different frequency, and that the neurons are only

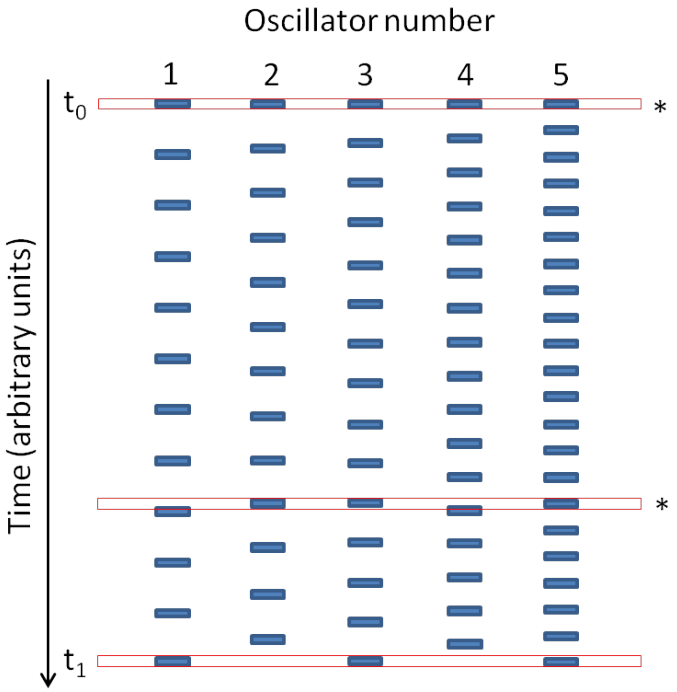

Figure 2.6: Schematic showing five oscillators of differing frequency whose activity is represented by the blue bars. Oscillators 1, 3 and 5 can be used to encode the duration of the interval from t_0 to t_1, as their activity is synchronous at both time points. The asterisks provide a further example of a beat frequency involving only oscillators 2, 3 and 5, which occurs at a different moment in time.

active for a brief time during each cycle, the "beat frequency" refers to the point at which two (or more) oscillators are active at the same moment (see Figure 1 red boxes). If all the oscillators are connected to a single readout unit, duration can be encoded by identifying which oscillators provide a beat frequency at both the event's onset and offset. An example is given by

oscillators 1, 3 and 5 in Figure 2.6, whose beat frequencies encode the interval between arbitrary time points t_0 and t_1. Retrieving a duration at a later stage can be achieved by resetting the phase of the oscillators, and waiting for that particular beat frequency to arise again. A group of oscillators can therefore encode a variety of different durations, as each duration would be associated with a different beat frequency.

An important aspect of this model is that the duration encoded by a beat frequency can be much longer than a duration encoded by the periodicity of a single oscillator, in a similar manner to the lowest common multiple of a series of numbers (see Figure 2.7, top image). As such, large populations of oscillators could potentially produce unique beat frequencies (peaks of coincident neural firing activity) for a variety of different durations across several orders of magnitude. Early computer simulations of the model showed promise, as peaks of activity occurred at a series of temporally distinct durations (Miall 1989). However, as the original model did not allow for any variance components (such as that attributed to background neural noise), it was criticised for being biologically implausible and of limited use in its present state (Matell and Meck 2004).

Building on the main concepts of the model, Matell and Meck created their own, neurobiologically inspired beat frequency model which they named the “striatal beat frequency” (SBF) model (Matell and Meck 2000) (see Figure 2.7). In this variation of the model, the oscillators are cortical neurons whose coincident activity is detected by spiny neurons in the striatum. The striatum is a nucleus within the basal ganglia, a subcortical region of the brain that has been implicated in duration processing (Matell et al. 2003; Chiba et al.

2008; Coull and Nobre 2008) (see Section 2.5). It primarily receives its input from the cortex (Cowan and Wilson 1994), and importantly, these inputs demonstrate ‘many-to-one’ convergence: between 10000 - 30000 cortical axons project to each spiny neuron in the striatum (Wilson 1995). The striatum is therefore well positioned to act as a coincidence detector for temporal patterns of activity arising from the sensory cortices (Matell and Meck 2004; Merchant et al. 2013a). Differentiation between these patterns would then be achieved through chronotopic arrangement of the spiny neurons, in which each neighbouring neuron responds preferentially to a different, specific range of durations. The firing activity of this striatal ‘timeline’ is then monitored and read out by neurons in the frontal cortex. It is proposed that dopamine could be involved in altering the effectiveness of cortical input on the spiny neurons of the striatum, and thus may have a modulatory role (Di Chiara et al. 1994; Matell and Meck 2004). Additionally dopamine may act as a ‘starting gun’ for interval timing, resetting the phase of the cortical oscillators when timing is initiated (Matell and Meck 2004; Parker et al. 2014). Initial computational simulations of the SBF model have successfully resembled human and animal behavioural data, and have also demonstrated the scalar property (Matell and Meck 2004; Buhusi and Oprisan 2013).

An advantage of this model is that it incorporates both modality specific (i.e. 'local' cortical oscillators) and modality independent (i.e. 'central' striatal coincidence detection) components. Oscillators residing at relatively peripheral locations within the brain could help to explain some of the

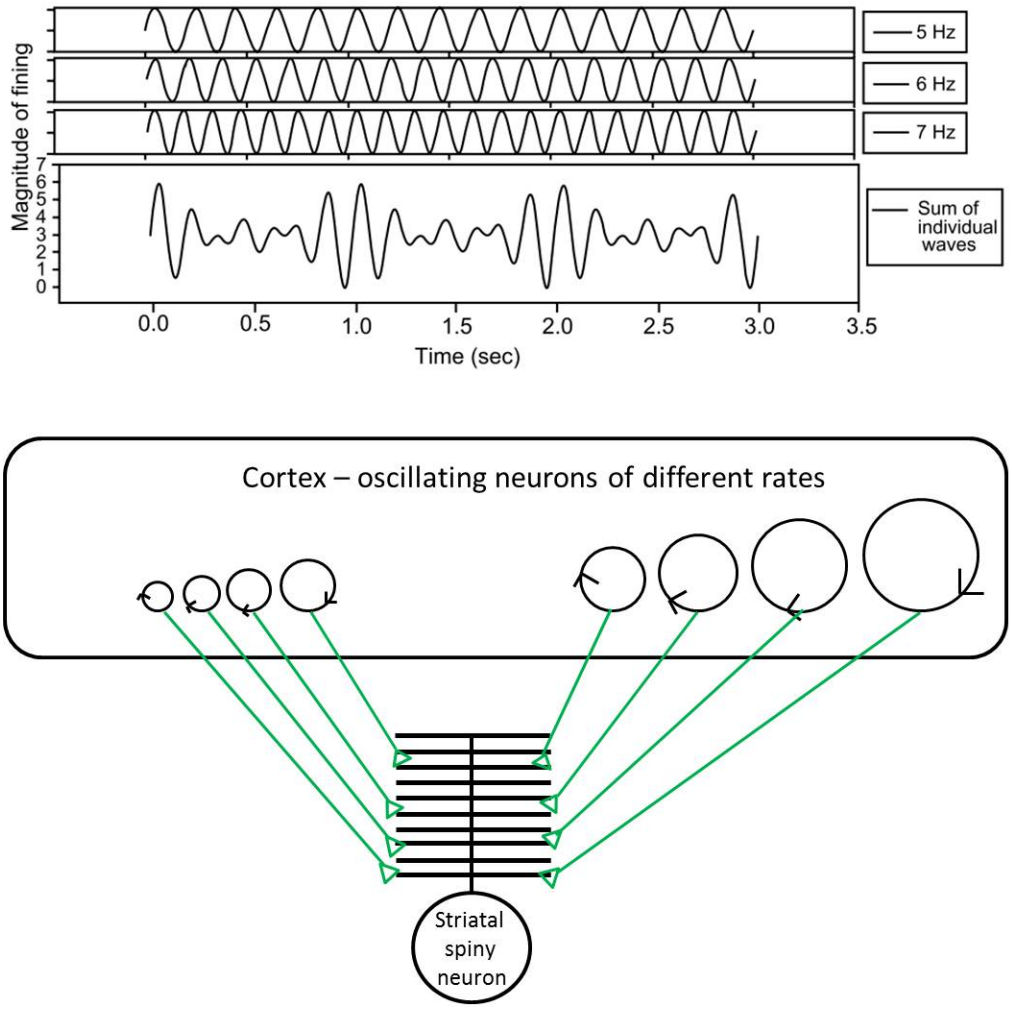

Figure 2.7: Schematic showing the components of the striatal beat frequency model. The top panel shows the signal from three different oscillators operating at 5Hz, 6Hz and 7Hz. The sum of these individual waves is shown below; where increases in signal amplitude correspond to the beat frequencies, (i.e. where the peaks of each oscillator signal are coincident). The periodicity of the summed waveform is therefore greater than that of the individual components, demonstrating the potential for this mechanism to encode a large range of durations. The lower panel shows how many cortical oscillator signals may converge onto a single spiny neuron, allowing it to act as a coincidence detector. Figure adapted from (Matell and Meck 2004), and reprinted with permission from Elsevier.

temporal illusions and differences between sensory modalities found in the literature (see Section 2.4), especially as neurons throughout the mammalian brain are known to exhibit a variety of oscillatory frequencies (Llinás 1988; Hari and Salmelin 1997). If oscillators within the auditory cortex oscillate at a different average frequency to those in the visual cortex, but beat frequencies are then detected by the same striatal mechanism, this could potentially explain phenomena such as the 'sound longer than vision' bias when the two signals are compared (see Section 2.4.3).

A difficulty facing the SBF model is that whilst it is biologically plausible, there is conflicting neurophysiological evidence regarding the functional role of the basal ganglia (and striatum) in duration processing (see Section 2.5). Furthermore, as with most neurobiological models of timing, one of the biggest criticisms is whether a mechanism built around specific patterns of oscillatory activity can withstand the variations in neural activity and fluctuating noise levels found in biological systems. However, recent model simulations (based on the known properties of neuronal activity) are beginning to address this issue, and have shown that one or more sources of noise (such as variation in the neuron firing frequency) are in fact necessary for the emergence of certain characteristics of interval timing, such as the scalar property (Oprisan and Buhusi 2013).

2.3.2 The state dependent network model

An alternative theory is that duration is processed intrinsically, without the need for a dedicated timing system. This possibility is raised by the

predictions of state-dependent network (SDN) models, which posit that millisecond range duration information can be inferred from time-varying information contained within ‘normal’ (ostensibly non-temporal) ongoing neuronal activity associated with a stimulus (Buonomano and Merzenich 1995; Buonomano 2000).

The key concept in the SDN model is that throughout the time period covering a stimulus’ presentation, interconnected groups (i.e. networks) of neurons create complex, multidimensional spatial distributions of activity (see Figure 2.8). At any given point following the start of a stimulus presentation, the balance of excitatory and inhibitory connections between neurons within a network creates time-locked activity within each neuron. This spread of activity can be exploited as it forms a unique temporal signature: the evolution of network activity will ensure that the spatial distribution of active and inactive neurons will be unique at each passing moment following the stimulus onset. This spatiotemporal specificity is analogous to the spatial pattern of ripples observed after throwing a series of pebbles into standing water (e.g. Figure 2.9). Interaction between the ripples caused by each pebble cause the overall spatial pattern of these ripples to constantly evolve. Taking a series of photographs of the water’s surface would provide a unique snapshot of the pattern’s evolution, a pattern that will contain critical spatial differences to those observed in any photographs taken before/after this point in time. If a system could extract the temporal sequence of events evolving from each snapshot’s spatial information, then duration could be inferred.

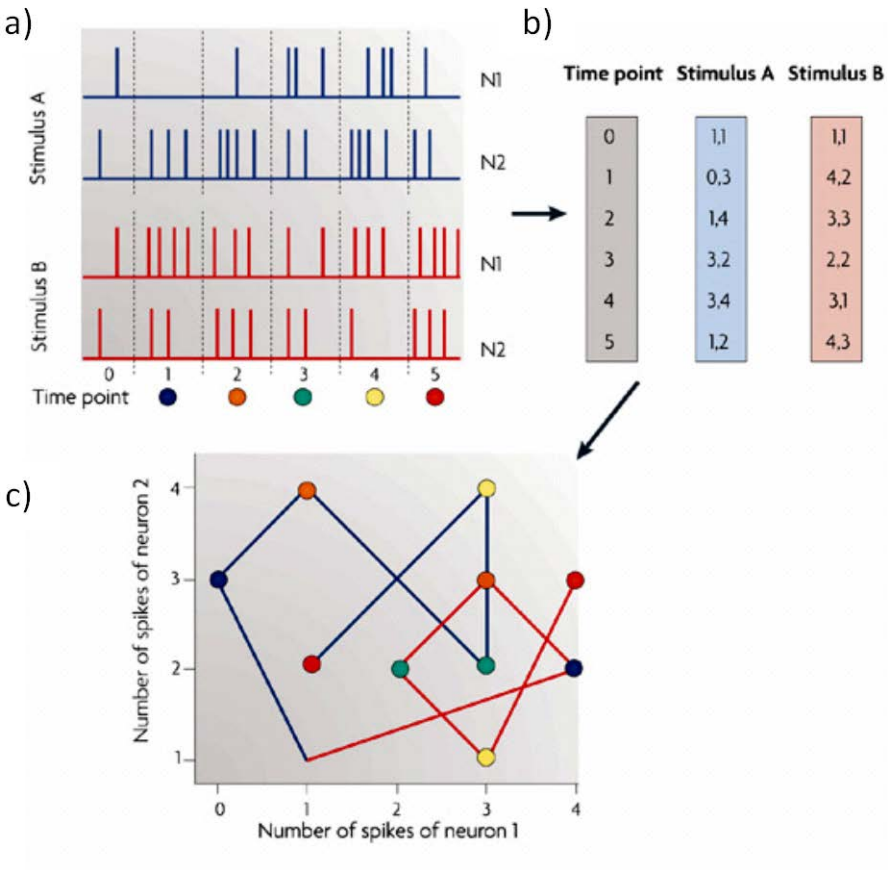

Figure 2.8: A schematic of how two neurons could encode duration through different spatiotemporal patterns of activity following stimulus presentation. **a)** The spikes of activity of two separate neurons (blue and red data) to two hypothetical stimuli (A&B) are shown in a series of time bins represented by the coloured circles. **b)** Here the number of spikes in each time bin are converted into spatial coordinates giving the trajectory of each neuron. **c)** Showing the different trajectories of neurons A&B, where the coloured circles represent the same time bins as previously. Each point along the trajectory can be used to determine how long ago the stimulus was presented, giving a measure of duration for the stimulus. Figure reprinted from (Buonomano and Maass 2009) with permission from Macmillan Publishers.

Some measure of variability in the strength of the input signal could be responsible for creating short term plasticity. For example, in the simple disynaptic network discussed by Buonomano, two very brief transient stimuli delivered 100ms apart could result in the generation of an action potential to either the first or second signal depending on the signal strength

(Buonomano 2000). If the signals are weak, then the first pulse would not be sufficient for a threshold to be met, and thus the signal would not be propagated to the next neuron. Only when the second pulse arrives and is summed with the decaying (yet subthreshold) signal from the first pulse is

Figure 2.9: Showing an example of the complex patterns of ripples when stones are thrown into a liquid. Source: http://wordservewatercooler.com/2012/05/15/advantages-of-book-tours/ Accessed: 23/06/2014

the action potential generated. However, if the original signals are strong, then the action potential will be generated by the first pulse but not the second pulse, since the neuron will be in a refractory period (i.e. recovery phase) when the second pulse arrives.

This principle, on a much larger scale, allows us to comprehend how inherent neuronal properties could give rise to different spatiotemporal firing patterns for identical successive stimuli: the pattern of firing for any given neuron at any given point in time is dependent on the current network state.

This has the potential to allow a system to ‘read-out’ the network state and thus elapsed time between successive action potentials.

An advantage of the SDN model is that timing could be performed ‘locally’, without recourse to dedicated neural hardware such as pacemakers, accumulators or switches etc. Theoretically, such a mechanism could yield rapid, efficient duration estimates that arise as a by-product of other ongoing neural activity.

A prediction arising from the model is that because neurons will be unable to respond during their refractory period, two successive brief stimuli with a short interstimulus interval (ISI) may be processed with less precision. Indeed, variability in a duration discrimination task has been shown to increase with very short ISIs as long as the stimulus characteristics remained constant (Karmarkar and Buonomano 2007; Burr et al. 2009b). When auditory stimuli were presented at different frequencies very little impairment was found, consistent with the idea that timing occurs in localised networks, and if stimuli are detected by different networks (i.e. as per the auditory system’s tonotopic organisation) they would be little interaction between the two (Karmarkar and Buonomano 2007).

Perhaps the greatest limitation of SDN models is that in order to interpret the temporal signature from a network state, the read-out system would need to learn and interpret an enormous (potentially infinite) number of stimulus-specific spatiotemporal neural activity patterns. For example, in order to compare auditory and visual durations, read-out neurons would need to rapidly interpret and compare complex patterns of neuronal activity,

potentially arising at different rates, then compare the two read-outs across modalities. Although the SDN model is suggested to be appropriate only for durations below 500ms (due to the known temporal properties of neurons and spike production (Karmarkar and Buonomano 2007)), this would still require knowledge of a huge number of network states.

Another challenge for SDN models is that they are highly susceptible to neural noise, and tiny fluctuations in background activity result in vastly different spatiotemporal neural trajectories making it extremely difficult to reliably reproduce consistent patterns of activity. Computationally, attempts have been made to model networks that can be trained to robustly reproduce trajectories in the presence of noise. However, these models are still thought to be biologically implausible, requiring learning rules that operate too quickly to be realistic (Laje and Buonomano 2013).

It is also unclear how SDN models could produce any of the distortions of perceived duration induced by adaptation, changes in attentional state or cross-modal comparison (see Section 2.4). In theory, this could involve a speeding up or slowing down of the temporal progression through different network states (or potentially via errors in the read-out of these states). Yamazaki and Tanaka (2005) provide one computational example of how varying certain parameters (e.g. the range over which neural activity is integrated) of a neural network model could change the speed at which neural patterns of activity were generated, hypothetically accounting for temporal compression/expansion of perceived duration. However this has yet to be replicated *in vitro*, and addressing such perceptual distortions remains largely unexplored territory in the SDN literature.

2.3.3 The channel-based processing model

Rather than relying on oscillatory processes or spatiotemporal patterns of activity, models of channel-based temporal processing centre on the idea

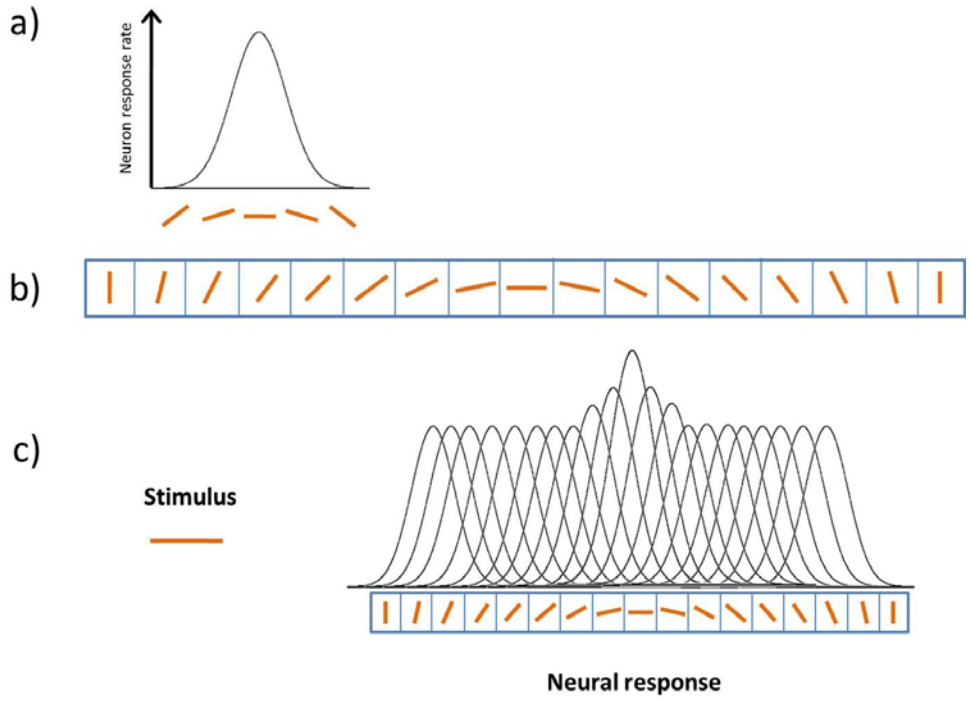

Figure 2.10: Schematic showing a channel-based processing system for orientation. **a)** An example of the neural response distribution for a neuron maximally responsive to the horizontal orientation (its preferred orientation). Its response amplitude declines progressively for orientations shifted clockwise or anticlockwise from horizontal. The rate of this decline dictates the width of the distribution and is referred to as the neuron's 'bandwidth'. The narrower the bandwidth the more sensitive the neuron is to changes in orientation. **b)** An example of the organised distribution of preferred stimulus orientations, where each box represents a different neuron's orientation tuning preference. **c)** A population of overlapping tuned neurons, each centred on a different 'preferred' orientation, could represent all orientations. The orientation of any given stimulus could then be inferred from examining the relative firing rates across the whole population. In the example above, the horizontal stimulus induces an increase in the firing rate of "horizontal tuned" neurons, with neighbouring neurons tuned to orientations slightly clockwise and anticlockwise from horizontal showing smaller increases in firing rate. Figure adapted from Snowden et al. 2012.

that populations of neurons could respond selectively to a narrow range of duration signals, and therefore encode brief intervals of time (Ivry 1996). In other words, groups of neurons could be “tuned” for certain durations therefore forming ‘duration channels’.

The conceptual foundations of this model are rooted in the channel-based systems evoked to explain the processing of spatial attributes such as stimulus orientation (see Figure 2.11), spatial frequency (Hubel and Wiesel 1962), motion direction (Hubel and Wiesel 1968; Gross et al. 1972) complex 3D shapes (Logothetis et al. 1995) and faces (Bruce et al. 1981; Rolls 1984). Visual spatial channels were originally described in a series of highly influential studies using single cell recording techniques in cat and monkey striate cortex (Hubel and Wiesel 1962; Hubel and Wiesel 1968). They documented the existence of neurons that respond selectively to a narrow range of stimulus orientations, with the maximum response corresponding to the ‘preferred’ orientation in the centre of the range (see Figure 2.10a).

These neurons are arranged in an orderly fashion within the visual cortex: neighbouring neurons each prefer an orientation that is slightly advanced from the previous one, causing a progressive shift in orientation tuning until all possible orientations are accounted for (Figure 2.10b). Comparing the relative firing rates of the population as a whole, and extracting the ‘channel’ exhibiting the greatest peak in activity will therefore give a measure of stimulus orientation (Figure 2.10c). One consequence of this orderly arrangement is that adapting to the extended presentation of a particular orientation results in a “tilt aftereffect” (Gibson and Radner 1937) (see Figure 2.11). If one continuously views a rightward tilting line for about 20 seconds

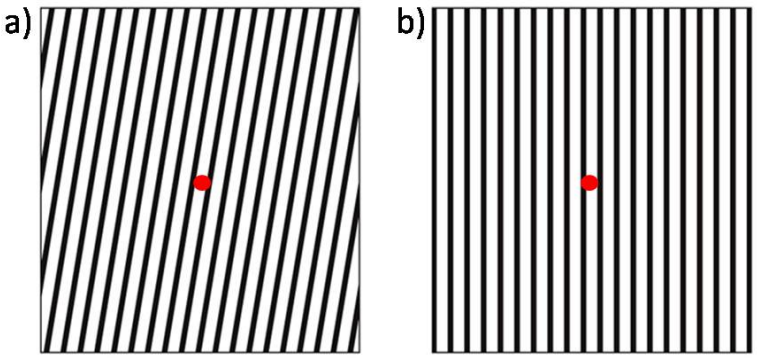

Figure 2.11: The tilt aftereffect. To generate the illusion, begin by fixating on the red dot in the centre of image **a)** for about 20 seconds. Then switch your gaze to the dot in image **b)**. The vertical lines should appear to tilt slightly to the left.

and then views a vertical line, the vertical line will be temporarily perceived as tilting to the left. This is a type of 'repulsive' aftereffect, as the perceptual distortion is in the *opposite* direction to the adapting stimulus. A neural explanation for this illusion, linking the responses of "orientation tuned" neurons to the adaptation-induced perceptual bias is given in Figure 2.12.

Aftereffects have been described as a "psychophysicist's microelectrode" (Frisby 1980; Thompson and Burr 2009), since they indicate underlying neuronal selectivity to a given stimulus feature. In the spatial domain aftereffects have been reported for a host of attributes including, but not limited to, visual motion (Wohlgemuth 1911; Levinson and Sekuler 1976), spatial frequency (Blakemore and Sutton 1969), facial features (Webster and MacLin 1999; Webster et al. 2004), auditory localisation (Kashino and Nishida 1998) and auditory motion (Grantham and Wightman 1979).

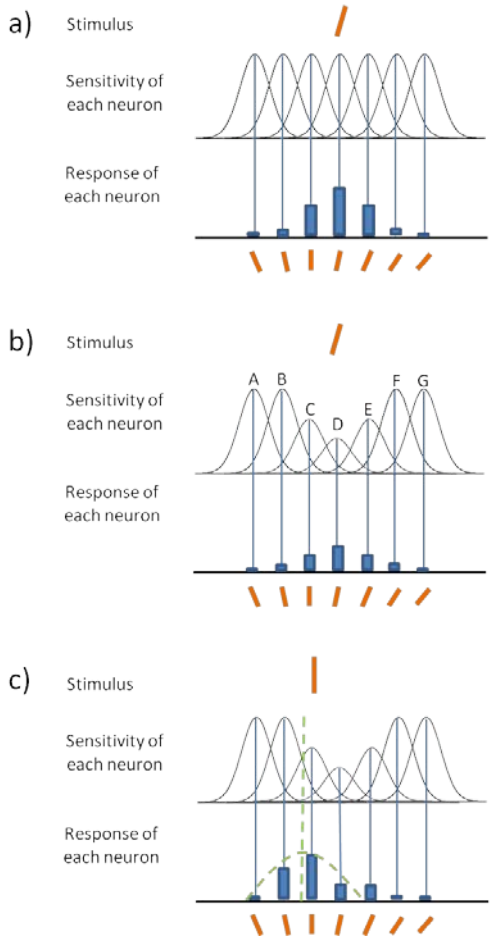

Figure 2.12: A neural explanation of the tilt aftereffect. **a)** Initially, viewing a line that is tilted 10° to the right will cause neurons tuned to a 10° rightward tilt to increase their response. **b)** Continuous presentation of this line (the period of adaptation) will cause a reduction in the response of neurons whose preferred orientation is closest to 10°, and to a lesser extent neurons with different preferred orientations but whose tuning curves overlap with 10° (neurons C and E in this example). **c)** If a vertical line is subsequently viewed, the neuron that responds maximally to this orientation in its unadapted state continues to show reduced activity (neuron C), and thus *its* neighbouring neuron (neuron B) shows a relatively greater response. When activity is compared across all neurons, the peak response is shifted leftward, as shown by the green dashed line. Since neuron B's preferred orientation is 10° to the left, the vertical line is perceived to have a leftward tilt. Figure adapted from Snowden et al. 2012.

In the temporal domain, applying a channel-based system to the processing of event duration would require clusters of neurons with preferences for slightly different (overlapping) duration ranges. These neurons would have activity levels that are dependent on the match between afferent duration

information and their preferred duration. Downstream read-out neurons could then determine the most likely event duration by comparing the relative firing rates of these neural populations (i.e. the activity of each channel). The existence of ‘duration aftereffects’, which show similar characteristics to those from the spatial literature (i.e. repulsive in nature and bandwidth limited) (Walker et al. 1981; Heron et al. 2012; Li et al. 2015b: see also Section 2.4.4 of this thesis), provide evidence in support of channel-based processing.

An advantage to channel-based models is that they are able to conceivably explain sensory modality differences, such as the higher variability of visual duration estimates compared to auditory estimates (see Section 2.4.3). For example, neurons in the visual cortex may be more broadly tuned for duration (i.e. have a wider duration bandwidth) compared to those in the auditory cortex (Figure 2.13a). Assuming that activity from these neurons is signalled to read-out neurons downstream once it reaches a certain threshold value, a larger range of durations would exceed this threshold if the bandwidth is greater (Figure 2.13 red arrows). Alternatively, the visual neurons may have a lower threshold for signalling to read-out neurons, also resulting in a wider range of durations exceeding this threshold value (Figure 2.13b). In both of these examples, the visual neurons would be less sensitive to signalling small changes in duration, resulting in poorer discrimination and more noisy duration estimates.

However, a potential difficulty for channel-based models arises from the continuous nature of time itself. Imagine a task that involves estimating the

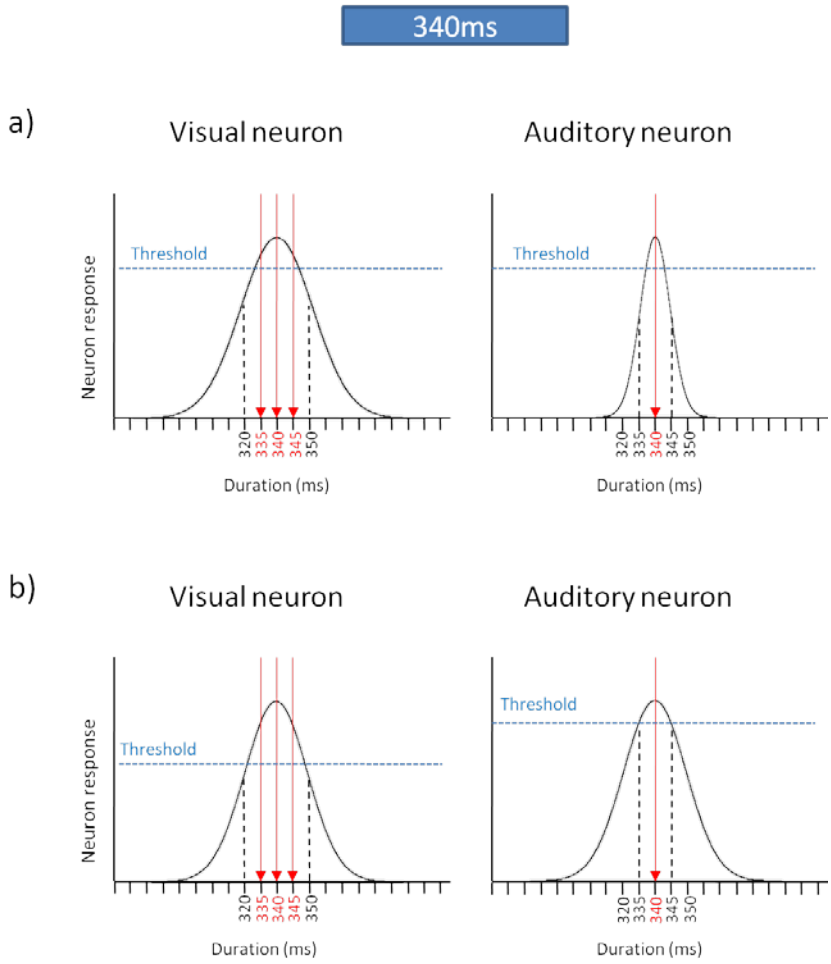

Figure 2.13: Simplified schematic demonstrating two ways in which the properties of visual neurons could result in lower temporal sensitivity compared to auditory neurons. The response of one visual and one auditory neuron is shown. Both neurons have a peak response to 340ms, and respond less strongly to other durations close to 340ms as determined by their bandwidth (the width of the tuning curve). Neurons send a signal (activating that 'channel') when a duration occurs that causes the neuronal response to exceed some threshold value (the blue dashed line). **a)** Here, the visual neuron has a wider bandwidth than the auditory neuron, and therefore a greater range of durations (335 – 345ms) would cause a neuronal response that exceeds the threshold. Thus, the auditory neuron is able to discriminate between 335ms and 340ms but the visual neuron cannot. **b)** In this example, both neurons have the same bandwidth, but the visual neuron has a lower threshold for signalling a response, again resulting in poorer sensitivity.

duration of a flash of light. If the physical duration of the flash is 100ms, we would assume that neurons tuned to 100ms (situated somewhere in the visual cortex) would become activated by the light, and this increase in channel activity would create a signal telling us that the duration is 100ms. However from the moment that the light appears (the onset) to when it disappears at 100ms, duration is progressively increasing. Therefore, moment by moment, neurons with successively longer preferred durations become activated in a 'domino-type' fashion. This would result in the centre of 'channel activity' being skewed towards the short duration range, creating spurious duration analysis at the read-out stage.

Additionally, channel-based models require some form of read-out mechanism further downstream to detect changes in the relative firing rates within the 'duration tuned' population, the properties and location of which are still unclear.

2.3.4 Summary

Our understanding of temporal processing still lags far behind that of spatial processing, and whilst numerous models exist to explain how event duration is estimated, none can convincingly explain all of the behavioural findings. In particular, this concerns the deviation from veridical duration estimation that occurs under certain experimental conditions such as when attention is manipulated (e.g. by the presence of distracter tasks), when signals are compared across sensory modalities or following repeated stimulus

presentation (see Section 2.4). Aligning model predictions with the existing behavioural data remains an ongoing challenge for researchers in this field.

2.4 Psychophysical Time

In this section we will examine some of the psychophysical characteristics of sub-second and supra-second timing, including some key examples of how distortions of perceived duration have provided useful insights into the possible mechanisms subserving temporal processing.

2.4.1 Conformity to the scalar property

Despite most internal clock models being centred on the concept of scalar timing, Weber fractions do not always remain constant across all durations, particularly those below 200ms (Getty 1975; Grondin 1993; Grondin 2010; Rammsayer 2010: see Figure 2.14) and above 1000ms (Getty 1975; Bangert et al. 2011; Grondin 2012) (see Figure 2.15). One proposed explanation is that the total variance resulting from a duration estimate can be divided into two independent components, one which arises directly from the timing mechanism and one which is independent of duration, and could be attributed to other background psychological processes (e.g. signal detection, implementing a response etc…) (Getty 1975). Whilst the variance arising from the timing mechanism is assumed to vary proportionally with duration (i.e. conform to Weber's Law), the duration-independent component is invariant across a range of different durations. This concept gave rise to a

generalised form of Weber's law for duration, which included an additive variance component that is independent of duration (Getty 1975):

$$\sigma^2 = kD + c$$

Where σ^2 is the total variance, k is a constant proportionality representing the rate of increase in duration dependent variance, D is the reference duration and c is a constant representing the duration-independent variance.

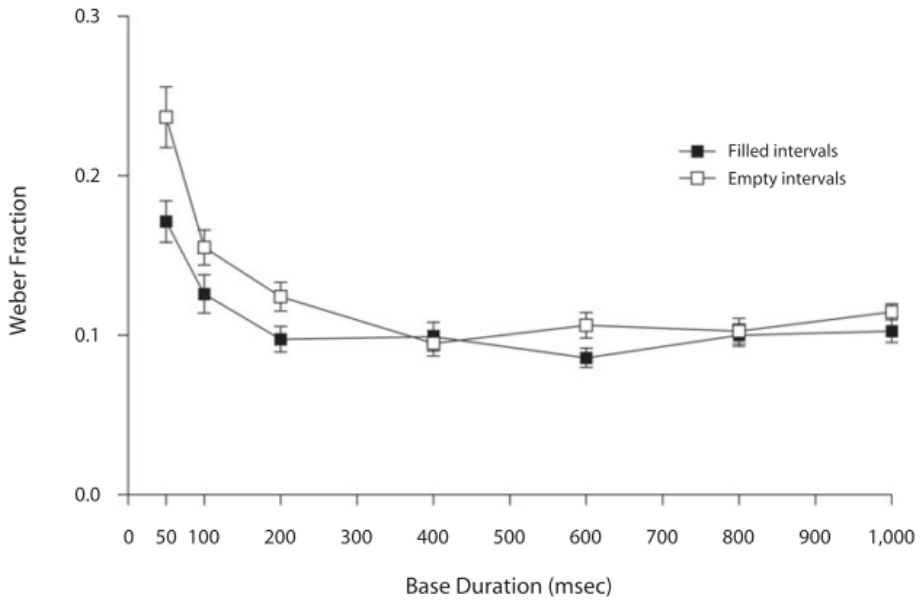

Figure 2.14: Increased Weber fractions are shown for base durations below 200ms during a duration discrimination task, in which observers had to decide "which was longer?" between a fixed base duration (7 base durations were tested in separate blocks) and a variable comparison duration (which varied around each base duration as per an adaptive staircase procedure). One group (n = 24) performed the task with filled auditory durations and another group (n = 24) with empty durations (marked by brief auditory markers). Error bars denote the standard error of the mean. Figure reprinted from (Rammsayer 2010) with permission from Springer.

In the context of ‘internal clock’ models, it is suggested that the duration-independent variance could represent variability in the opening and closing of the switch between the pacemaker and accumulator (see Section 2.2 Figure 2.4). This can be thought of as “start/stop” variance, in that it relates

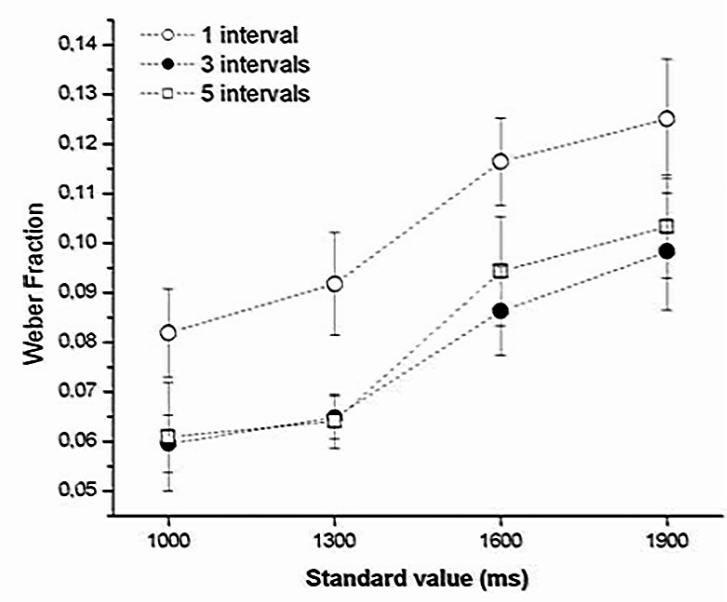

Figure 2.15: Increasing Weber fraction with increasing stimulus duration in a task where observers had to decide whether the second (comparison) of two intervals (marked with brief auditory tones) was longer or shorter than the first (standard) interval. In each case the duration of the comparison interval varied in seven steps around the duration of the standard interval. The number of presentations of the standard interval was also varied in separate blocks (1, 3 or 5 intervals). Irrespective of the number of standard intervals, the Weber fraction increased with increasing standard duration. Figure reprinted from (Grondin 2012) with permission from the American Psychological Association.

to the starting and stopping of the timing process rather than from the timing process itself (Wearden and Lejeune 2008a). Whilst it might vary slightly from trial to trial, it is not proportional to the duration being timed. Thus, when encoding short durations, this duration-independent variance would make up a proportionally larger amount of the total variance compared to when timing

longer durations, which could explain the increased Weber fractions for durations below 200ms.

Yet, this doesn't explain monotonic decreases in the Weber fraction with increasing duration (Wearden et al. 1998; Wearden et al. 2007; Lewis and Miall 2009) or increases with durations greater than 1000ms (Getty 1975; Grondin 2012: see Figure 2.15), (referring to increases and decreases in observer sensitivity respectively). Instead, this could arise as a result of humans adopting chronometric counting strategies, which is proposed to be particularly advantageous when timing durations beyond 1200ms (Grondin et al. 1999), and can reduce variability in an observer's judgements (Wearden et al. 1997; Grondin et al. 2004). The decision to count would be beneficial when the sum of the variance associated with each (smaller) segment is lower than the variance associated with the duration in its entirety, and this becomes increasingly the case as duration lengthens. Studies which have not explicitly controlled for counting could therefore find violations in the scalar property as a result.

Other researchers posit that the brain uses two distinct internal clock mechanisms for processing short (e.g. sub-second intervals) and long (e.g. supra-second intervals) durations. If the quantity of variance associated with each stage of processing (i.e. clock, memory, decision) is different for each clock, the combined variance associated with the 'long clock' may not readily exhibit the scalar property. This could explain why the relationship appears to break down beyond ~1000ms. Support for this "distinct timing" theory comes from a variety of studies (Rammsayer and Lima 1991; Rammsayer 1999; Lewis and Miall 2003b; Wiener et al. 2010). However, if two distinct clock

mechanisms did exist, a shift in observer performance should be seen representing the point at which duration processing switches from one mechanism to the other (referred to as a “break point”). Despite measuring the CV across a wide range of durations spanning 68ms to over 16 minutes, Lewis and Miall (2009) failed to find a significant break point to evidence this switch.

Despite conflicting evidence, close approximations to the scalar property have been demonstrated across large swatches of the human and non-human sub-second timing literature (e.g. Rousseau et al. 1983; Wearden and McShane 1988; Ivry and Hazeltine 1995; Grondin et al. 2001; Leon and Shadlen 2003; Merchant et al. 2008; Piras and Coull 2011; Dolores de la Rosa and Bausenhart 2013), with the result that it is generally still accepted as a necessary requirement for any successful model of time-dependent behaviour.

2.4.2 Non-temporal magnitude

Adjusting the intensity of a stimulus can cause a change in its perceived duration. For example, increases in stimulus brightness (Brigner 1986), size (Mo and Michalski 1972; Xuan et al. 2007), complexity (Schiffman and Bobko 1974), numerosity (Mo 1975; Dormal et al. 2006) and sound intensity (Zelkind 1972; Goldstone et al. 1978) are positively correlated with increases in perceived duration.

In internal clock models of temporal processing (Treisman 1963; Gibbon et al. 1984), all of these distortions could be explained by modulating a putative pacemaker's pulse emission rate, perhaps via changes in the observer's state of arousal or attentional focus (Treisman 1963; Penton-Voak et al. 1996; Matthews et al. 2011). This would result in more (or fewer) ticks collecting over a given duration and a subsequent distortion of duration perception based on these accumulated ticks. For example, if the number of ticks corresponding to 500ms is counted when the pacemaker is running at a 'normal' baseline rate, the system will remember the number of accumulated ticks and store this value in reference memory. If stimuli of greater non-temporal magnitude then cause an increase in pacemaker rate, this would cause the accumulator's rate of tick accrual to increase. The same 500ms interval will now see a greater number of ticks being passed to working memory. When this number is compared to stored durations within reference memory, a match will be found at say 550ms, and thus a perceptual distortion occurs where perceived duration undergoes an expansionary bias of 50ms.

However the notion that non-temporal magnitude can directly alter the pacemaker rate through increased attention has recently been contested using a reproduction task, in which observers first viewed a 'target' stimulus of fixed duration and then attempted to reproduce its duration (the reproduced interval) (Rammsayer and Verner 2015). For half of the trials within a block the target interval contained a filled circle which subtended either 1.2° or 10° of visual angle (randomly interleaved), and the reproduced interval contained a fixation cross. In the remaining trials the target interval

contained the cross and the reproduced interval contained the circle of varying size (see Figure 2.16). Trials were presented in a random order.

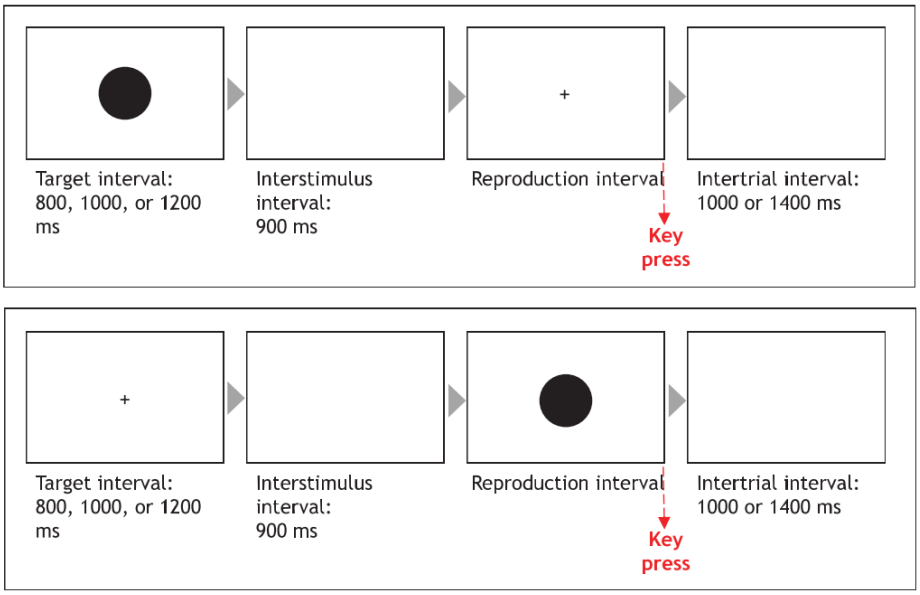

Figure 2.16: Schematic from the study by Rammsayer and colleagues (2015) demonstrating the two experimental conditions. The upper and lower panels show the experimental conditions in which stimulus size (either 1.2° or 10° of visual angle) was varied in the target interval and reproduction interval respectively. Note that in this schematic only the large 10° circle size is shown for simplicity. Figure reprinted with permission from ARVO.

According to internal clock models, the number of pulses corresponding to the initial target interval would be stored in memory during the first part of the task. Then, during the reproduced interval, the number of pulses entering the accumulator would be compared with the representation in memory, and the reproduction ends when a match occurs. If stimulus size directly affected the pacemaker rate, (and thus pulse accrual), it should cause a change in perceived duration regardless of whether size is experimentally varied in the target or reproduced interval (as both involve relative differences in pacemaker rate). Rammsayer and Verner found evidence to the contrary,

such that durations in the range of 800 – 1200ms were on average reproduced 3.2% longer for large 10° stimuli (compared to small 1.2° stimuli), but only when size was varied in the target interval. In addition, observer attention was examined using two groups of observers: one attending to changes in the target size and the other to changes in target shape (Rammsayer and Verner 2015). The shape group acted as an experimental control, as observers were still required to attend to the stimuli but there was no variation in magnitude. Reproduced durations were made using a different key press depending on the observed stimulus' size or shape. There was no significant difference in the reproduced durations between the two groups in any of the experimental conditions, suggesting that the observed effect of stimulus size on perceived duration (when size was varied during the target interval) could not be explained by differences in attention/arousal increasing a putative pacemaker rate.

Similar findings have also been described when varying stimulus numerosity between the target and reproduced interval (Cai and Wang 2014). As non-temporal magnitude only appears to bias perceived duration when it is varied during the encoding phase, both aforementioned studies point to the memory stage of the internal clock as an explanation of their data. It is suggested that as the duration is transferred from the accumulator it becomes biased by non-temporal magnitude, leading to a distortion of the representation in reference memory. Varying stimulus size in the reproduction interval doesn't result in a similar perceptual bias as this value is merely compared to previous values of the target interval (which is now a constant size on every trial) without being transferred to reference memory.

An alternative explanation is that non-temporal magnitude has no direct effect on an underlying temporal processing mechanism, but instead reflects an experimental decision bias. If the brain uses a common metric for magnitude across a range of perceptual dimensions, humans may be predisposed to associating increases in magnitude in one dimension with increases in another dimension. Descriptions of the ‘kappa’ effect, where empty intervals marked by brief visual flashes have a longer perceived duration when the markers are more spatially disparate (Cohen et al. 1953), indicate that magnitude in the dimensions of space and time do interact. Expanding upon this further, “A Theory of Magnitude” (ATOM) argues that the brain uses a common metric for measuring space, time and quantity (Walsh 2003). For example, higher numbers (e.g. 10) would be perceived to be larger or brighter or longer in duration etc. than lower numbers (e.g. 3), and larger stimuli would be perceived to be louder or longer in duration than smaller stimuli etc. Observers performing a comparative duration discrimination task (i.e. “which was longer, the first or second stimulus?”), could therefore bias their responses in favour of the stimulus with the greatest (non-temporal) magnitude.

This strategy would be particularly difficult to detect in studies that use a performance measure (e.g. temporal discrimination threshold or percentage of correct responses), as both responding to the non-temporal attribute and

Figure cannot be displayed due to copyright law. To view the original figures please see Xuan et al. (2007), Figures 1&2 (p2 & p3).

Figure 2.17: Data from a duration discrimination task in which observers had to decide "which was longer, the first or second duration?" (Xuan et al. 2007). Four standard durations were used throughout the experiment (600ms, 650ms, 700ms and 750ms), and these were always paired with a comparison duration that was 1.25 times longer than the corresponding standard. Observers were told to ignore the visual pattern of the stimuli. **a)** Non-temporal magnitude was varied in four ways: number of dots (dot #), size, brightness (lum) and numerosity (digit). These were then divided into "small" and "large" depending on magnitude. **b)** Performance on the duration discrimination task shown as error rate (percentage of incorrect responses). Congruent conditions refer to the shorter standard duration being paired with a "small" stimulus (and longer duration paired with a "large" stimulus). Incongruent conditions refer to the opposite (short duration with "large" stimulus and long duration with "small" stimulus). Performance was significantly worse during incongruent trials, suggesting that non-temporal magnitude was influencing the perceived duration of the stimuli.

genuine perceptual bias would reduce task performance. For example, Xuan and colleagues (2007) investigated the effect of non-temporal magnitude on performance during a duration discrimination task (see Figure 2.17). They used four standard durations (ranging from 600 – 750ms) and paired each one with a comparison stimulus that was 1.25 times longer in duration.

Observers were tasked with deciding “which was longer, the first or second duration?” During congruent trials the short duration had a “small” non-temporal magnitude and the long duration a “large” non-temporal magnitude, and during incongruent trials the reverse was true. Significantly more errors were made on incongruent trials across all four types of non-temporal magnitude, which the authors interpreted as evidence that non-temporal magnitude influenced perceived duration (i.e. “large” stimuli caused a perceptual expansion of the short durations, and “small” stimuli a perceptual compression of the long durations, making discrimination between the two more difficult). However the same pattern of results could have arisen from the observers responding to the non-temporal magnitude rather than from changes in perceived duration. Since perceptual biases were not measured during the study, it is impossible to determine which strategy the observers were using.

The issue of decision type in Xuan et al.’s study was also addressed by Yates et al. (2012), who suggested that the use of a comparative judgment could make it more likely that observers would be influenced by non-temporal magnitude when they are uncertain. Initially they replicated the results of Xuan and colleagues using a comparative judgment (“which was longer?”). However when the experiment was repeated using an equality judgement (“are the two durations the same or different?” – in which the response options are no longer bound to one particular stimulus), the opposite result was found: larger stimuli now had a slightly *shorter* perceived duration than smaller stimuli. The authors suggest that this result may reflect an increased likelihood for observers to respond ‘same’ in congruent trials

(i.e. small square with short duration, large square with long duration), as both the spatial and temporal features of the stimulus are complimentary in this condition. Certainly the results of this study indicate that at least one of the two duration judgements was affected by a decisional bias, and therefore careful considerations should be made when designing duration judgment tasks. Nevertheless, if duration does form a common perceptual metric with quantity and space, this implies that any underlying temporal processing mechanism must have access to signals arising in these other dimensions, and perhaps is more likely to reside in a central location.

2.4.3 Sensory modality

A large proportion of the temporal processing literature is devoted to the comparison of temporal estimates between different sensory modalities, particularly focusing on vision and audition. One robust temporal phenomenon is the “sound longer than vision” bias (see Figure 2.18), in which auditory stimuli are perceived to be longer in duration than visual stimuli of the same physical duration (Goldstone and Goldfarb 1963; Goldstone and Lhamon 1974; Walker and Scott 1981; Wearden et al. 1998; Penney et al. 2000; Ulrich et al. 2006a; Wearden et al. 2006).

Internal clock explanations of these sensory differences again tend to centre on the speed of a hypothetical pacemaker, which is proposed to tick at a faster rate for auditory stimuli (Wearden et al. 1998; Penney et al. 2000; Wearden et al. 2006). If more pulses accumulate during an auditory interval than a visual interval, the auditory interval will seem longer when the two are

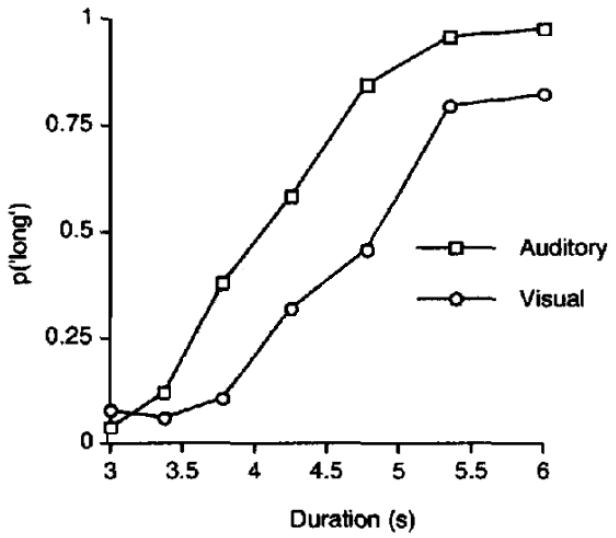

Figure 2.18: Data from a study by Penney *et al.* (2000) demonstrating the sound longer bias. In a training phase, observers were trained to recognise two anchor durations, 3 seconds ('short') and 6 seconds ('long'), presented randomly in both auditory and visual modalities. During a subsequent test phase they were presented with single durations (sampled from a range of durations falling between these 'short' and 'long' durations), and asked to judge whether they were closer to the short or long duration. The plot above shows the proportion of 'long' judgements as a function of physical test duration, demonstrating that stimuli of the same physical duration were more likely to be classified as 'long' when presented in the auditory modality compared to the visual modality. Figure reprinted with permission from the American Psychological Association.

compared. If perceived (subjective) duration is plotted against physical (objective) duration, this would manifest as a difference between the slopes of auditory and visual estimates (see Figure 2.19a). A competing explanation

for modality differences is that the switch lying between the pacemaker and accumulator (see Section 2.2 Figure 2.4) opens and closes with a different latency for visual and auditory signals (Wearden et al. 1998). This would be independent of duration magnitude, and would therefore manifest as an additive difference between visual and auditory estimates (i.e. the intercepts would differ – see Figure 2.19b).

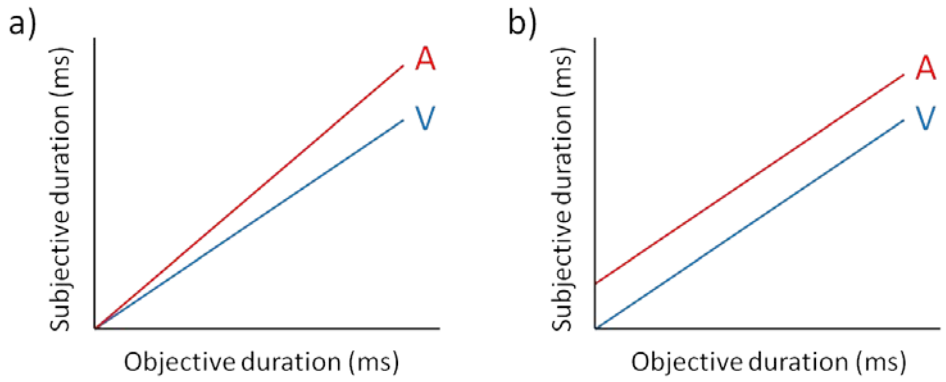

Figure 2.19: Schematic showing perceived (subjective) duration against physical (objective) duration for auditory (A) and visual (V) modalities under two different encoding scenarios. **a)** The pacemaker ticks at a faster rate for auditory stimuli than visual stimuli, manifesting as a difference in the slopes. **b)** The pacemaker ticks at the same rate for both modalities, but the switch latency differs between auditory and visual stimuli (note that for illustrative purposes the visual estimates are shown to be veridical in the above schematic, which is not necessarily representative).

To investigate these two possibilities, Wearden and colleagues (1998) presented observers with a variety of both visual and auditory durations (randomly intermixed within a block) and asked them to verbally estimate each duration in milliseconds. The mean estimates were then plotted against the physical durations for both modalities, allowing a comparison of both

functions. The results showed a significant difference between the slopes but not the intercepts, suggesting that it was the pacemaker rate and not switch latency that accounted for the sound longer bias (see Figure 2.20).

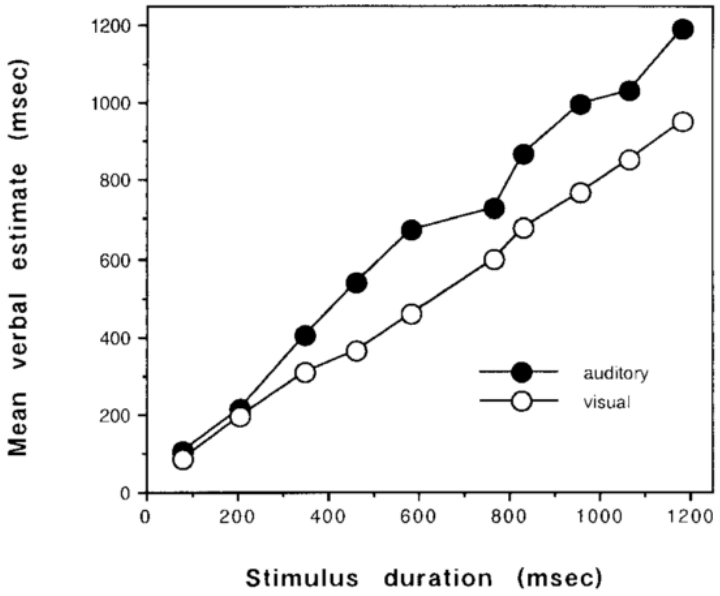

Figure 2.20: Data from Wearden and colleagues (1998) showing verbal estimates of duration plotted against physical duration. The slope of the function is related to the rate of the putative pacemaker, whilst the intercept gives an indication of switch latency. In the example above the slope for audition is steeper than the slope for vision, suggesting that the 'sound longer' bias is due to a faster auditory pacemaker rate (Wearden et al. 1998). Reprinted with permission from Taylor and Francis (http://www.tandfonline.com/)

Penney et al. (2000) went on to propose that both pacemaker rate and the mixing of duration representations in reference memory were responsible for the sound longer bias, as they found it only arose when both auditory and visual stimuli were presented within the same experimental session. Assuming that auditory stimuli cause the clock to tick at a faster rate than visual stimuli, a greater number of pulses would accrue in response to the same physical duration. Therefore if both auditory and visual estimates

contribute to a ‘supramodal’ representation of the duration in reference memory, the peak of this distribution would correspond to the average of both auditory and visual values (a ‘mixed’ value - see Figure 2.21). If an auditory (visual) duration is then presented in isolation at a later stage, the value in the accumulator will contain more (less) pulses when compared to

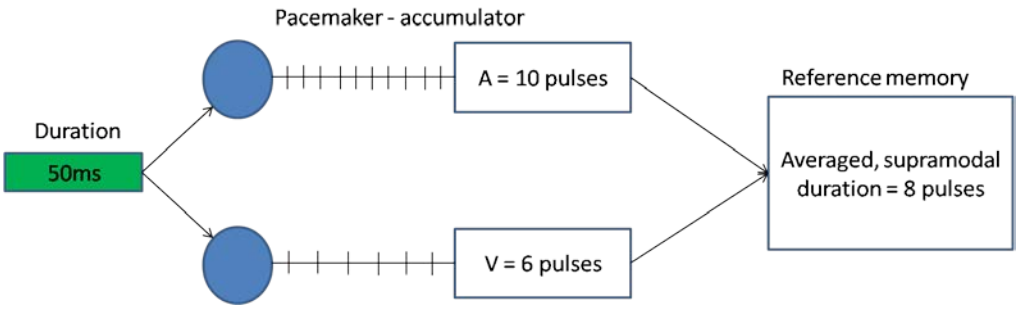

Figure 2.21: Schematic showing the concepts of memory mixing. A 50ms duration presented in the auditory modality (A) results in 10 pulses entering the accumulator. Comparatively, the pacemaker runs at a slower rate when the same 50ms duration is presented in the visual modality (V), such that only 6 pulses enter the accumulator. Yet both of these values contribute to the same shared distribution of the duration in reference memory, the peak of which is an average of both modalities.

the *average* value stored in memory, resulting in a perceptual expansion (compression) of the auditory (visual) stimulus. However, a later report that the ‘sound longer bias’ can arise between different experimental groups during a verbal estimation task (in which one group estimated only a series of visual standard durations and another group estimated only auditory standard durations) suggests that the ‘memory mixing’ account may not be appropriate in all cases (Wearden et al. 2006).

Another well documented characteristic of inter-sensory timing is the finding that auditory duration discrimination thresholds are significantly lower than visual thresholds (Tanner et al. 1965; Grondin and Rousseau 1991; Grondin et al. 1998; Mattes and Ulrich 1998; Wearden et al. 1998; Ulrich et al. 2006a; Ortega et al. 2009; Stauffer et al. 2012). Internal clock models account for this by proposing that 'modality-specific' switches are activated by auditory and visual stimuli, and the visual switch operates with greater variability (Rousseau and Rousseau 1996; Wearden et al. 1998). This would subsequently cause greater variation in the number of pulses entering the accumulator, resulting in a noisier representation of each visual duration and poorer discrimination ability.

However, modality differences in both perceived duration and discrimination thresholds also lend support to localised, sensory-specific models of temporal processing (see Section 2.3). These posit that duration is encoded at a relatively peripheral neural location, and differing temporal sensitivity could result from inherent variations in neural processing within each sensory cortex. For example, the distance from the retina to the primary visual cortex is physically greater than from the cochlea to the primary auditory cortex, creating a longer 'pathway' over which the visual signal must travel. This increases the opportunity for the signal to become degraded or more variable, potentially resulting in a lower signal-to-noise ratio. Electrophysiological studies have indicated shorter processing latencies in response to an auditory stimulus compared to a visual stimulus (Perrault and Picton 1984; Vidal et al. 2008), and auditory reaction times are known to be

shorter than visual reaction times (Shelton and Kumar 2010; Jain et al. 2015). Therefore faster, more accurate auditory processing could account for superior temporal sensitivity in the auditory modality.

Yet independent, modality specific timing mechanisms are difficult to reconcile with reports that temporal learning can demonstrate crossmodal transfer (Warm et al. 1975; Bartolo and Merchant 2009; Bratzke et al. 2012). For example, variability in duration reproduction estimates for an auditory task using durations centred on 450ms, 650ms and 850ms may be reduced by intensive training, and this increased sensitivity can be carried over to untrained durations in the visual modality (Bartolo and Merchant 2009). This would imply that visual and auditory timing relies on a common supramodal resource (e.g. reference memory), and these stored duration estimates are improved by the auditory training. Alternatively, it may be explained by a hierarchical system such as the example proposed by Stauffer and colleagues, in which duration is initially processed in a modality specific manner before being passed to a supramodal processing system downstream (Stauffer et al. 2012). Nevertheless, caution should be taken to separate out genuine intermodal transfer from task learning effects, as repeated training on a task can increase precision between pre-training and post-training stages independent of any transfer effects (Lapid et al. 2009).

2.4.4 Temporal adaptation

Our perception of time can be distorted by our recent sensory history. For example repeated presentations of a short/long duration can cause a

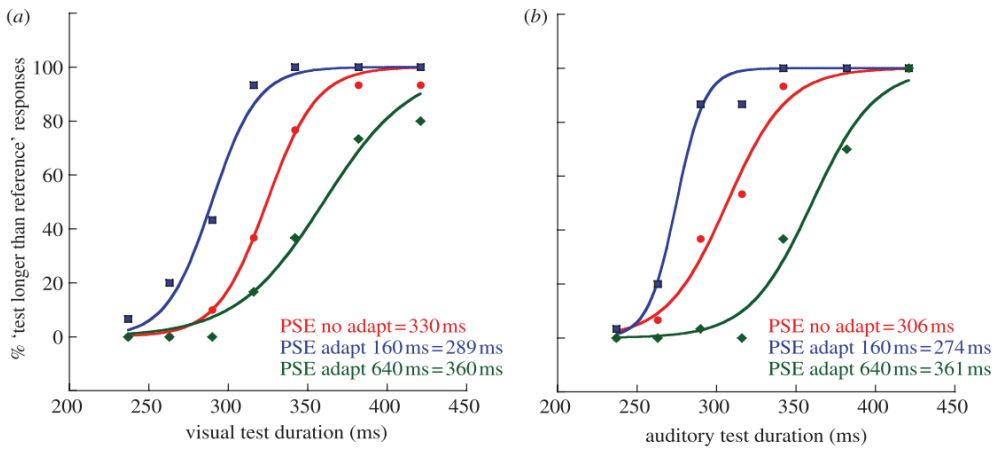

Figure 2.22: Data from a representative observer in the study by Heron et al. (2012) demonstrating the bidirectional duration aftereffect following adaptation to short (adapt 160ms) and long (adapt 640ms) durations. Here, observers adapted to repeated presentations of a visual (a) or auditory (b) duration before performing a duration discrimination judgment in which they had to decide "which was longer?" between a 320ms reference duration in the opposite modality, or a variable duration test stimulus in the same modality. The above functions plot the number of 'long' responses as a function of physical test duration. Differences in the PSE in the 'no adapt' condition (red data) reflect the 'sound longer' bias (Wearden et al. 1998) – see Section 2.4.3. Adapting to short durations (blue data) caused an increase in the number of 'test' longer durations, and a leftward shift of both functions. Figure 2.22a shows that in the 'short' condition the PSE was reduced to 289ms, representing a temporal expansion: test stimuli had to be shortened to 289ms to feel perceptually equivalent to the 320ms reference. The opposite is true for adapting to long durations (green data), where test stimuli had to be expanded to 360ms in order to feel equivalent to 320ms. Figure reprinted with permission from the Royal Society.

temporal expansion/compression of a subsequently presented moderate duration (Walker and Irion 1979; Walker et al. 1981; Allan 1984; Heron et al. 2012) (see Figure 2.22). This phenomenon is referred to as 'duration adaptation', and results in a 'duration aftereffect'. Adaptive aftereffects in the temporal domain were introduced in Section 2.3.3 of this thesis, and have been documented for both duration (Johnston et al. 2006; Burr et al. 2007;

Heron et al. 2012; Heron et al. 2013; Latimer et al. 2014; Li et al. 2015b) and temporal rate (Recanzone 2003; Becker and Rasmussen 2007; Levitan et al. 2015).

Duration aftereffects may also be contingent on other, non-temporal stimulus features such as auditory pitch (Walker and Irion 1979; Li et al. 2015b). For example, after adapting to alternative presentations of 600ms at 600Hz (low pitched tone), and 200ms at 900Hz (high pitched tone), a high pitched tone of 400ms was perceived to be longer than a low pitched tone of physically equal duration (Walker and Irion 1979). These results indicated that the underlying neurons (suggested to be located within the auditory system) were tuned to more than one stimulus feature: duration *and* auditory pitch, since the net effects of adaptation would otherwise cancel out. Neurons tuned to pitch may be located early in the auditory pathway (Kelly and Beaver 1991; Casseday and Covey 1992), suggesting that duration signals could also be extracted at an early (pre-cortical) stage of processing. Given the hierarchical arrangement of feature processing in the early auditory (and visual) cortices (Hubel and Wiesel 1968; Merzenich and Brugge 1973; Gattass et al. 1981; Rauschecker et al. 1997; Kaas and Hackett 2000), examining the selectivity of the duration aftereffect to a range of non-temporal features may help to pinpoint a possible neural locus for duration encoding.

The concept of ‘localised timing’ is further supported by evidence that duration aftereffects do not transfer across the sensory modalities (Walker et al. 1981; Heron et al. 2012), and in the visual system have been shown to arise prior to multisensory integration (Heron et al. 2013). In addition, Li and

colleagues (2015b) demonstrated that two independent duration aftereffects could be elicited in different modalities simultaneously. Observers adapted to alternate presentations of 160ms visual durations and 640ms auditory durations before judging whether a subsequently presented (variable) test duration (either auditory *or* visual – mixed within a block) was “long or short?” Measurable duration aftereffects of differing magnitudes occurred in both modalities, implying that two independent, local mechanisms (rather than one common, supramodal clock) were driving the aftereffects.

Recently, research in the visual domain has focused on the spatial selectivity of temporal adaptation, utilising the systematic increases in receptive field size from pre-cortical through to extrastriate brain regions (Hubel and Wiesel 1968; Gattass et al. 1988; Logothetis et al. 1995; Xu et al. 2001) to infer a possible neural locus for duration encoding. Initial explorations showed that adapting to a continuously presented grating drifting at 20Hz caused the perceived duration of a 10Hz grating to be temporally compressed when both the adapt and test stimuli were presented at the same spatial location (Johnston et al. 2006).The aftereffect is unidirectional: 5Hz adapting stimuli also induced temporal compression, albeit of a reduced magnitude. A follow up study found that this compression was tightly tuned to retinotopic spatial locations within 1° of the adapting stimulus, consistent with an early neural locus (Ayhan et al. 2009). Additional reports that the aftereffect was not contingent on stimulus orientation (Johnston et al. 2006), and did not display interocular transfer (Bruno et al. 2010), led to the suggestion that the aftereffect may have arisen within the magnocellular pathway at a pre-cortical location such as the lateral geniculate nucleus (LGN). This was also

consistent with the finding that the aftereffect survived adaptation to perceptually invisible flicker of approximately 60Hz, which is beyond the high temporal frequency cut-off of cortical neurons but detectable by magnocellular neurons in the LGN (Johnston et al. 2008).

However, a pre-cortical locus has been contested by evidence that these compressive aftereffects only arise when both the adapt and test stimuli are presented at the same spatiotopic coordinates, suggestive of a higher cortical location such as extrastriate area V5 (Burr et al. 2007). Additionally Curran and Benton (2012) used random dot kinematograms and drifting plaids with a temporal frequency of $3°s^{-1}$ to demonstrate that this compressive aftereffect is contingent upon both the adapt and test stimuli moving in the same direction. This again pointed to a cortical locus for duration encoding, as the earliest location of direction selective neurons is the primary visual cortex (Hubel and Wiesel 1968).

Importantly, these studies use flickering or drifting adapting stimuli which vary in one temporal metric (temporal frequency) to create a bias in another temporal metric (perceived duration), producing a duration aftereffect which is unidirectional: resulting in temporal compression. In contrast, adaptation to stimuli of fixed duration results in a duration aftereffect which is bidirectional, repulsive (i.e. adapting to short/long durations leads to a perceptual expansion/compression of a subsequently presented medium duration) and tuned around the duration of the adapting stimulus (Walker et al. 1981; Heron et al. 2012: see also Figure 2.23). Thus it is suggested that these two aftereffects are independent phenomena, and may be generated by separate temporal mechanisms, potentially at different neural locations.

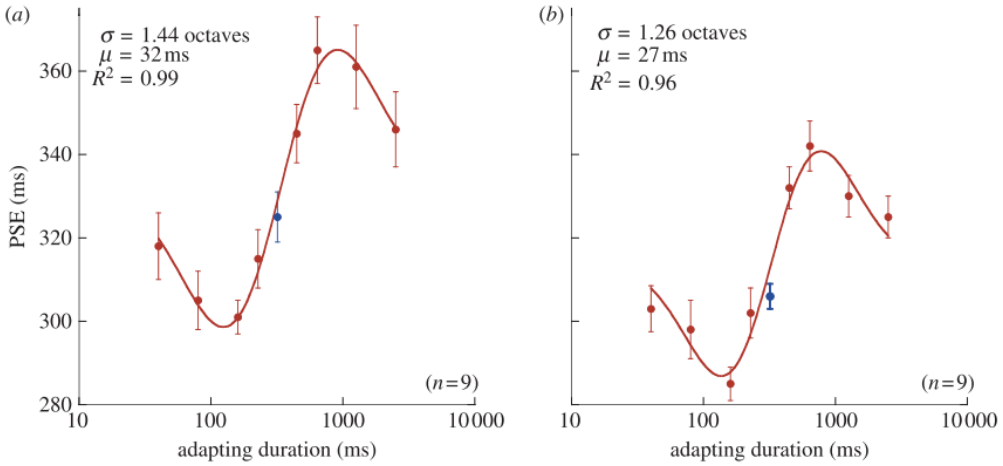

Figure 2.23: Data from the study by Heron and colleagues (2012). Perceived duration (PSE) of a 320ms test stimulus is plotted without adaptation (blue data) and following visual (a) or auditory (b) adaptation to a range of durations from 40 – 2560ms (red data). The value μ gives the amplitude (height of the peaks/troughs), and σ (the standard deviation of the function) gives a measure of aftereffect bandwidth. For both modalities the duration aftereffect is shown to be bidirectional in nature, and bandwidth tuned. Figure reprinted with permission from the Royal Society.

That being said, any adaptive aftereffect which is restricted to the spatial location occupied by the adapting stimulus is difficult to reconcile with a single, central dedicated mechanism that pools inputs across visual space. Adaptation-induced changes in the function of a centralised mechanism would presumably influence all subsequent judgments that depended on said mechanism, regardless of spatial location. Similarly the perceptual bias induced by adaptation cannot be attributed to a global change in cognitive factors such as attention as there is no reason to see why this would induce bandwidth-limited repulsive effects.

2.4.5 Attention and time

Anecdotally, time often appears to slow down when it is attended to (Mattes and Ulrich 1998), just as it appears to move faster when attention is directed elsewhere (Brown 1997). The popularity of the adage “time flies when you’re having fun” is supported by the fact that associations between attention and time form a particularly robust area of the temporal processing literature. Experiments in which observers are told beforehand that the task involves a duration judgement (i.e. a prospective paradigm) often invoke selective or “endogenous” attention, in which the observer voluntarily directs their attention to time. Under these conditions, duration estimates are longer and less variable than when observers were unaware that a task required a temporal judgement, and are asked retrospectively to judge elapsed duration (Hicks et al. 1976; Brown 1985; Block and Zakay 1997).

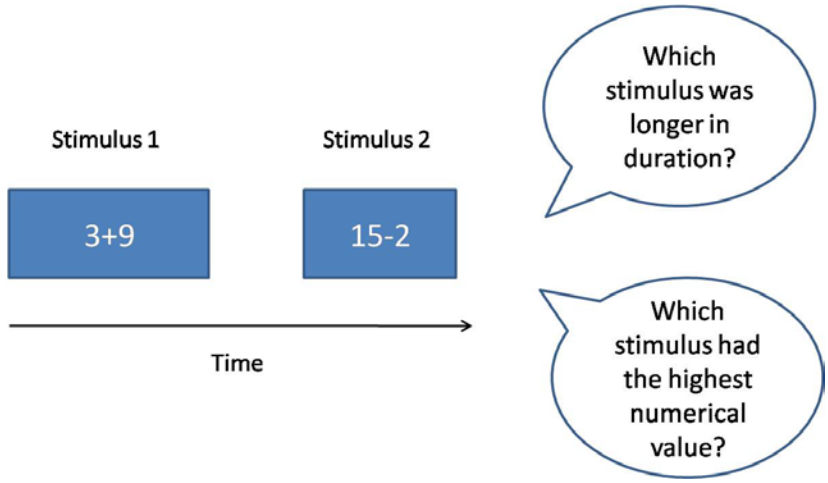

Figure 2.24: Schematic of an interference task involving mental arithmetic. Observers must divide their attention between the two tasks in order to judge which stimulus had the longest duration and also calculate which had the highest numerical value.

A popular method to study the effects of attention on subjective duration is to divide attention between a timing task and a secondary (non-temporal) task such as mental arithmetic (Rammsayer and Ulrich 2005) (see Figure 2.24). These interference tasks are suggested to work by increasing the cognitive load (amount of information processing required) over a specified time period. Typically, the effect of the secondary 'distracter' task sees an observer's duration estimates become shorter and more variable (Brown 1997; Brown 2006). Assuming that the brain has a finite amount of processing resources, the 'attentional allocation model' proposes that increased allocation of resources to the interference task comes at the cost of fewer resources allocated to duration processing (Thomas and Weaver 1975; Brown and West 1990; Zakay 1993). This could disrupt the operation of a putative clock, such that some pulses are missed or not processed accurately, leading to an incomplete and unreliable record of the temporal event. If non-temporal and temporal processes compete for the same resources, it then follows that performance on the secondary (non-temporal) task should also decrease in split attention conditions. Such 'bidirectional interference' has been evidenced, particularly if the secondary task requires comprehension, mental arithmetic or sequencing (Brown and Merchant 2007). Given that these 'executive' functions are often attributed to high-level processing in the frontal cortex (Stuss and Benson 1986; Jurado and Rosselli 2007), it implies that this brain region may play a role in temporal processing.

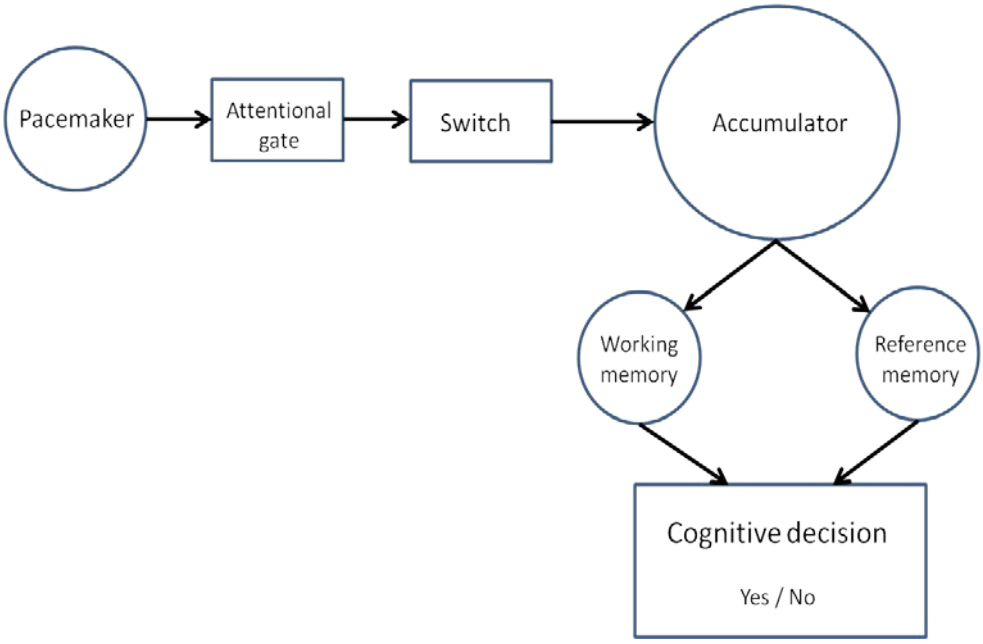

Figure 2.25: Schematic of the 'Attentional Gate Model' model proposed by Zakay and Block (1995)

Zakay and Block more formally integrated the 'attentional allocation model' into an internal clock framework with their 'attentional gate model' for prospective time judgements (Zakay and Block 1995; Block and Zakay 1996: see Figure 2.25). This model proposes that an attentional gate exists between the pacemaker and switch that controls the flow of pulses into the accumulator. The more attention is directed to timing, the wider the gate opens, allowing a greater number of pulses to pass through from the pacemaker. The switch then operates independently in an all-or-nothing manner, responding only to the onset/offset of a temporal event (i.e. the external signal). However, it has been questioned whether an additional component to the original clock model is actually necessary, or whether the activity of the switch itself could be modulated by attention (Lejeune 1998).

In the sub-second range, interference tasks may only influence perceived duration under particularly cognitively demanding conditions such as when

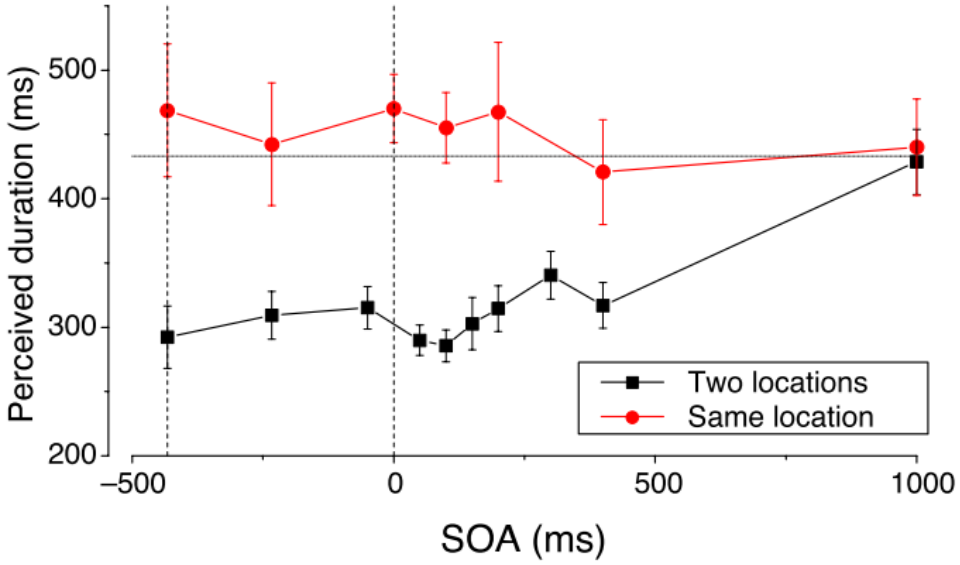

Figure 2.26: Showing data from the study by Cicchini and Morrone (2009), in which observers performed a duration discrimination task between a 433ms 'target' and a variable 'probe' duration. Durations were empty intervals marked by visual flashes, which could either appear at two spatial locations (black data) or the same spatial location (red data). Perceived duration is plotted as a function of stimulus onset asynchrony (SOA) between the duration task and a concurrent distracter task. Temporal compression of the target is found only when attention was spatially divided during the task. Figure reprinted with permission from ARVO.

spatial attention is concurrently manipulated (Cicchini and Morrone 2009). Cicchini and Morrone designed a dual-task procedure in which observers performed a motion discrimination task (the interference task), whilst also judging the duration of an empty 'target' interval (433ms) delineated by two brief visual markers (3x30° horizontal red bars each of 17ms duration – designed to minimise any apparent motion cues). These markers could either be presented at the same spatial location, or different spatial locations (with 24° vertical separation between the bars), and the asynchrony (SOA)

between the duration discrimination task and the interference task was varied. Across a range of SOAs spanning approximately ±500ms a perceived temporal compression of the target interval was found, but only when the visual markers were spatially separated (see Figure 2.26). Visual markers presented in the same spatial location resulted in near veridical duration perception. The authors interpreted this as evidence that multiple, spatially selective clocks are initially utilised during duration encoding, and attentional modulation occurs at a later stage of processing when comparing the output of these clocks. If the same cognitive/memory resources are common to both the duration task and the motion discrimination task, and monitoring the output of two 'clocks' (each responsible for a different spatial position) generates a greater cognitive load than monitoring one clock/spatial position, there will be even greater competition for resources in the spatially separated condition. Subsequently, monitoring the output of multiple clocks will be less accurate, resulting in a shortening of perceived duration.

Rather than attempt to reduce the attentional resources allocated to a duration task, it is also possible to direct attention towards a particular stimulus. The presentation of a novel stimulus has been shown to cause an involuntary or "exogenous" attentional shift towards the stimulus (Remington et al. 1992). This attentional shift is used in "oddball" paradigms, which study the effects of a novel or "oddball" stimulus occurring within a sequence of identical reference stimuli (Tse et al. 2004; Seifried and Ulrich 2010) (see Figure 2.27). The oddball is frequently perceived to have a longer duration than the other stimuli in the sequence, in line with the idea that the capture of exogenous attention by a stimulus causes an expansion in its perceived

duration. This concept could also explain why the first of a sequence of identical visual stimuli is temporally overestimated (Rose and Summers 1995). In traditional pacemaker models the appearance of an unexpected or "oddball" stimulus is thought to increase observer arousal, and heightened

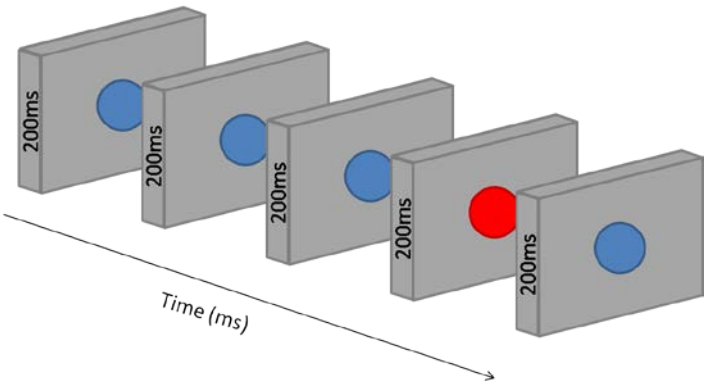

Figure 2.27: Schematic of an oddball paradigm in which the oddball is the red circle within a train of blue circles.

arousal then increases the rate of the pacemaker (Treisman 1963; Treisman et al. 1990; Ulrich et al. 2006b; Seifried and Ulrich 2010). This leads to a greater number of pulses entering the accumulator and thus a subjective expansion of time.

An opposing theory argues that repeated presentations of the non-oddball stimuli leads to a contraction of their perceived duration, so that they appear relatively shorter than the (veridical) oddball (Eagleman and Pariyadath 2009). The "neural energy model" proposes that perceived duration is proportional to the amount of neural energy expended (Pariyadath and Eagleman 2007; Pariyadath and Eagleman 2008: see Figure 2.28). When a stimulus is repeatedly presented, the neural firing rate in response to this stimulus reduces (known as 'repetition suppression') (Fahy et al. 1993; Grill-

Spector et al. 2006; Summerfield et al. 2008). If neural energy is proportional to perceived duration, the duration of the stimulus also becomes subjectively compressed. Then, if a novel stimulus is presented, a 'normal' neural response is elicited and thus its duration is perceived veridically. However, because the neural activity in response to the novel stimulus is relatively

Figure cannot be displayed due to copyright law. To view the original figure please see Eagleman and Pariyadath (2009), Figure 1 (p1842).

Figure 2.28: Predictions of the 'neural energy model' (Eagleman and Pariyadath, 2009). All stimuli have the same physical duration. The repeated presentation of the strawberry results in repetition suppression, in which the neural response to the stimulus declines over time. This results in a subsequent compression of perceived duration. When the novel stimulus (the car) is presented, neural activity (and perceived duration) rises again to the original level, such that the duration of the car stimulus feels longer than the duration of the strawberry.

greater, it is perceived to be longer in duration. Irrespective of the underlying mechanism, an important criticism of oddball paradigms is that the illusion is very susceptible to the choice of stimulus. For example, oddballs characterised by grossly suprathreshold changes in low-level stimulus characteristics may fail to generate an expansion in perceived time (Aaen-Stockdale et al. 2011).

Manipulating an observer's emotional state has also been linked to distortions of temporal processing. For example, negative sounds (e.g.

crying) are judged longer than positive sounds (e.g. laughing), and in general emotional sounds are judged longer than neutral sounds (e.g. street noises) (Noulhiane et al. 2007). In a similar experiment using vision, “unpleasant” pictures were judged to be longer than “pleasant” pictures when chosen from the 6.5 – 7.5 ‘high arousal’ rating of the International Affective Picture System (Angrilli et al. 1997). Emotional stimuli are thought to be more arousing, and thus expansions of perceived duration are again attributed to the speeding up of an internal clock (Droit-Volet et al. 2004; Fayolle et al. 2015). However, emotionally salient stimuli are not always judged to be significantly different from neutral stimuli, (Pariyadath and Eagleman 2007), and it has been demonstrated that emotional stimuli have a smaller impact on perceived time in experiments with a larger working memory requirement (e.g. where observers must remember a previously presented reference duration) (Gil and Droit-Volet 2011). It is proposed that under these conditions, storing durations in working memory decreases the ‘emotional contrast’ between stimuli (perhaps due to a larger division of cognitive resources), resulting in smaller differences in perceived duration.

2.4.6 Timing multiple independent intervals

In addition to understanding how the brain might encode the duration of a single event, several studies have focused on the processing of multiple, temporally overlapping events (Brown and West 1990; Van Rijn and Taatgen 2008; Gamache and Grondin 2010; Klapproth 2011; Cheng et al. 2014; Bryce et al. 2015). These tasks often require an observer to monitor several durations of different magnitude and with temporally staggered onsets, and

then either reproduce or make a temporal judgement about one of the durations from the group. As such, they are akin to the ‘interference’ tasks (i.e. “dual-tasks”) described in Section 2.4.5, only rather than utilising a non-temporal interference task such as mental arithmetic, the secondary task now requires a temporal judgement.

Variations of the internal clock have been proposed in an attempt to deal with the challenge, including suggestions of multiple accumulators or multiple pacemaker-accumulator mechanisms which all feed into the same memory

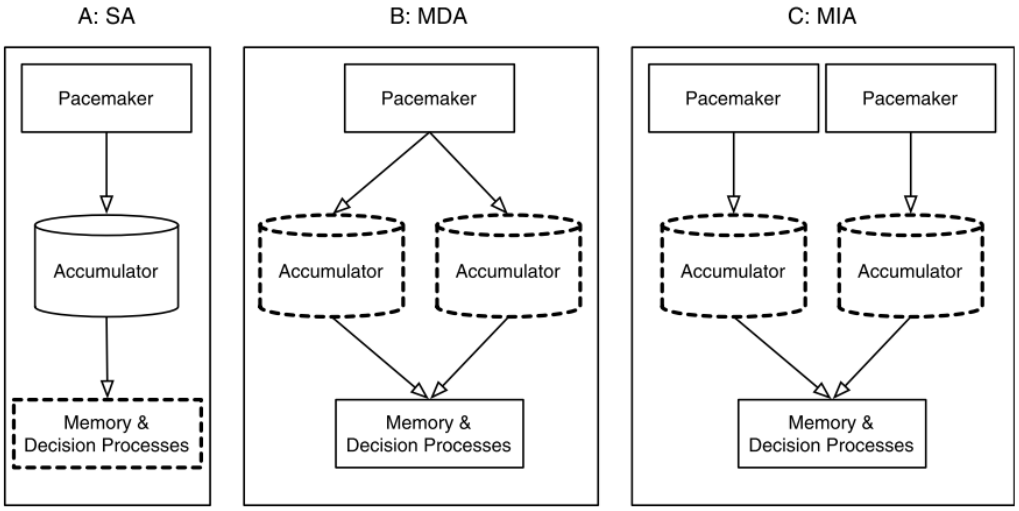

Figure 2.29: Three possible pacemaker-accumulator scenarios. **A)** A single accumulator (SA) model. **B)** A multiple dependent accumulator (MDA) model, where accumulated ticks from separate stimuli are dependent on a single pacemaker but counted separately. **C)** A multiple independent accumulator (MIA) model where multiple stimuli may be timed by independent pacemakers each with their own accumulator. Figure reprinted from (Van Rijn and Taatgen 2008) with permission from Elsevier.

component (Ivry 1996; Rousseau and Rousseau 1996; Ivry and Richardson 2002; Klapproth 2011; Cheng et al. 2014) (see Figure 2.29). However the concept of multiple processing units would require a different processor (or clock) for each event requiring duration estimation. As these events will be

defined by their spatial locations or auditory frequencies etc., a large (potentially infinite) number might be required in real world environments; a scenario which may not be biologically plausible.

A single clock with a single accumulator could instead be capable of processing multiple overlapping durations if these durations could be treated

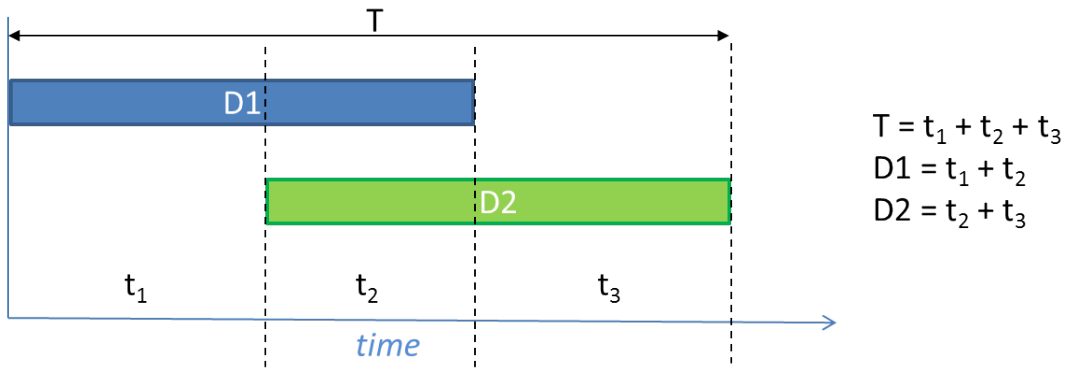

Figure 2.30: Schematic demonstrating how two overlapping durations (D1 and D2) could be treated as a sequence of temporal segments (t_1, t_2 and t_3). The total elapsed time T is given by the sum of these segments, and the individual durations can be determined by summing the relevant segments (some examples of which are shown on the right).

as a sequence of events consisting of several temporal segments (Van Rijn and Taatgen 2008; Bryce and Bratzke 2015; Bryce et al. 2015). For example, two overlapping durations could be split into three segments, composed of the first interval alone, the overlapping section and the second interval alone (see Figure 2.30). Each segment would be processed independently and then the component parts could be summed to give an estimate of each duration. Early behavioural data is reasonably well supported by model predictions (Van Rijn and Taatgen 2008), however this 'temporal arithmetic' would presumably require additional processing demands, and it is not

known how the brain would behave when timing additional (i.e. 3 or more) overlapping intervals particularly in the sub-second range.

Alternatively, several overlapping durations could be processed by local timing mechanisms, where different neural populations/regions encode the duration of each independent stimulus. Within a single modality this would be viable if the underlying duration selective neurons were also selective for non-temporal stimulus features such as auditory pitch, visual orientation or spatial location etc., allowing multiple neural populations to be activated simultaneously. Cheng and colleagues examined the ability to process temporally overlapping visual stimuli across a variety of spatial locations. Based on measures of observer performance they calculated the capacity limit for simultaneous timing to be 3 – 4 spatial locations, or 3 – 4 'local clocks' (Cheng et al. 2014). They also reported no significant correlation between each individual's timing capacity and their visual working memory capacity, suggesting that the limits of temporal processing are not constrained by an ability to store representations in working memory. These results appear to conflict with an earlier study who used a similar paradigm to show that when two temporally overlapping visual durations were presented during a duration discrimination task, observer thresholds showed a large increase when the number of judged durations increased from one to two (Morgan et al. 2008). This was interpreted as an inability for humans to accurately judge the duration of more than one visual event at a time, supporting the notion of a single clock. However, whilst these two studies appear to disagree in their conclusions, this may result from differences in interpretation (e.g. Cheng et al. (2014) chose to adopt a mathematical

formula (Cowan's K) to determine capacity whilst Morgan et al. (2008) reported a trend in observer thresholds), as both studies reported a similar drop in observer performance when timing more than one overlapping interval. This would support the notion that humans struggle to process more than one duration signal simultaneously. Importantly, throughout both tasks observers not only had to monitor multiple overlapping durations but also maintain divided spatial attention throughout the full presentation of the stimuli. This suggests that noisier duration estimates could be the product of competition between attentional and time specific resources, in a similar manner to the increases in observer thresholds witnessed in attention-based 'interference tasks' (see Section 2.4.5).

2.4.7 Integrating multisensory durations

The world is comprised of multiple concurrent sensory inputs, and our ability to navigate through and make sense of the world around us requires that the brain is able to process multiple temporal signals or 'cues' in parallel. We have previously discussed how the brain might process multiple, temporally overlapping durations under 'temporal dual-task' conditions, in which each duration must be monitored independently (see Section 2.4.6). The following discussion will now focus on the brain's ability to integrate temporal signals from multiple sources of sensory information to create a unified percept of duration. This process of combining multiple sensory inputs is often referred to in the literature as 'cue combination'.

In humans, the ability to combine complementary signals certainly has behavioural advantages, and combining incoming sensory information across multiple modalities has been shown to improve sensory estimates. For example, our ability to interpret speech against a noisy background is improved when it is accompanied by visual cues (Sumby and Pollack 1954; Ross et al. 2007) and detection thresholds for weak visual signals are enhanced by simultaneous bursts of auditory white noise at the same spatial location (Frassinetti et al. 2002).

Performance on a motion perception task also improves significantly faster when observers are trained with audio-visual stimuli compared to visual stimuli alone (Seitz et al. 2006). Observers were trained to perform a visual motion detection task across 10 days ("did the first or second interval contain the directional motion?"). One half of the group received a mixture of audio-visual, auditory and visual stimuli in separate trials, whilst the others received only visual stimuli. Performance from 'visual stimulus only' trials was measured each day (percent correct) for each group, and the multisensory group were found to improve significantly faster than the vision alone group, suggesting that our ability to learn new tasks is increased with multisensory input (Seitz et al. 2006).

One robust finding from the multisensory spatial integration literature is the 'ventriloquist illusion', where perception of the spatial location of an auditory source is biased towards concurrently presented visual information (Pick et al. 1969; Warren et al. 1981; Alais and Burr 2004). The illusion is driven by the ventriloquist's ability to synchronise the dummy's mouth movements with their own speech. This temporal correspondence between sound and vision facilitates 'visual capture' of perceived auditory location - the spatial origin of the ventriloquist's voice appears to coincide with the visual location of the

Figure 2.31: Example of a ventriloquist act. The dummy is perceived to talk through the synchronisation of the dummy's mouth movements with the ventriloquist's speech, resulting in 'visual capture' of perceived auditory spatial location. Source: http://tv.bt.com/tv/tv-news/british-ventriloquist-paul-zerdin-voices-his-joy-at-americas-got-talent-win-11364004727160. Accessed: 08/12/2016

dummy's mouth (see Figure 2.31). We also experience this illusion every day through our television sets, as we perceive the actor's speech to arise from their on-screen spatial locations rather than the (spatially disparate) speakers.

The ventriloquist illusion can be explained by one of the leading models of cue combination, the maximum likelihood estimation (MLE) model. This posits that visual capture arises via asymmetrical allocation of perceptual weight between the sensory modalities (Ernst and Banks 2002; Alais and Burr 2004). This asymmetry is driven by differential levels of cue reliability (see Figure 2.32). For example, visual spatial localisation thresholds are up to ten times lower than their auditory counterparts (Mills 1958; Westheimer and McKee 1977; Alais and Burr 2004). The far lower variance (or higher reliability) in the visual (relative to auditory) positional estimates lead to it being allocated much greater perceptual weight and thus its dominance of the integrated (multisensory) percept.

The concept of perceptual weighting gives the MLE model a greater degree of flexibility than earlier models of multisensory integration such as the "modality appropriateness" hypothesis, in which the modality with relatively superior reliability dominates the multisensory percept, with other sensory input exhibiting little or no influence (see Welch and Warren 1980). Allocating perceptual weighting based on reliability can explain why under conditions of visual uncertainty, it has been shown that visual spatial dominance is reversible: when presented with uncertain visual positional information (and thus noisy visual positional estimates) visual dominance declines relative to auditory or haptic signals (Ernst and Banks 2002; Alais and Burr 2004).

In the temporal domain the auditory system provides less variable estimates than its visual counterpart (Wearden et al. 1998; Grondin et al. 2001; Lapid et al. 2009), and has been shown to bias visual estimates in a variety of studies (Walker and Scott 1981; Donovan et al. 2004; Klink et al. 2011;

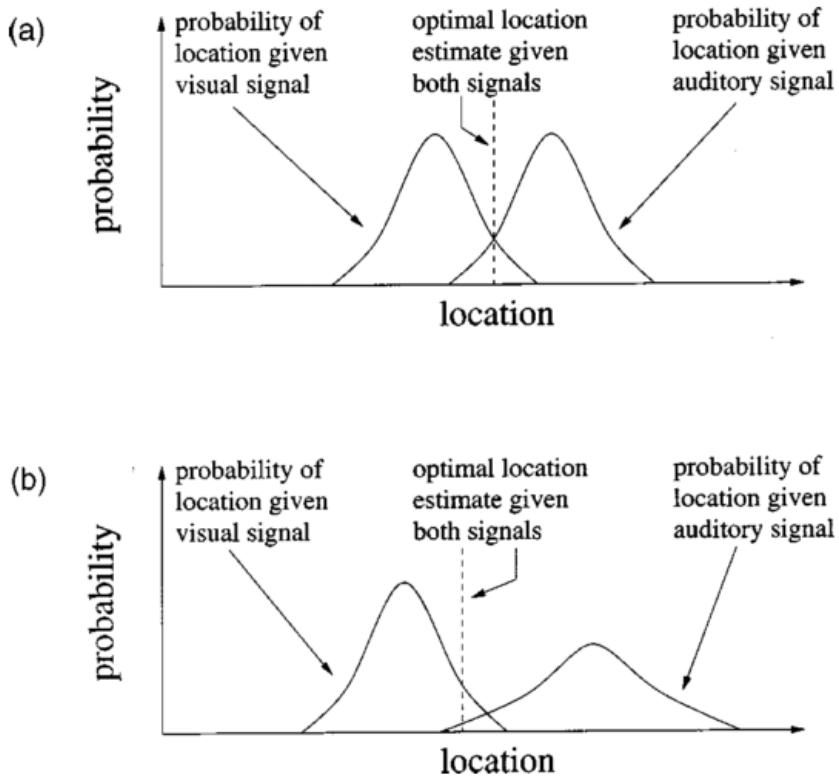

Figure 2.32: Schematic showing the principles of the maximum likelihood estimation (MLE) model of integration in determining the probable spatial location of an event. **a)** Both the auditory and visual signals are equally reliable, and thus the optimal perceived location (see dashed line - based on a weighted average) lies midway between the auditory and visual locations. **b)** The auditory signal is less spatially reliable than the visual signal (demonstrated by the relatively broader function), and therefore the weighted average is closer to the visual location estimate. Reprinted with permission from (Battaglia et al. 2003), and The Optical Society of America.

Romei et al. 2011; Sarmiento et al. 2012; Dolores de la Rosa and Bausenhart 2013). For example, the perceived onset time of a visual flash is significantly earlier when preceded by a sound compared to the condition when both stimuli occur simultaneously (Fendrich and Corballis 2001). The

perceived rate of visual flicker is drawn to concurrently presented auditory flutter (Gebhard and Mowbray 1959; Recanzone 2003), and - as per the "double flash illusion" - multiple auditory beeps presented with a single visual flash leads to the appearance of multiple (illusory) visual flashes (Shams et al. 2000). Frequently, reports of auditory temporal estimates 'attracting' visual temporal estimates are referred to as examples of "temporal ventriloquism".

A prediction of MLE models is that bimodal estimates will be more precise than either of the contributing unimodal estimates, such that the combined percept is said to be 'statistically optimal' (Ley et al. 2009). According to MLE, if two modalities are represented by P and Q, the variance (σ^2) of the bimodal estimate (PQ) can be calculated from the variance of each unimodal estimate using the following equation:

$$\sigma_{PQ}{}^2 = \frac{\sigma_P^2 \sigma_Q^2}{\sigma_{P+}^2 \sigma_Q^2}$$

Furthermore, the maximum reduction in the variance of the bimodal estimate (i.e. the greatest precision) will occur when the variances of both unimodal estimates are matched. Whilst the MLE model has been shown to be near-optimal in accounting for the variance of multisensory estimates in a variety of spatial tasks (Ernst and Banks 2002; Alais and Burr 2004), there is some debate as to how relevant it is for temporal integration.

In a study by Burr and colleagues (2009a), observers performed a temporal bisection task in which they had to decide whether the second (of three) transients was closer in time to the first or third transient. In unimodal

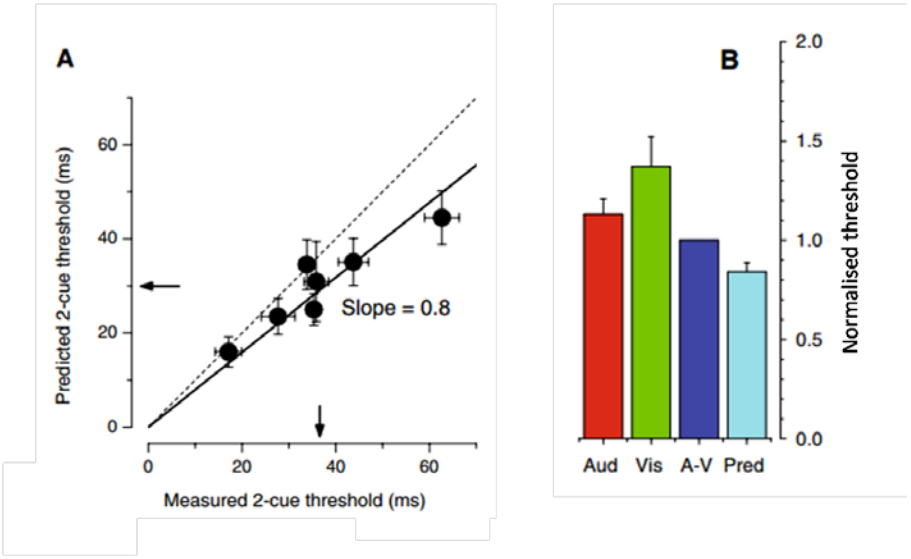

Figure 2.33: Data from an audio-visual bisection task by Burr and colleagues (Burr et al. 2009a). **A)** Predicted versus measured values of each observer's bimodal threshold (where predictions were calculated from the unimodal thresholds according to MLE). If the measured values exactly matched the mathematical predictions, the data points would fall on the dashed line (representing a line of equality). The solid black line shows the best fitting regression through each observer's bimodal threshold. The gradient of this line is less than one, signifying that whilst the bimodal thresholds were lower than the unimodal thresholds, they were higher than those predicted from the MLE model. Arrows close to each axis represent the group averages (where again the measured threshold is visibly higher than the predicted threshold), and error bars represent bootstrapped standard errors. **B)** Mean observer thresholds for each experimental condition, and predicted thresholds (light blue data) based on MLE calculations. Prior to averaging across observers, individual values were normalised by dividing each observer's thresholds by their bimodal (A-V) threshold. Error bars represent the SEM between observers. Figure reprinted with permission from Springer.

conditions, the transients were either auditory or visual, whilst in bimodal conditions the transients were audio-visual and the temporal order of the auditory and visual component was randomly varied between trials from +60ms (auditory leading) to -60ms (visual leading). This allowed the authors to examine the allocated weighting of audition and vision under conditions of temporal conflict (i.e. did the bisection point track the temporal position of the

auditory or visual stimulus, or a weighted average of the two). All conditions were mixed within a block of trials. MLE predictions for the bimodal thresholds were calculated from the unimodal thresholds (where the threshold refers to the amount that the second transient could be temporally adjusted back and forth before observers could reliably state that it was closer to the first (or third) transient).The authors reported that observer thresholds in the bimodal condition (Figure 2.33b – dark blue data) were only slightly lower than auditory-alone thresholds (Figure 2.33b – red data), and were higher than the predicted thresholds (Figure 2.33a and Figure 2.33b – light blue data). Additionally, more perceptual weight was assigned to the auditory stimulus in bimodal conditions than was predicted, suggesting that MLE models may not accurately account for temporal integration (Burr et al. 2009a).

It is argued that the MLE model violations may be specific to unfilled duration intervals, as these are suggested to have more ambiguous start and end points depending on whether (on a trial to trial basis) observers also include the markers delineating the empty interval in their estimate of duration (Hartcher-O'Brien et al. 2014). Hartcher-O'Brien and colleagues (2014) found that multisensory duration estimates were optimally integrated using filled durations in a discrimination task (i.e. the temporal estimates demonstrated the highest possible precision, and increasing noise in the auditory signal resulted in a greater reliance on visual information). However, a report that observers do not optimally integrate filled visual-tactile durations (i.e. bimodal precision remained higher than MLE model predictions), suggests that MLE

cannot explain the integration of filled durations across all modalities (Tomassini et al. 2011).

Additionally, a recent study of visual-tactile duration discrimination reported that bimodal task performance was superior to that of the best unimodal duration estimate, but worse than the statistically optimal weighted predictions derived from MLE (Ball et al. 2017). Instead it was determined that across the group of observers, the best fitting model to the bimodal data was one that relied on the weighted averaging of cues, but used tactile weights that were estimated empirically (derived from the difference in PSEs obtained under two types of duration discrepancy, i.e. $tactile_{long}visual_{short}$ and $tactile_{short}visual_{long}$) rather than weights that were optimally derived using MLE. Whilst this indicates that the integration of visual-tactile durations may not be optimal, any model of cue integration in which the bimodal estimate is improved through averaging requires that each contributing unimodal estimate has its own independent noise source. This suggests that independent duration estimates must be formed for each cue modality, which certainly has important implications for traditional internal clock models, which posit that all signals are encoded by a single, supramodal mechanism (Treisman 1963; Gibbon and Church 1984: see Section 2.2). Although independent noise sources could potentially arise within a single ‘clock’ framework, if duration estimates are transferred from the accumulator into modality-specific stores in memory (see Section 2.2, Figure 2.4), Ball et al. (2017) argue that their data may also be explained by modality-specific duration processing (e.g. Heron et al. 2012).

When duration discrepancies arise from perceptual rather than physical mismatches (in which temporal conflicts occur due to the ‘sound longer bias’ – see Section 2.4.3), cue integration may be resolved according to the modality appropriateness (auditory dominance) model (Welch and Warren 1980) rather than MLE. A study by Walker and Scott (1981) utilising perceptual duration discrepancies found that reproductions of a bimodal (audio-visual) stimulus were significantly longer than reproductions of the visual stimulus alone, and almost identical to audition alone, regardless of whether the stimuli were filled durations or empty gaps within a continuous presentation.

More recently Ortega et al. (2014) ramped the onsets and offsets of auditory stimuli (i.e. gradually increasing or decreasing the sound intensity) in order to equate unimodal auditory and visual thresholds (see Figure 2.34a), and then used these values to create audio-visual stimuli with very small duration discrepancies. The authors reported that even when auditory and visual thresholds were matched, the perceived duration of the bimodal (audio-visual) stimuli were indistinguishable from auditory stimuli alone during a temporal bisection task (i.e. “was the stimulus long or short?”), also favouring the auditory dominance hypothesis (Ortega et al. 2014: see Figure 2.34b). However, it should be noted that in both of these studies the same results could have been obtained if observers had simply responded to the auditory cue during bimodal stimulus presentation (i.e. ignoring the visual cue entirely), making their conclusions regarding auditory dominance models of cue integration less substantive.

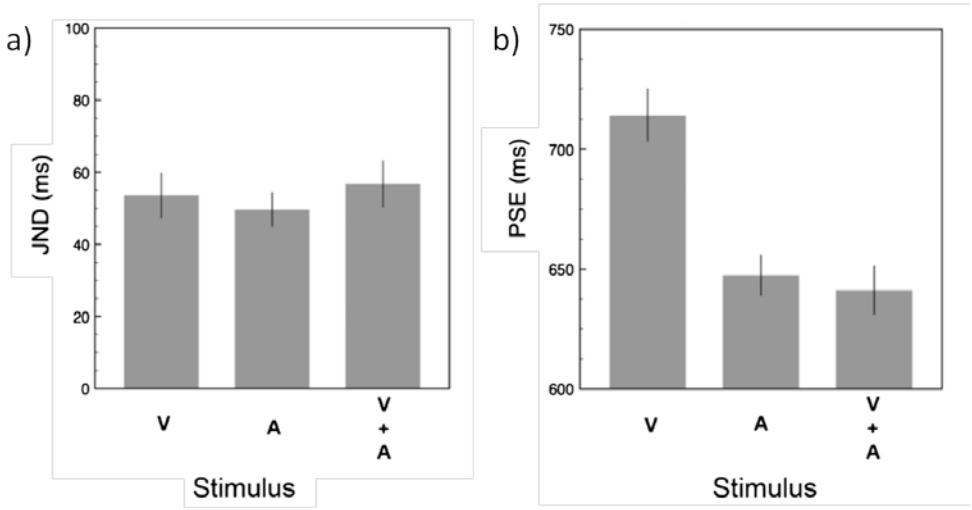

Figure 2.34: Data from Ortega et al.(2014), in which the reliability of the auditory stimuli was equated with that of visual stimuli by ramping the sound intensity of the auditory onsets/offsets. a) Observer thresholds (JND) for unimodal visual (V) and auditory (A) and bimodal audio-visual (AV) duration discrimination judgements. b) PSE data for the same three stimulus types, demonstrating that perceived duration of the audio-visual stimulus is similar to that of audition alone. Figure reprinted with permission from Springer.

A further challenge facing MLE models of integration is accounting for temporal cue discrepancy, as these models predict that both auditory and visual signals would continue to contribute to the perceived temporal estimate of a bimodal stimulus irrespective of cue discrepancy magnitude. It is clear that a putative multisensory integration mechanism should allow for a certain amount of discrepancy between signals, as despite being generated simultaneously, auditory and visual signals do not always arrive simultaneously at their respective receptor surfaces (e.g. light travels ≈ 1000000 times faster than sound). In terms of stimulus duration, this issue leads to misalignment in their physical onset times and is further complicated by audition's shorter neural transduction and transmission latencies. In addition, stochastic variations in internal neural noise levels mean that the

neural correlates of physically matched auditory and visual stimulus durations are likely to be discrepant. However, if two signals are highly discrepant they are less likely to have arisen from a single event, and thus in these circumstances cue segregation would be favourable in order to avoid potentially hazardous integration of unrelated temporal cues.

Most multisensory investigations into temporal discrepancy centre on the amount of audio-visual asynchrony that is tolerated before asynchrony is perceived (e.g. Stone et al. 2001; and for a review see Vroomen and Keetels 2010). It is suggested that when the amount of discrepancy between two signals exceeds a certain threshold, the signals will fall outside of the "temporal window of integration" (TWI), and be perceptually segregated. As a result, these reports are primarily concerned with measuring the temporal resolution of a putative crossmodal 'asynchrony detection/discrimination mechanism', rather than the tolerance of a 'crossmodal duration integration mechanism' to increasing temporal discrepancy.

Studies that have examined both perceptual bias *and* temporal discrepancy have done so using different strategies. For example, Morein-Zamir and colleagues (2003) designed a visual temporal order judgment (TOJ) and presented irrelevant, brief auditory stimuli at varying intervals before and after each visual stimulus (ranging from 0ms – 225ms) (see Figure 2.35). Observers were instructed to determine "which visual stimulus was presented first?" whilst ignoring the auditory stimuli. Performance on the TOJ task significantly improved for audio-visual intervals ranging 75 - 225ms compared to baseline (0ms). This suggests not only that the auditory stimuli attracted the visual stimuli towards them in time (a type of "temporal

ventriloquism"), but this interaction occurred over a range of audio-visual temporal discrepancies extending up to 225ms.

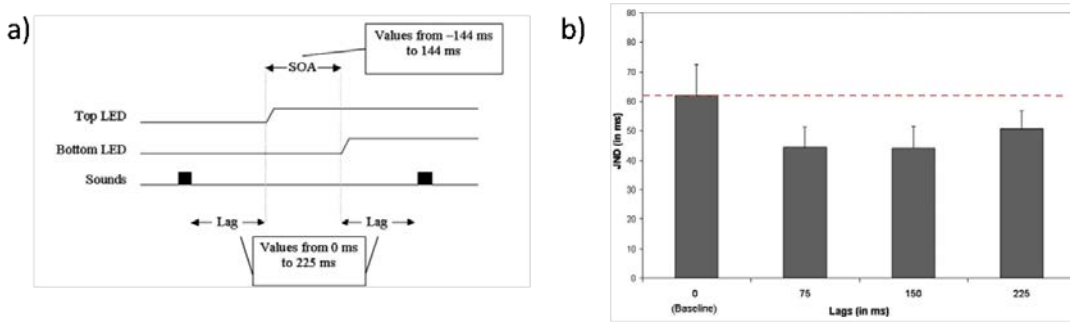

Figure 2.35: a) Schematic showing the experimental paradigm used by Morein-Zamir and colleagues (2003). Observers performed a temporal order judgment (TOJ) between two visual LEDs whose stimulus onset asynchrony (SOA) varied from +/- 144ms (positive SOAs referred to trials where the top LED came on first). At varying intervals or "lags" (0 – 225ms), a brief auditory stimulus was presented before the onset of the first LED and after the offset of the second LED. Observers decided "which visual LED was presented first?" **b)** Observer thresholds (JND) for the visual TOJ task as a function of auditory stimulus lag. Relative to baseline (0ms), all lag intervals resulted in significantly lower JNDs. Figure reprinted from (Morein-Zamir et al. 2003) with permission from Elsevier.

Another approach to mapping cross-modal bias across a range of temporal discrepancies was employed by Roach and colleagues (2006), who designed a rate discrimination task in which observers were asked to ignore the auditory component of a bimodal stimulus (whose rate varied between trials) and make purely visual judgements. Their results were found to be inconsistent with MLE predictions, as auditory bias was demonstrated across a finite range of rate discrepancies (Roach et al. 2006: see Figure 2.36). This supports the idea that multisensory cues in the temporal domain are not integrated based on their reliability alone: the degree of discrepancy between the signals must also be involved. The authors instead proposed an alternative Bayesian model of integration that considered relative reliability

and prior knowledge of the correspondence between auditory and visual signals. Predictions from this model were reported to provide a better fit to the experimental data (see Figure 2.36).

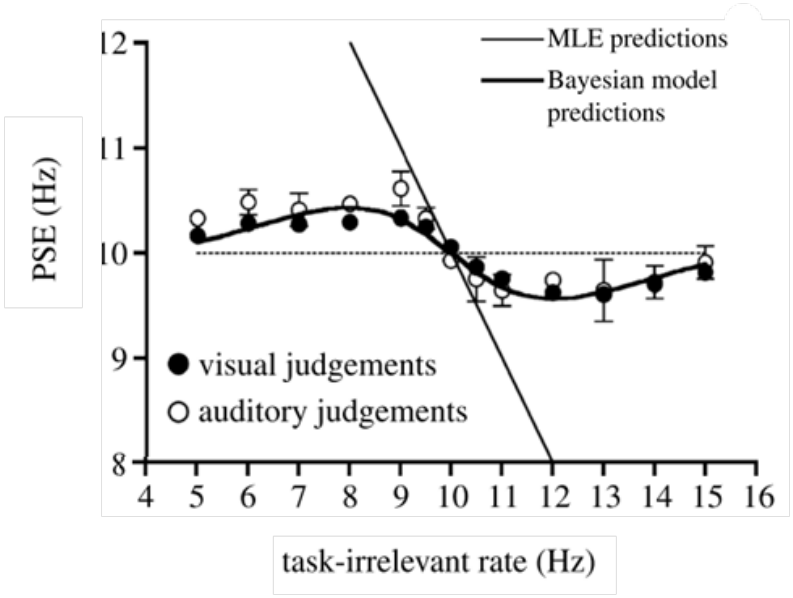

Figure 2.36: Mean data from a study by Roach and colleagues (2006) showing perceived auditory (white circles) and visual (black circles) temporal rate (indicated by the PSE) in the presence of a distracting 'irrelevant' stimulus of variable rate in the opposite modality. Models based on MLE (in which discrepancy should be irrelevant) and Bayesian predictions are shown as the thin and thick black lines respectively. The data is best described by the Bayesian model which takes into account prior observer knowledge of auditory and visual rate signals, thus predicting the observed cue segregation with increasing temporal discrepancy. Figure reprinted with permission from the Royal Society.

More recently, Klink et al. (2011) also used 'irrelevant' auditory stimuli to investigate the range of audio-visual asynchronies over which overlapping, filled durations may interact. Observers performed duration discrimination judgments ("which was longer?") between two 500ms visual stimuli, one of which was always paired (via the temporal midpoint) with an auditory 'distracter' of varying duration (150 – 850ms). They were instructed to ignore the auditory stimulus and make purely visual judgements. The proportion of

'visual+audio perceived as longer' responses was plotted as a function of audio-visual asynchrony (see Figure 2.37). Compared to baseline (no audio-visual asynchrony), longer distracters caused an increase in the number of 'long' responses and shorter distracters caused a decrease in 'long' responses, which the authors interpreted as evidence of audio-visual interaction across the full range of asynchronies. However, it should be noted that a similar pattern would be expected if the observers had switched to

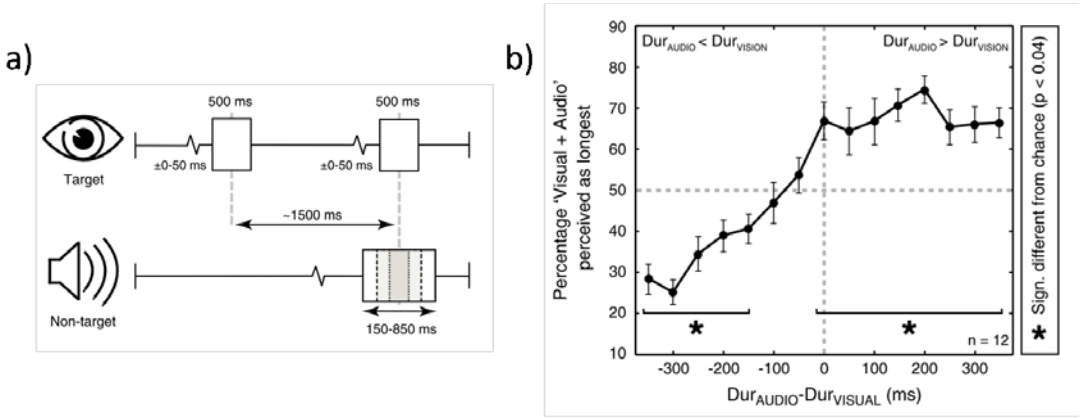

Figure 2.37: a) Schematic showing the experimental paradigm of Klink and colleagues (2011). Observers made duration discrimination judgements between two 500ms visual stimuli ('targets'), whose interstimulus interval (ISI) was 1500ms +/- a 50ms jitter. One of the visual stimuli was paired with an irrelevant auditory stimulus (the non-target), which varied in duration between 150 – 850ms. This is presented with the second visual stimulus in the above schematic, but in reality was randomly presented with either visual target. **b)** The proportion of "visual+audio" 'long' responses plotted as a function of audio-visual asynchrony. Negative values denote conditions where the auditory stimulus was shorter than the visual stimulus. Stars denote conditions which were significantly different from chance performance (50% longer responses). Figure reprinted with permission from Springer.

making audio-visual judgments (i.e. comparing the auditory distracter with the visual duration); particularly as both visual stimuli were physically equal in duration. Unfortunately as the authors employed a performance measure which cannot distinguish between this response strategy and true auditory

perceptual bias of visual duration, it is impossible to rule out this possibility. Thus, the range of temporal discrepancies over which auditory stimuli may bias perceived visual duration continues to remain unclear.

In summary, the brain must be flexible enough to accommodate small amounts of physical/neural signal discrepancy which are likely to arise despite the two signals being co-generated by a common external source. At the same time, it must avoid potentially hazardous misperceptions that could occur from integrating two signals that do not belong together. Thus, the decision to integrate is likely to be determined by the spatial, temporal and semantic congruence of the afferent sensory information, and the combined perceptual estimate governed (at least in part) by relative cue reliability. The exact nature of the relationship between these factors is yet to be determined.

2.4.8 Summary

It is likely that on some level, duration encoding is sensory specific. These unimodal mechanisms should have some capacity to encode multiple, overlapping durations (perhaps arising at different spatial locations), although this may come at a cost (i.e. increased variance of estimates). Each mechanism may also be affected by recent stimulus history, such that repeated exposure to a fixed duration (or visual flicker) results in a distortion of perceived duration. Yet it is also probable that signals arising from different modalities converge (perhaps at a more central location) allowing for crossmodal comparison (and perhaps integration – see Section 2.4.7).

Additionally, the documented interactions between duration and higher-level processes such as attention and non-temporal magnitude suggest that neural regions such as the frontal lobe are likely to exert some influence over duration processing.

2.5 The neurophysiological basis of temporal processing

The search for the neural correlates of duration processing has taken in a raft of diverse neural structures via a combination of techniques including electrophysiology, neuroimaging, pharmacological and lesion studies.

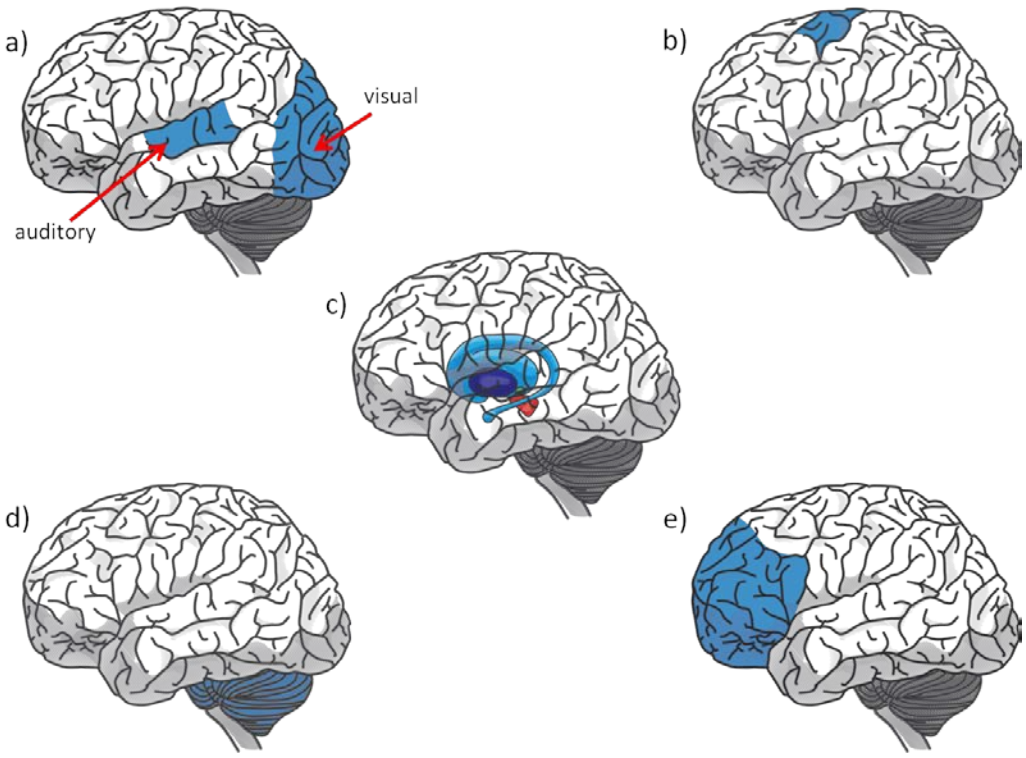

Figure 2.38: Schematic showing the approximate anatomical locations of some of the neural areas that have been implicated in duration processing, including **a)** the visual and auditory cortices, **b)** the supplementary motor area, **c)** the basal ganglia, **d)** the cerebellum and **e)** the prefrontal cortex

Despite this, no single structure has emerged that could be said to have temporal processing as its primary function. This has led sections of the neurophysiology literature to challenge the assertions of classic 'internal clock' models (see Section 2.2) that rely on modular components such as a pacemaker or accumulator.

Instead, neurophysiological approaches have implicated that numerous brain areas are involved in our temporal estimates, leading to suggestions that timing could be achieved through interactions between spatially distributed networks of neural activity (Mauk and Buonomano 2004; Buhusi and Meck 2005; Coull and Nobre 2008; Wiener et al. 2010; Coull et al. 2011; Wiener et al. 2011). The location of these interacting areas could vary with the nature of the timing signal (e.g., motor vs. sensory), the nature of the temporal metric (e.g., duration vs. temporal rate), the sensory modality of the temporal signals (i.e. auditory or visual) or the range of durations involved (i.e. sub-second or supra-second) (Ivry and Hazeltine 1995; Lewis and Miall 2003b; Jantzen and Kelso 2005; Bueti et al. 2008; Wiener et al. 2010; Teki et al. 2011).

The following sections will examine evidence for the neural correlates of duration processing utilising a variety of techniques, and primarily focusing on durations in the sub-second range.

2.5.1 Electrophysiological studies

Electrophysiological techniques record the electrical properties of individual neurons *in vivo* (within a living animal). Neurons are not electrically neutral,

but exhibit small, measureable voltages which change depending on the state of the cell (Levitan and Kaczmarek 2002). When a neuron fires in response to a stimulus (referred to as an action potential or “spike”), there is a measurable jump in voltage compared to the baseline ‘resting potential’ (see Figure 2.39). These changes in electrical activity can be measured using microelectrodes inserted into the neuron.

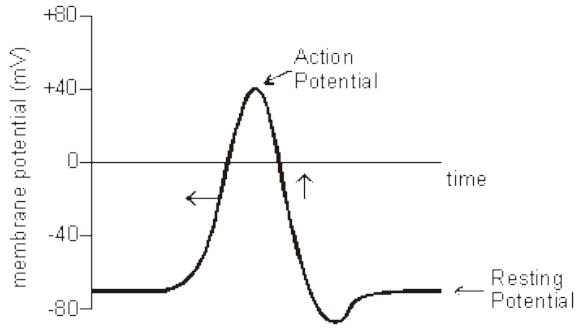

Figure 2.39: Schematic of an action potential showing the change in voltage over time. The resting potential of the neuron is around -70mV, and during an action potential this jumps to approximately +40mV. Source: http://www.mrothery.co.uk/module4/webnotes/Mod4Notes.htm. Accessed: 24/09/2014

Neurons that vary their rate of firing in response to the duration of a stimulus have been evidenced in a variety of species, and may be described as ‘tuned’ for duration. Neurons that demonstrate duration tuning can be categorised into three classes of electrophysiological response. Neurons that respond preferentially to a limited range of durations centred on a ‘best duration’ (BD) are referred to as ‘band-pass’ neurons (Figure 2.40b). Short-pass neurons then respond to durations shorter than a given duration, but their firing rate rapidly declines in response to durations longer than this value (Figure 2.40a), and long-pass neurons respond only when a duration

exceeds a minimum threshold value (Figure 2.40c). Importantly, increasing non-temporal signal strength does not cause these neurons to respond at increasingly shorter durations (Faure et al. 2003; Perez-Gonzalez et al. 2006). For example, the band-pass neuron centred on a duration of 5ms in Figure 2.40b shows the same tuning preference across a variety of auditory

Figure cannot be displayed due to copyright law. To view the original figure please see Faure et al. (2003), Figure 4 (p3056).

Figure 2.40: Data from Faure et al. (2003) showing three different types of duration tuning in the inferior colliculus of a big brown bat. 'Threshold' refers to the minimum sound level (decibels) at which the neuron will respond, and the different symbols denote increases in sound level above this threshold. **a)** A short-pass neuron that responds to durations up to 5ms, but whose response rapidly declines for durations longer than this value. **b)** A band-pass neuron, which is tuned around 5ms (its best duration or "BD"). **c)** A long-pass neuron, which responds only to durations beyond 5ms.

intensities (Faure et al. 2003). Consequently, the invariable tuning profiles of these neurons would allow them to distinguish between weaker signals presented for long durations and stronger signals presented for short durations.

Neurons tuned to duration have been found in a range of vertebrate species. The locations of these neurons include auditory midbrain (Casseday et al. 1994; Chen 1998; Brand et al. 2000; Ma and Suga 2001; Perez-Gonzalez et al. 2006; Leary et al. 2008), auditory cortex (Galazyuk and Feng 1997; He et al. 1997), visual cortex (Duysens et al. 1996) and premotor cortex (Merchant

et al. 2013b). Duration tuning of this type offers potential neurological support for the notion of channel-based duration processing: where banks of bandwidth-tuned neurons - each centred around progressively different

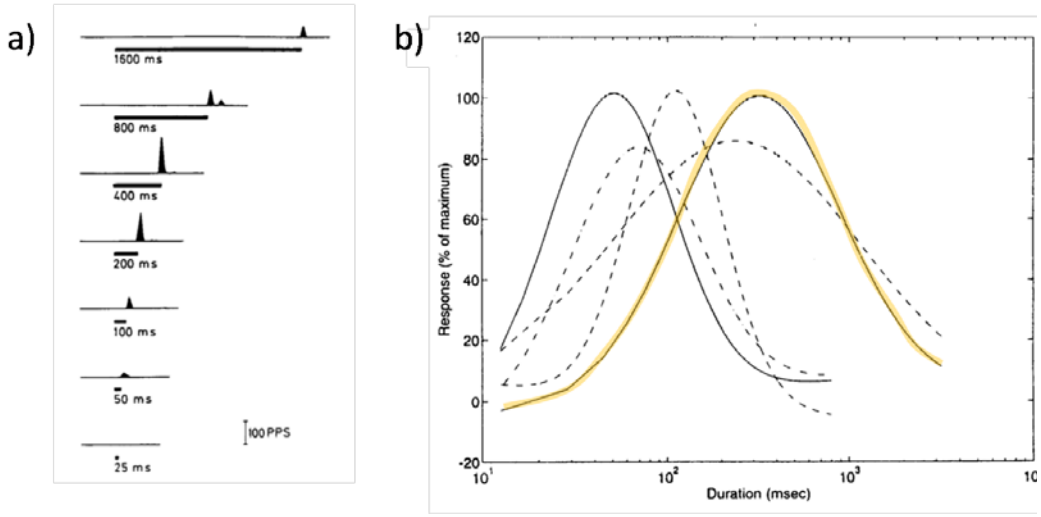

Figure 2.41: a) Off-response of a single neuron from cat visual cortex in response to a variety of stimulus durations. Bold black horizontal lines represent the presentation of a stimulus, above which the response of the neuron is shown in spikes per second (PPS). Not only does the neuron respond at the offset of stimulus duration, but it also shows an increase in firing rate (greater amplitude) with increasing duration, peaking at 400ms (its preferred duration). Beyond this duration, firing rate declines, indicating that this neuron is band-pass tuned. **b)** The tuning function of the neuron from **(a)** is shown highlighted in orange, alongside four additional band-pass neurons from cat visual cortex (area 17 = dashed lines and area 18 = filled lines) which respond to duration offset. Each neuron has a slightly different 'best duration', at which they exhibit peak firing rate. Figure reprinted from (Duysens et al. 1996) with permission from Elsevier.

durations to their neighbours - facilitate the processing of duration within overlapping perceptual channels (see Section 2.3.3). Given that these tuned neurons tend to be evidenced at relatively early stages of processing within the sensory cortices, they also provide a potential explanation for why some perceptual phenomena show selectivity for sensory modality.

However, channel-based processing is challenged by the fact that time is continuous, and therefore neighbouring ‘channels’ with increasing preferred durations would become activated successively in a ‘domino’ sequence as physical event duration progressed. One possible solution to this problem would be if the neurons became activated by the stimulus offset rather than the onset. Precisely this pattern of activity has been recorded from cat visual cortex where some neurons respond maximally when stimulus offset occurs at their preferred duration (Duysens et al. 1996: see Figure 2.41). Single cell recordings taken from a sample of 174 neurons in areas 17 and 18 of the cat (corresponding to striate and extrastriate cortex), showed that approximately one third of these neurons responded to the offset of a range of stimulus durations (from 10 – 3200ms).

Neural off-responses have also been demonstrated within the auditory system of the cat (He et al. 1997), big brown bat *Eptesicus fuscus* (Faure et al. 2003) and rat (Perez-Gonzalez et al. 2006). In one report, extracellular recordings in the bat inferior colliculus identified 73 duration-tuned neurons that responded to the varying duration of an auditory tone. All short-pass and band-pass neurons in the sample (89% of all duration tuned neurons tested) were classed as “offset responders” (Faure et al. 2003). As tone duration increased, this type of response was characterised by increasing latency between the start of the trial and the neuron’s first ‘spike’ (increase) of activity, suggesting that the neurons were able to track the stimulus offset (see Figure 2.42).

Figure cannot be displayed due to copyright law. To view the original figure please see Faure et al. (2003), Figure 6A (p3058).

Figure 2.42: Firing activity of a single band-pass neuron in response to the offset of an auditory tone of varying duration (Faure et al. 2003). Dots represent 'spikes' of activity recorded from the neuron, and are plotted as a function of time from stimulus onset (0ms). The rastergram shows that for durations ≥4ms (the preferred duration) the latency of spike activity (from 0ms) increases linearly with stimulus duration, as the neuron tracks the stimulus offset. The vertical spread of the dots demonstrates the tuning bandwidth of the neuron, which extends to around 12ms (beyond which the cell is almost entirely unresponsive).

Another well documented type of neural activity that can occur during timing tasks (often with a motor component), is known as 'ramping activity'. This refers to gradual (often linear) increases or decreases in firing rate that occur in conjunction with the evolving likelihood of a temporally relevant event. Animals performing tasks in which predictions must be formed about the required delivery time of a motor response (resulting in a reward), have shown that neural activity climbs (or falls) slowly during the interval before stimulus delivery. This activity then reaches a peak (or trough) which is almost coincident with an anticipated motor response. Ramping activity has also been proposed as a potential mechanism for encoding duration (Durstewitz 2003; Reutimann et al. 2004).

Shuler and Bear (2006) reported ramping activity during an associative learning task, which required animals to make predictions about the anticipated delivery time of a stimulus (see Section 2.2). In the study, single cell recordings were made in the primary visual cortex of rats trained to

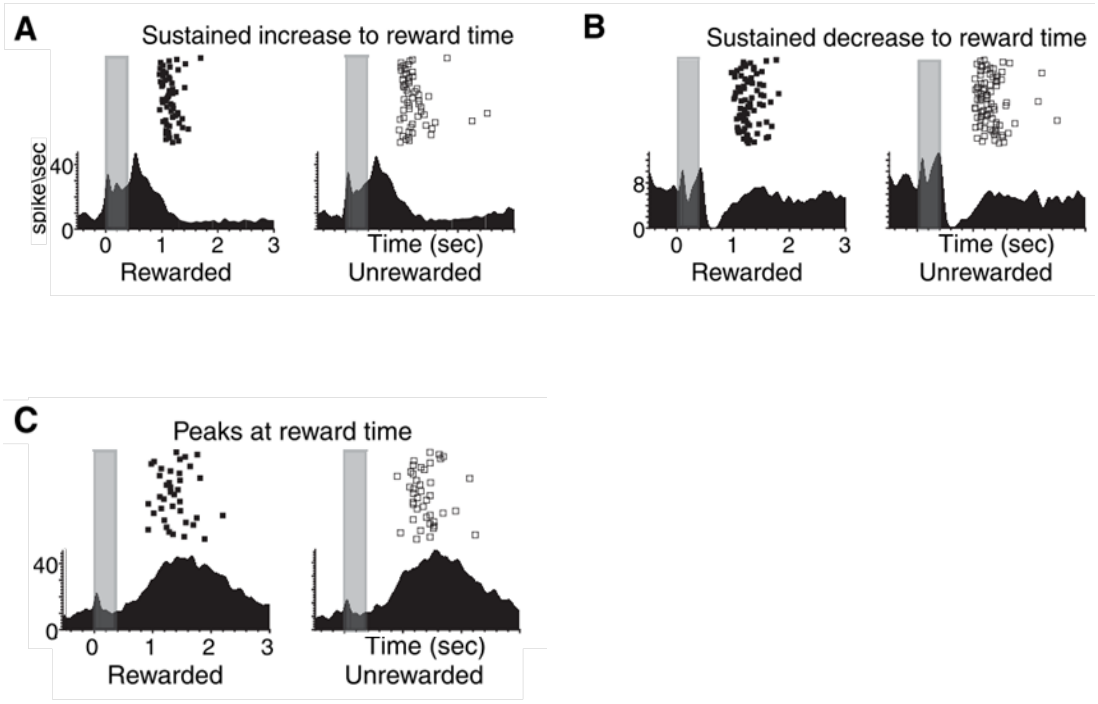

Figure 2.43: Three different types of reward timing in rat visual cortex (Shuler and Bear, 2006). Plots show the firing activity of 3 individual neurons in response to expected time of reward following a period of training (where reinforced durations varied between each rat). For each type of reward timing two plots are shown, with 'rewarded' trials shown on the left (where the reward was delivered) and 'unrewarded' trials shown on the right (where the reward was withheld). Filled black squares indicate the time of reward delivery, and empty squares indicate the time that reward would have been delivered on unrewarded trials. The grey rectangles indicate the presentation of the visual cue. **A)** Example of a sustained increase in firing up to the expected time of reward. **B)** Example of a sustained decrease in firing up to the expected time of reward. **C)** Example of a peak in firing rate that coincides with the expected time of reward. In all three plots the same response profile is seen in both rewarded and unrewarded trials, indicating that this activity is not dependent on reward delivery itself. Figure reprinted with permission from AAAS.

associate a 400ms visual cue with a water reward, following a fixed temporal interval. The temporal interval was defined by a predetermined number of licks on the water dispenser, and varied between rats (roughly equating to either 1 or 1.6 seconds). On half of the trials, no reward was given, so that neural responses relating to reward delivery itself could be distinguished from those relating to temporal expectancy of the reward. 130 neurons were found to show ‘reward timing’, of which three types of neural activity were evidenced: 50% of neurons showing a sustained increase in firing until expected time of reward, 22% showing a sustained decrease in firing and 28% showing a peak in firing rate that coincided with the expected time of reward (see Figure 2.43). This evidence suggests that neural activity early in the visual processing hierarchy may (collectively) provide a population estimate of temporal expectancy that can be used to guide behaviour. Importantly, it also indicates that neurons in the primary visual cortex may have a role beyond that of simple feature detection.

In addition to the primary visual cortex, ramping activity has also been reported in prefrontal cortex (Kalenscher et al. 2006; Oshio et al. 2008; Jin et al. 2009), striatum (Jin et al. 2009), motor and premotor cortex (Renoult et al. 2006; Lebedev et al. 2008; Mita et al. 2009; Merchant et al. 2011), lateral intraparietal area (Leon and Shadlen 2003; Janssen and Shadlen 2005) and thalamus (Tanaka 2007) across a variety of timing tasks (for a review see Wittmann 2013). The ramping activity can also take various forms. In a rhythmic synchronisation-continuation (SC) task, Merchant et al. (2011) trained monkeys to tap a button along to a beat (defined by brief auditory tones) with a fixed interstimulus interval (synchronisation phase), and then

continue to tap at the same rate without the aid of the sound (continuation phase). This ability required that the animal kept track of the elapsed time from the previous tap as well as the time remaining to the next tap, so as to maintain the rhythm. Electrophysiological recordings were then made from 1083 neurons in supplementary motor area (SMA) and pre-SMA during a series of SC tasks - each with a different inter-tap interval (ranging from 450 - 1000ms). From these recordings it was determined that a subset of neurons varied their activity as a function of inter-tap duration during either the synchronisation or continuation phase. These neurons were divided into five groups, depending on the different type of ramping activity identified. In one group (termed “relative timing cells”), individual neurons showed increases in firing rate leading up to the button press. Increases in inter-tap interval were associated with earlier onsets and more gradual increases in ramping activity (see Figure 2.44a), which had the effect of causing convergence around a common firing rate, irrespective of inter-tap interval. This rate value could therefore be utilised as a threshold value for initiating a motor action (button press, in this case) after a specific interval had relapsed since execution of the previous motor action. In another group (termed ‘time accumulator cells’), neurons showed linear increases in activity *following* the button press, of which the magnitude (the peak activity) increased with increasing inter-tap interval (see Figure 2.44b). It is suggested that this activity could signal the passage of time following the last button press, where the magnitude of firing activity directly correlates with elapsed duration.

In total, Merchant et al. reported three groups of neurons that showed ramping activity leading up to the button press, and two groups following the

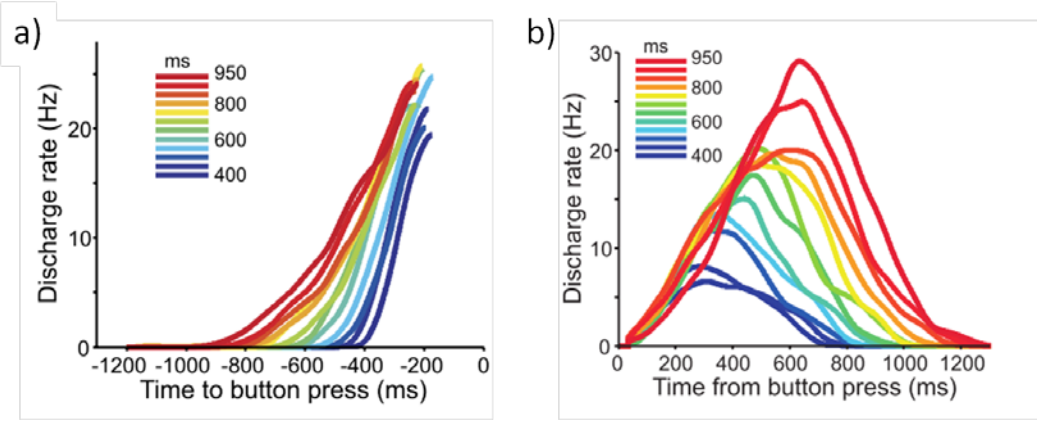

Figure 2.44: Neural population response functions showing two types of ramping activity found in primate supplementary motor area during a rhythmic timing task (Merchant et al. 2011). Population activity is shown as a function of time, either leading up to a button press (a) or following a button press (b).Different colours represent the responses of the population to different inter-tap intervals on a synchronization-continuation task (see main text for details). **a)** Predicted responses of a population of relative timing cells, in which neural activity gradually increases leading up to the button press. Longer inter-tap intervals result in earlier onsets of activity, and flatter response gradients. **b)** Predicted responses for a population of accumulator cells, in which peak activity has an increasingly higher magnitude with increasing inter-tap interval. Figure reprinted with permission from the author.

button press. Collectively, these different types of neural activity could form a temporal code in response to the rhythmic SC task, allowing the primate to quantify the amount of time that had elapsed from the last button press and predict when the next button press should commence. However, given that the ramping activity of individual neurons was very noisy, it is unlikely that individual neurons alone would provide an accurate temporal measure. Merchant et al. instead suggested that a downstream read-out mechanism might sum together the activity of multiple neurons (showing the same type of ramping activity) over time, to determine a ‘population response’.

Population response functions for the aforementioned ‘relative timing cells’ and ‘time accumulator cells’ are shown in Figure 2.44, for a variety of possible inter-tap intervals.

However, it should be noted that the neural activity described by Merchant et al. may be specific to the rhythmic SC task utilised in their study, in which primates had to learn and respond to the temporal frequency of auditory tones. Similarly, most associative learning tasks involve repetitive, cyclical behaviours or sequences. How this type of ramping activity might play a role in non-rhythmic temporal tasks (i.e. duration discrimination) is unclear. Furthermore, the climbing or ‘ramping’ activity does not necessarily indicate that duration is being encoded by neurons in these brain regions, it may simply reflect preparatory activity in readiness for a motor response or a decision making process. It is possible that temporal information is encoded elsewhere in the brain, and passed to these areas in anticipation of a temporally relevant event.

In summary, electrophysiological evidence has shown that neurons are capable of adjusting their rate of firing in a manner that is time-specific and thus provides a potential metric upon which we might base our temporal judgments. The apparent importance of the auditory and visual cortices in temporal processing appears to contradict models of timing that rely on a central, amodal ‘internal clock’ (see Section 2.2) and thus provides support for localised timing operating in parallel at multiple stages of the sensory cortical pathways. This distributed approach may help to explain why some perceptual distortions of duration do not transfer between modalities (Walker et al. 1981; Becker and Rasmussen 2007; Heron et al. 2012) and why a

recent cue combination study found duration integration patterns consistent with visual and tactile unisensory duration estimates forming separate, independent noise sources (Ball et al. 2017).

2.5.2 Neuroimaging

Functional magnetic resonance imaging (fMRI) is a neuroimaging technique that measures changes in blood flow in order to detect neurological activity. Depending on the oxygenation of the blood, different magnetic properties can be detected by the fMRI scanner. Increases in blood oxygenation are associated with increases in neurological activity, and therefore measuring the blood-oxygenation-level-dependent (BOLD) signal allows scientists to correlate specific tasks/functions with specific regions of the brain. Positron emission tomography (PET) uses a different technique to image neural activity. An analogue of glucose containing positron-emitting radio-nucleotides (a radiotracer) is ingested, and the radiotracer accumulates in areas of increased neural activity. The PET scanner then images this activity by detecting Gamma rays that are emitted by the radiotracer.

fMRI and PET studies have implicated numerous neural regions in sub-second duration processing (for reviews see Lewis and Miall 2003b; Wiener et al. 2010). However, consistent reports of activation within a subset of regions including the supplementary motor area (SMA) (Schubotz et al. 2000; Macar et al. 2002; Ferrandez et al. 2003; Lewis and Miall 2003a; Pouthas et al. 2005; Tregellas et al. 2006; Coull et al. 2008; Morillon et al. 2009; Shih et al. 2009; Wiener et al. 2010), the basal ganglia (BG) (Schubotz

et al. 2000; Ferrandez et al. 2003; Nenadic et al. 2003; Harrington et al. 2004; Jahanshahi et al. 2006; Bueti et al. 2008; Coull et al. 2008; Shih et al. 2009; Harrington et al. 2010) and prefrontal cortex (PFC) (Maquet et al. 1996; Ferrandez et al. 2003; Lewis and Miall 2003a; Nenadic et al. 2003; Pouthas et al. 2005; Tregellas et al. 2006; Coull et al. 2008; Morillon et al. 2009; Shih et al. 2009) suggest that these areas in particular may have a key role.

In a study by Coull et al. (2008) activation was reported in multiple brain regions during a duration discrimination task, in which fourteen healthy observers had to decide whether a visual 'probe' duration was "shorter, longer or the same?" as a visual 'sample' duration. Randomly interleaved

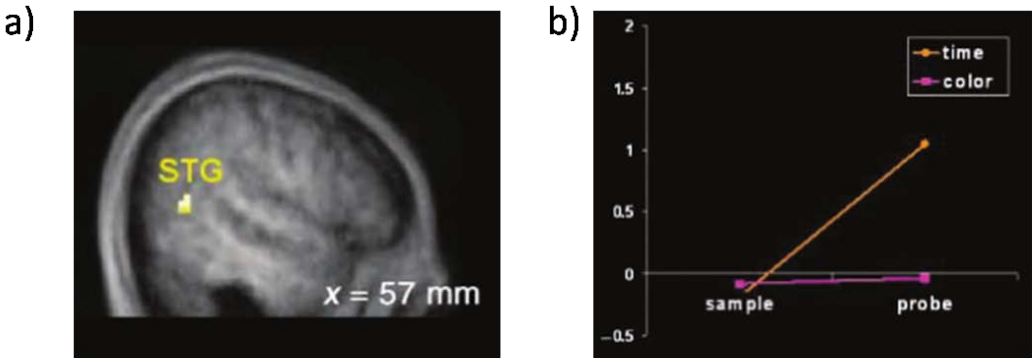

Figure 2.45: Data from a study by Coull et al. (2008) in which observers completed a duration discrimination task between a visual 'sample' and a visual 'probe' duration. In addition, they performed a control task in which they had to decide whether the probe was 'bluer' or 'redder' than the sample stimulus. **a)** Significant activation (depicted by yellow pixels) was found in the superior temporal gyrus (STG) during the duration task (but not during the control task). The value 'x' refers to a spatial co-ordinate used to define the anatomical location of the scan slice. **b)** Illustrating the mean level of activity in the STG during presentation of the sample and probe stimuli, for both the time (orange data) and colour (pink data) tasks. Figure reprinted with permission from MIT Press Journals.

within ‘duration trials’ were ‘colour trials’ which acted as a control task, in which observers had to decide whether the probe stimulus was “bluer, redder or the same?” as the sample stimulus. To make the cognitive demands of the colour task as similar as possible to the duration task, the hue of the visual stimulus varied throughout its presentation, requiring that the observer attend the stimulus to determine its ‘average’ colour. All visual stimuli had durations of either 540ms, 1080ms or 1620ms, counterbalanced between trials. The results demonstrated significantly greater activation of the SMA, bilateral dorsolateral prefrontal cortex (DLPFC) and right superior temporal gyrus (STG) during the duration task compared to the colour task. Additionally, the left putamen (part of the basal ganglia) and right STG (see Figure 2.45) were selectively activated during the duration task but only during the presentation of the sample stimulus *or* probe stimulus respectively. This would suggest that the putamen may be tasked with encoding and/or storing duration estimates, whereas the STG (an area that includes the primary auditory cortex) may be involved in retrieving and/or comparing duration estimates (see Figure 2.45).

Additional support for the role of the putamen and SMA in temporal processing comes from an earlier study by Ferrandez and colleagues (2003), who had used a similar duration discrimination paradigm and also found significant activation in these neural areas. However, activity was also identified in four additional regions: ventrolateral prefrontal cortex, the left premotor cortex, inferior parietal cortex and temporal cortex, suggesting the involvement of a wider network of areas in temporal processing.

The effect of task difficulty on a putative timing network has also been reported for durations in the sub-second range (Tregellas et al. 2006). Observers made duration discrimination judgements ("was the second tone longer or shorter than the first tone?") between a 200ms auditory standard and a variable auditory comparison duration. In the 'easy' condition, the comparison durations ranged from 70ms – 330ms and in the 'difficult' condition they ranged from 160 – 240ms (a smaller step size). Judgements were made using a key press. Control conditions included the presentation of two successive 200ms tones, and observers were merely asked to respond at random. Behavioural data showed that observer performance was significantly higher in the 'easy' task compared to the 'difficult' task. During the 'easy' task, activation was observed in the cerebellum (a motor area located in the hindbrain (Brooks and Thach 1981; Hore et al. 1996; Ivry et al. 2002)), medial occipital cortex and bilateral STG, yet in the 'difficult' task additional activation was noted in the SMA, insula/opercular cortex, premotor cortex, DLPFC, basal ganglia and thalamus (see Figure 2.46). As the cerebellum was active irrespective of task difficulty, it is argued that this area plays an important role in perceptual, sub-second timing, a finding that is corroborated by other neuroimaging studies (Lewis and Miall 2003a; Bueti et al. 2008). The finding that SMA was only significantly activated in the 'difficult' task appears to contradict reports that this region is consistently engaged across all timing tasks (Wiener et al. 2010). However when less conservative analyses were applied to the data from the 'easy' task, the same neural areas showed activation as with the 'difficult' task (Tregellas et al. 2006). This highlights a potential difficulty in comparing data across fMRI

studies that do not use equivalent timing tasks. It could also explain why such a wide range of neural regions have been evidenced in event timing, if

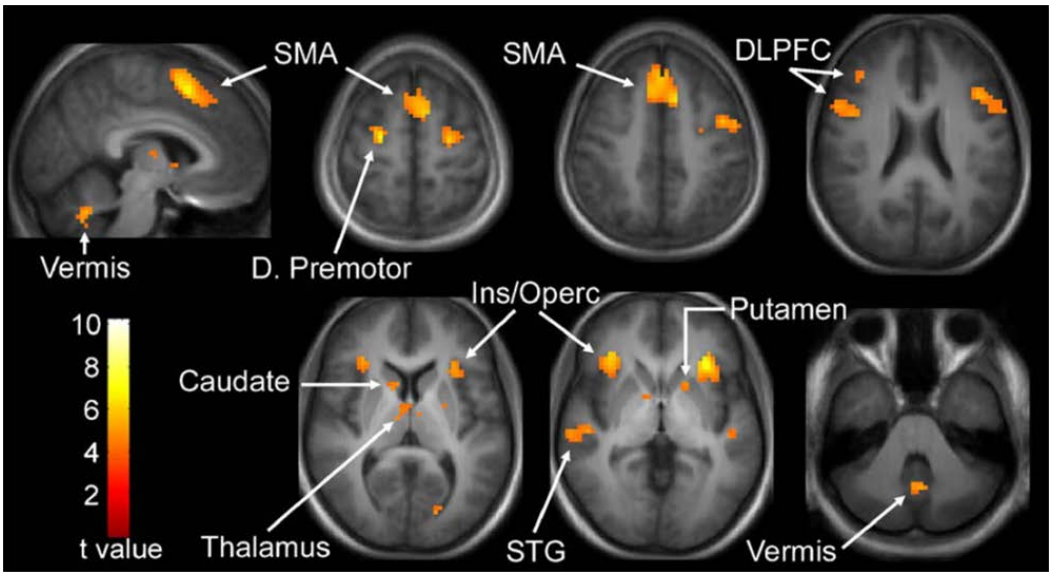

Figure 2.46: fMRI Data from the study by Tregellas and colleagues, showing areas of neural activity during the 'difficult' duration discrimination task. The 't value' reflects neural activity that is significantly greater than baseline as determined by a series of statistical t-tests – with whiter colours indicating greater significance. Figure reprinted with permission from Elsevier.

studies have utilised different timing tasks, control tasks, duration ranges, and/or statistical analyses.

2.5.3 Lesion studies

It is argued that if a neural region is involved in duration processing, damage to this area should result in distortions of time perception and/or reduced performance on timing tasks (for reviews see Coull et al. 2011; Allman and Meck 2012). Lesion studies can be broadly classified into those involving

temporary lesions (e.g. using transcranial magnetic stimulation) and those involving physical lesions arising through neurological damage or disease (e.g. clinical populations).

2.5.3.1 Transcranial Magnetic Stimulation

Transcranial magnetic stimulation (TMS) involves applying a localised magnetic pulse to temporarily disrupt neural activity by affecting the electrical signal. It therefore creates a 'temporary lesion' within a targeted area of the brain, allowing its underlying function to be investigated. A limitation of TMS is that it can only be applied to outer neural areas, as targeting inner areas would also affect any overlying neurological activity.

In a TMS study by Jones and colleagues (2004), participants reproduced empty visual durations (delineated with brief visual markers) from either a sub-second range (400 – 600ms) or a supra-second range (1600 – 2400ms), during which TMS was applied to either the right dorsolateral prefrontal cortex (DLPFC), supplementary motor area (SMA) or a control site (involved in leg movement) (Jones et al. 2004). Additionally, in different blocks the TMS was either applied during the encoding phase (during presentation of the standard duration) or during the reproduction phase (when participants had to reproduce the standard duration). The results showed that applying TMS to the SMA had no significant effect across either duration range, but there was a significant effect of TMS to the right DLPFC for supra-second durations (but not sub-second durations) when applied during the reproduction phase. This was argued to implicate the right DLPFC in the

memory/decision processes involved in supra-second timing, but makes the role of SMA in temporal processing more questionable.

Additionally, disruption of cerebellar activity through TMS can cause significant impairment of sub-second timing (Fierro et al. 2007; Koch et al.

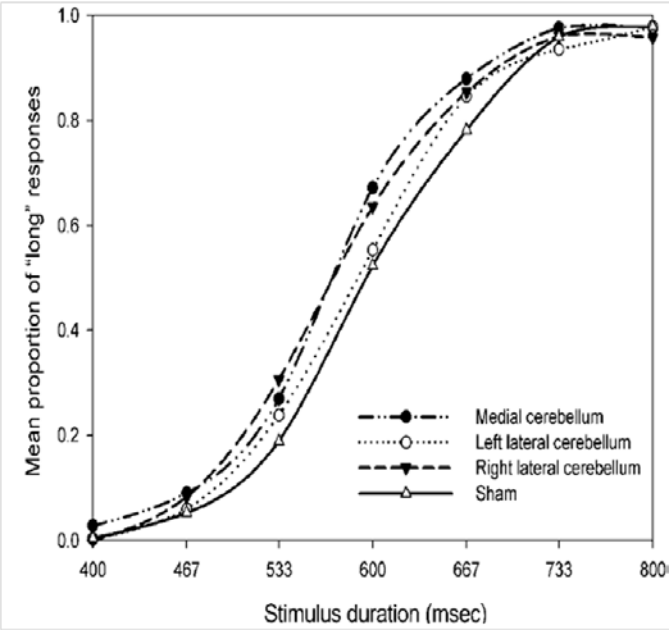

Figure 2.47: Mean data from a study by Lee and colleagues, showing the proportion of 'long' responses against physical test stimulus duration in a temporal bisection task. Repetitive TMS was applied to the right, left and medial cerebellum in separate blocks, and the 'sham' TMS acted as a control (in which the TMS coil was held 90° from the scalp). Figure reprinted with permission from MIT Press Journals.

2007) as well as an expansion of perceived duration (Lee et al. 2007). In the study by Lee and colleagues (2007), eleven healthy participants were initially trained to identify two standard durations, one 'short' (400ms) and one 'long' (800ms). In separate blocks, repetitive TMS was then applied to either the right, left or medial cerebellum, after which the participants performed a temporal bisection task. This involved judging whether the duration of an auditory 'test' stimulus was closer to the 'short' or 'long' duration, and responding with a key press. Sham TMS acted as a baseline, during which

the TMS pulses were directed perpendicular to the cerebellum. The results showed a leftward shift of the psychometric functions (indicative of an expansion of perceived duration) when TMS was applied to the right and medial cerebellum compared to the sham TMS (see Figure 2.47), although interestingly there was no significant effect of TMS to the left cerebellum. A subsequent experiment examined a longer, supra-second duration range (1000 – 2000ms), and found no significant effect of TMS on any region of the cerebellum. Lee et al. therefore concluded that the cerebellum has a role in encoding sub-second (but not supra-second) durations, supporting the notion that the processing of different duration ranges may be dissociated within the brain.

A more recent study has also implicated the visual areas V1 and V5 in the encoding and short term memory of sub-second durations (Salvioni et al. 2013). Observers performed a 2AFC duration discrimination task between two empty intervals marked by brief visual flashes: a 200ms standard and 200+Δms comparison (which varied in duration via an adaptive procedure). TMS was applied to V1, V5 or a control site in separate blocks (see Figure 2.48). In addition the temporal position of the TMS was varied in two separate experiments, so that it either followed the offset of the first flash (i.e. throughout the first empty duration – Experiment 1) or the offset of the second flash (i.e. during the interstimulus interval separating the two unfilled intervals – Experiment 2), at three possible delays (50ms, 85ms or 120ms). The former was hypothesised to disrupt the temporal encoding of the first stimulus and the latter to disrupt the retention of the temporal information in short term memory. Performance on the task was significantly impaired by

TMS for both cortical locations in Experiment 1, suggesting that both V1 and V5 have a role in encoding temporal information. Similarly, performance was reduced for both cortical locations in Experiment 2, but now at different delays (50ms for V1 and 85ms for V5), suggesting that both are *independently* involved in retaining temporal information. TMS was

Figure cannot be displayed due to copyright law. To view the original figure please see Salvioni et al. (2013), Figure 1 (p12425).

Figure 2.48: Schematic showing the experimental paradigm employed by Salvioni and colleagues (2013). In the above example the 200ms standard duration (T) is shown prior to the variable comparison (T+ΔT), although this was randomly interleaved within a block. In different experiments TMS was applied following either the offset of the first marker (during the encoding phase – Exp 1) or the offset of the second marker (during the storage of the first duration in short term memory – Exp 2). Participants then made a duration discrimination judgement as to "which was longer?" between the first and second duration, after which they received feedback.

ineffective in disrupting performance on a brightness discrimination control task (demonstrating that reduced temporal performance was not due to difficulty in stimulus detection). The results of this study imply that areas within the early visual sensory cortices have a role in encoding and storing duration information, supporting the notion of local, sensory specific timing (see Section 2.3).

2.5.3.2 Clinical populations

To establish a link between certain brain areas and timing ability, patients with specific neurological disease or neurological damage have been subject to investigation. However, interpreting these results must be approached with caution, to distinguish true temporal processing impairments from the general functional decline associated with these neurological diseases. For example, lesions of the prefrontal cortex have been associated with poor timing performance (Nichelli et al. 1995; Harrington et al. 1998b; Mangels et al. 1998; Casini and Ivry 1999), but this may reflect a more generalised deficit in memory, attention or decision processes. In addition, most lesion studies focus on neural regions known to be involved in motor function (e.g. cerebellum, supplementary motor area, basal ganglia). Whilst reduced performance following damage to these areas has been reported for motor timing tasks in the sub-second range (Ivry and Keele 1989; Halsband et al. 1993; Harrington et al. 1998a; Spencer et al. 2003), it is important to make a distinction between general motor deficits associated with damage to these regions (e.g. bradykinesia, tremor, poor coordination etc…) and genuine timing deficits. One way to achieve this would be to examine perceptual timing tasks, which don't rely on a speeded or accurate motor response.

Basal ganglia dysfunction is a characteristic of Parkinson's disease, in which gradual atrophy of the substantia nigra results in loss of neurons and a diminished supply of the neurotransmitter dopamine to the striatum. Parkinson's disease has been shown to result in impaired discrimination performance in both duration (Artieda et al. 1992; Rammsayer and Classen

1997; Harrington et al. 1998a), and rhythm discrimination tasks (Grahn and Brett 2009).

For example, Harrington and colleagues (1998a) investigated the ability for 34 medicated patients with Parkinson's disease to perform a duration discrimination task using empty stimuli marked with brief auditory tones. Standard durations were either 300ms or 600ms (tested in separate blocks), and comparison durations varied according to an adaptive staircase procedure. Observers indicated whether the comparison duration was longer or shorter than the standard. A frequency discrimination task was performed

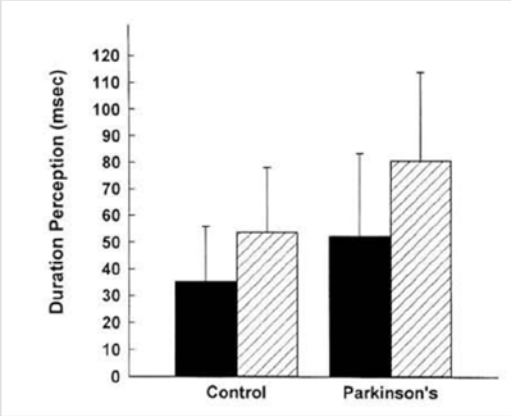

Figure 2.49: Observer thresholds ("Duration Perception") for a duration discrimination task, in which standard durations were either 300ms (solid blocks) or 600ms (hashed blocks). Patients with Parkinson's disease had significantly higher thresholds compared to age matched controls. Figure reprinted from (Harrington et al. 1998a) with permission from the American Psychological Association.

as a control, in which observers had to determine whether the comparison tone was higher or lower in pitch. This served as a measure of general cognitive or auditory processing impairment. Compared to age matched controls, Parkinson's sufferers had significantly higher discrimination thresholds for both standard durations on the duration task (see Figure 2.49), but measures of perceptual bias were similar across both groups (with no

significant differences found). No differences were found between controls and Parkinson's sufferers on the frequency task, suggesting that the basal ganglia are involved in sub-second duration processing, and impaired performance on the duration discrimination task was not a reflection of a generalised cognitive impediment.

Yet the importance of the basal ganglia to temporal processing has been challenged by Wearden and colleagues (2008), who investigated the ability for patients with Parkinson's disease to perform a series of perceptual tasks (e.g. verbal estimation, temporal generalisation, temporal bisection, duration discrimination). Wearden et al. found that regardless of whether the patient was on or off medication, timing ability remained reasonably comparable to age matched controls. Furthermore, experiments conducted on patients with bilateral basal ganglia lesions found no performance deficit in time estimation tasks compared to healthy controls (Coslett et al. 2010), and Smith et al. (2007) have reported that Parkinsonian performance deficits on a duration bisection task were limited to supra-second durations, suggesting that healthy basal ganglia may not be necessary for accurate sub-second timing.

Damage to another motor area, the cerebellum, has also been associated with human performance deficits in perceptual timing (Ivry and Keele 1989; Nichelli et al. 1996; Mangels et al. 1998). Mangels and colleagues (1998) demonstrated that patients with cerebellar lesions had significantly higher discrimination thresholds on a sub-second duration discrimination task compared to healthy controls, yet showed no impairment on a frequency discrimination task. However, the type and/or location of cerebellar damage may be important when interpreting the role of the cerebellum in duration

processing. Harrington and colleagues examined temporal reproduction (motor task) and duration discrimination (perceptual task) in 21 patients with focal cerebellar damage following stroke. Performance was not consistently impaired in either task across the group of patients, perhaps as a result of variation in disease location. Arguably, in cases where damage to the cerebellum is unilateral, temporal processing could be maintained by the intact half of the cerebellum (Ivry and Spencer 2004b). This in turn would explain why greater reductions in timing performance (but not performance on a loudness perception task) are seen in patients with more generalised cerebellar atrophy (e.g. Ivry and Keele 1989).

2.5.4 Summary

The relatively large number of neural regions implicated in timing, coupled with an inability to find any convincing components of an 'internal clock', lends weight to suggestions that temporal processing is mediated by a network of distributed neural regions working in conjunction with one another, possibly on a task dependent basis (Teki et al. 2011; Wiener et al. 2011; Merchant et al. 2013a). Two proposed circuits (based on persistent involvement of these neural regions across the literature) are a fronto-striatal loop involving the premotor cortex and basal ganglia (Coull et al. 2004), or a cortico-thalamic-basal-ganglia (CTBG) core timing circuit that is linked to other specialised sensorimotor circuits depending on task relevance (Meck et al. 2008; Merchant et al. 2013a) (see Figure 2.50). An ongoing challenge for researchers exploring temporal processing within these regions will be to consolidate the neurophysiological data with the psychophysical data.

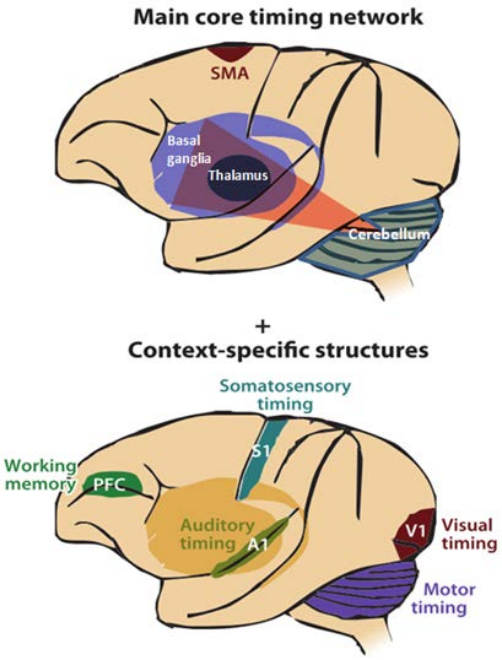

Figure 2.50: Proposed timing circuit involving a central core network with access to specialised areas on a task-dependent basis. Figure reprinted from (Merchant et al. 2013a) with permission from Annual Reviews.

Chapter 3: General Methods

3.1 Introduction

Psychophysics is a technique that measures the relationship between the world around us and our corresponding sensory experience. By physically altering a stimulus along a given perceptual dimension (e.g. orientation, weight, colour etc.) and measuring an observer's judgements along the same dimension, it is possible to map out the relationship between sensation and reality. Since this thesis is concerned with the perception of time, this chapter will first introduce the central concepts in psychophysics and then discuss the methods available for the psychophysical investigation of duration.

3.2 Signal detection theory and decision types

Sensory information is always embedded in 'noise', which may arise via moment-to-moment variation in neural activity (be it stimulus driven or a product of ongoing background activity) or external variation in the signal itself. Thus, the observer must decide whether their sensory experience is due to noise alone (N), or the combination of a signal plus noise (SN). An observer's ability to detect an afferent signal under conditions of uncertainty can be modelled mathematically (Green and Swets 1966).

Take, for example, a task where the observer must detect the presence of a light. On each trial they have two choices: "yes, the light is present" or "no, the light is absent". For each trial this observation can be modelled by two probability density functions (see Figure 3.1). The observer must set some

‘cut off’ point, or internal criterion, beyond which the probability that the light is present sufficiently outweighs the probability that the change in neural response is attributable to noise alone. Whilst this criterion may result from an unconscious bias, it may equally arise from a conscious decision to respond a certain way, for example in order to maximise the number of correct responses. The internal criterion is not necessarily a fixed value, and may vary from trial to trial.

The discriminability of the signal from the noise is given by the value d-prime (*d’*), which relates to the degree of overlap between the two functions. When the signal is weak (e.g. the light is dim), *d’* is small, and the probability of the resulting sensory experience occurring from either N or SN is similar. However as the signal strength increases (i.e. the light becomes brighter), *d’* also increases, and the probability of the sensory experience originating from SN becomes greater. The discriminability varies not only with the lateral separation of the two functions (i.e. increase in signal strength), but also with the spread of each function.

On each trial there can be one of four possible outcomes: the observer will correctly identify the light (a ‘hit’), correctly identify when there is no light (‘correct rejection), report a light when only noise is present (a ‘false alarm’) or miss when the light is present (a ‘miss’). A liberal observer might set a low criterion value, resulting in more hits, but also more false alarms. However a conservative observer might set a high criterion value, resulting in fewer false alarms but more misses. The position of this criterion value can therefore have a direct influence on any data obtained from the task. This is an

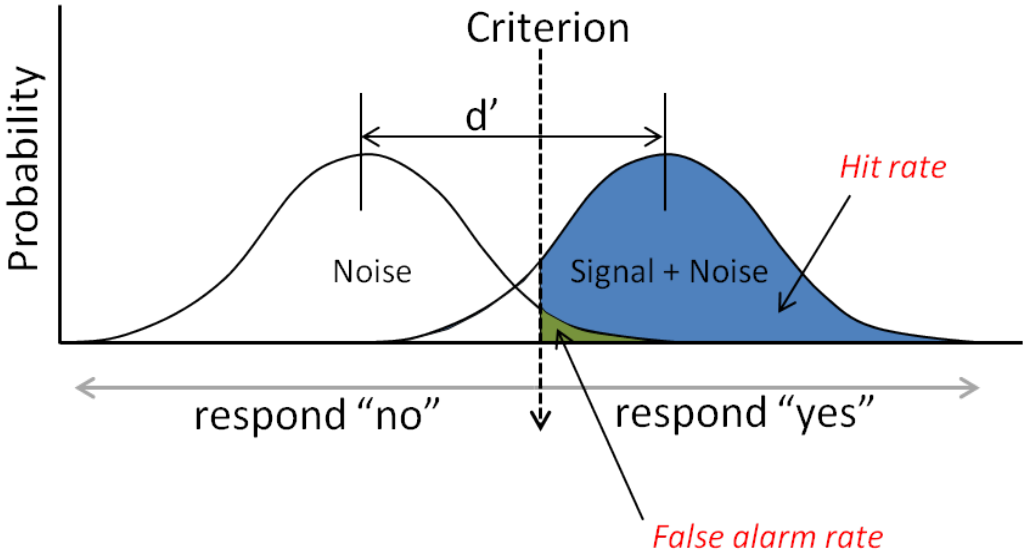

Figure 3.1: Schematic showing probability density functions for noise alone (N) and signal plus noise (SN). The x axis represents the observer's internal response, such that 'yes' responses will be given if an observation falls to the right of the criterion value, and 'no' responses will be given to the left of the criterion value. The hit rate refers to 'yes' responses which correctly identify the signal against the background noise. The false alarm rate refers to 'yes' responses which have incorrectly been made in the presence of noise alone. The discriminability of the signal is given by d-prime (*d'*), which relates to the degree of overlap between each function.

important consideration when designing a psychophysical experiment, as the type of decision can directly affect measurements of performance. Importantly, although the above example describes a detection task, the influence of criterion is equally significant during discrimination tasks that require a 'yes/no' response, where the observer might be required to decide whether or not two stimuli are the same.

Since all decisions are made against a background of signal noise, no method is truly objective, but it is important to be aware that methods requiring a 'yes/no' response from the observer are more 'criterion

dependent'. In these instances it is helpful to gain a measure of the observer's criterion by inserting 'noise only' trials, in which the signal is not present. The probability of the observer responding 'yes' during these trials can be used to determine the observer's criterion (calculated from the false alarm rate), which can then be taken into consideration alongside measures of their detection / discrimination ability.

Psychophysical judgements that do not require a 'yes/no' response limit the influence of criterion. A popular method is to use a "forced choice" comparative judgement, in which the observer is shown two (or more) choices and has to decide between them. An example from the temporal domain could be asking an observer to report "which flash of light has the longest duration, the first or second?" This removes the option to respond "no" or "I'm not sure", and forces the observer to guess even when uncertain (e.g. when the durations are close to their discrimination threshold). Response biases are less common using this paradigm, as there is no obvious advantage to selecting one option over the other (i.e. it removes any strategy to maximise the 'hit rate' or minimise 'false alarms'). However, in some circumstances an observer may demonstrate a "time order error" (TOE) (Fechner 1860), which refers to a bias in judgements that can occur when two stimuli are presented sequentially. In the temporal domain, it manifests itself as an under/overestimation of one stimulus' duration which depends on whether it was presented first or second. In the 'method of constant stimuli' paradigm (see Section 3.4.5), observers are presented with a reference stimulus of fixed duration, and a variable duration test stimulus

(which varies around the reference). On each trial the observer must decide "which stimulus was longer?" If the reference stimulus is always presented before the test stimulus a 'positive' TOE may result: an overestimation of the reference stimulus. A negative TOE would refer to an underestimation of the reference stimulus.

The stimulus range used may influence the TOE, as small ranges can be associated with a positive TOE and large ranges with negative TOE (Jamieson and Petrusic 1975). Also the magnitude of the duration may also affect the TOE with short reference durations (<100ms) giving positive TOEs and long durations (>1000ms) giving negative TOEs (Hellström and Rammsayer 2004). Models of the TOE suggest that it occurs because the observer compares the second stimulus with a weighted average of both the first stimulus and some internal estimate of sensory magnitude, which may change as a result of recent stimulus history (Hellström 1985; Hellström and Rammsayer 2004).

Therefore, when designing an experiment it is important to take precautions to eliminate TOE biases. One way to achieve this is to randomly interleave the presentation order of the stimuli (Gescheider 1976). This way the positive and negative TOEs should cancel each other out when averaged across trials.

3.3 Psychophysical measures

3.3.1 Sensory threshold

A central concept in psychophysics is the measurement of the sensory threshold. This may refer to the amount of stimulus energy required to first elicit a sensation (the detection threshold), or the change in stimulus energy required for the change to be detected by the observer (the difference or 'discrimination' threshold). The lower an observer's threshold, the more 'sensitive' they are to changes in the stimulus, and thus threshold and sensitivity are inversely linked. The experiments included in this thesis are chiefly interested in the smallest detectable difference between two stimuli, which may also be referred to as the 'just noticeable difference' (JND) between two stimuli.

Weber's law describes a proportional relationship that can exist between a stimulus and the observer's discrimination threshold for a change in that stimulus, and states that the threshold will increase proportionally with the magnitude of the stimulus being judged. Conformity to Weber's law may be demonstrated by calculating the Weber fraction, which is achieved by dividing the threshold value by the stimulus magnitude. This ratio would remain constant across stimulus magnitude if Weber's law holds (Gescheider 1976). This fraction can be useful when comparing sensory discrimination across different modalities as it normalises the measurements. As discussed previously in Chapter 2, in timing studies it is a variation of Weber's law known as the scalar property which is often reported (see Section 2.2.1).

This states that the standard deviation of an observer's judgements is proportional to the magnitude of the estimated duration.

3.3.2 Perceptual bias

In addition to measuring the sensory threshold (JND), psychophysical methods may also aim to measure perceptual bias. This can be achieved by collecting quantitative estimates of duration, (e.g. where the observer estimates the stimulus duration on each trial), or by making a comparative judgement (e.g. perceived duration is derived from the probability of the observer responding a certain way). The non-temporal content of the stimulus (e.g. pitch, luminance, magnitude, colour etc…) can also be modified to examine how perceptual bias may be altered under different experimental conditions. A popular method for measuring perceptual bias is the construction of a psychometric function, which plots the probability of a certain observer response type (e.g. 'stimulus 1 was longer than stimulus 2') against physical stimulus duration. From this, a measure of the subjective duration can be found by determining the physical duration that corresponds to chance performance – known as the Point of Subjective Equality (PSE). This technique will be discussed in further detail in Section 3.5.

3.4 Methods used for the study of duration

There are a number of methods that may be used to study duration perception in humans. These may first be categorised as either prospective or retrospective depending on whether the participants are explicitly aware

that the task involves a temporal judgment. Prospective paradigms describe tasks where the participants are informed at the start of the experiment that they will be making a duration judgment (e.g. Klink et al. 2011). In retrospective paradigms, participants often perform a non-temporal task and are then asked afterwards to make a judgment about duration (e.g. Dong and Wyer 2014). As such, retrospective tasks often rely heavily on memory processes, and temporal judgments tend to be shorter and more variable (Brown 1985). The experiments described in this thesis use prospective paradigms, and therefore the main methods for examining duration in a prospective manner will now be discussed.

3.4.1 Method of limits

For ease of explanation, we will arbitrarily refer to two stimuli as A and B. The method of limits requires that stimulus A remains constant while the magnitude of B is adjusted relative to A in either an ascending or a descending manner to find a measure of the JND (e.g. Zihl et al. 1983; Rammsayer 1999). The amount by which B is adjusted on each trial is known as the ‘step size’, and is a predetermined, fixed amount. In the ascending version A and B may have physically identical durations at the start of the task. To find the JND, B is then adjusted until the observer first reports that they can detect a difference between the two. At this point the difference in duration between A and B gives a measure of JND. In the descending version the experimenter will ensure that they begin from a position above the observer’s threshold (i.e. there is an obvious difference

between A and B), and reduce the magnitude of B until the observer first reports that they can no longer detect a difference in the stimuli.

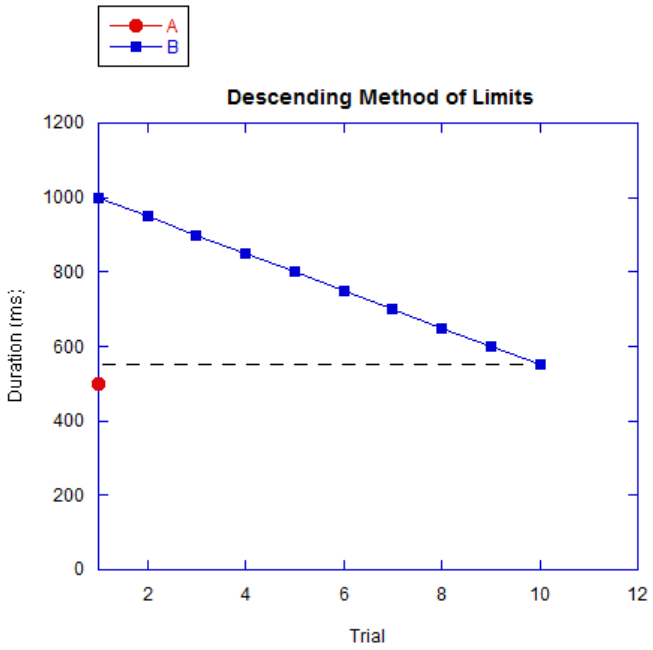

Figure 3.2: Showing the procedure for measuring the JND using the descending method of limits. The dotted line represents the point where the observer can no longer perceive a difference in stimulus A (the red data point) and B (blue data points), and hence this gives the threshold. The JND is the difference between this value and the magnitude of stimulus A (500ms in the above example).

For example, Figure 3.2 shows that the representative observer reports no difference between A and B when B is 550ms and A is 500ms, and so consequently the JND could be described as 50ms. Often multiple repeats of the task are performed and an average of all the readings is given as the final value.

The method of limits is simple to perform, yet it can be time consuming and is often subject to two types of observer error. The error of habituation occurs when observers start to habitually give the same response on each repeat of the task, which may induce a loss of selective attention. For example, if the

step size is too small for the observer in a descending limits situation the observer will initially experience a large number of trials where the difference between A and B is grossly suprathreshold, leading to the task being performed with a reduced degree of attentional focus. This could lead the observer to continue to respond ‘A and B are different’ after a point where heightened attention would have detected that the two are now indistinguishable. This would lead to a spuriously high measure of the observer’s sensitivity.

An error of expectation occurs when observers begin to anticipate the stimulus threshold and give a premature response, causing either a spuriously high threshold in descending trials or a spuriously low threshold in ascending trials. Also, as mentioned earlier, because the method of limits depends on the observer’s internal criterion of whether the stimulus can be detected (e.g. ‘are the stimuli the same? yes/no?’), less confident observers could decide to wait until they are certain that there is a difference between two stimuli before responding, and more confident observers might respond at the merest hint of a difference. This can lead to a large variation in observer responses. Another disadvantage of this method is that the exact measurement of the threshold might in fact lie somewhere between the penultimate and the final presentations, and therefore the accuracy is dependent on the step size used.

3.4.2 Staircase method

A variation of the method of limits is the ‘staircase procedure’, where the stimulus magnitude of B is first adjusted either in an ascending or descending manner until the subject responds. At this point the direction of adjustment is reversed until the subject responds again, and this continues until a pre-determined number of reversals (e.g. fifteen) have been made. The average of the last few predetermined reversals (e.g. the last five transition points) gives the threshold (e.g. Kanai et al. 2011). This method can also be used to gain a measure of perceptual bias by finding the duration at which the observer is equally likely to respond that A and B are the same/different (or that B is shorter/longer than A). In this manner, staircase reversals for stimulus B include durations that are both shorter and longer than stimulus A, which remains constant.

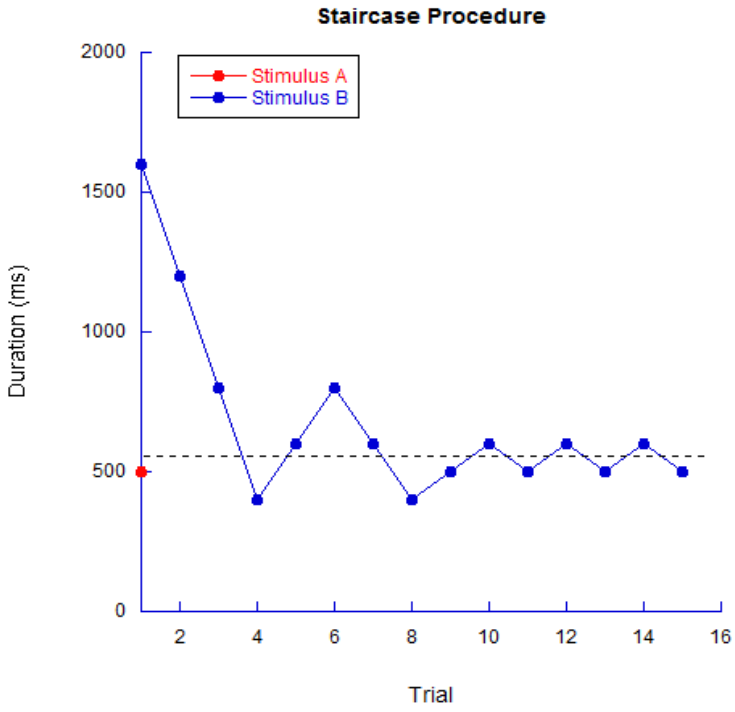

Figure 3.3: Showing the procedure for measuring the JND using the staircase method of limits. In the above example stimulus A (the red data point) is 500ms, and stimulus B (blue data points) is adjusted in a descending or ascending manner until a predetermined number of reversals have occurred. The dotted line represents the average of the last (in this case) five reversals, which can be taken as a measure of discrimination threshold. Once again it is the difference between this value and the magnitude of A that gives the JND.

An example of measuring the JND using a staircase procedure is shown in Figure 3.3 (using the same arbitrary values as for Figure 3.2). This method is usually faster than the standard method of limits as large step changes can be made initially prior to 'homing-in' on the threshold via progressive reduction in step size. However staircase techniques could be limited by the fact that the final number of trials are all close to threshold, and therefore difficult to discriminate. In this situation, the observer may make more criterion dependent errors towards the end of the trial sequence, and it is the average of these results which contribute to the threshold.

3.4.3 Magnitude estimation methods

Magnitude estimation can measure the perceived temporal extent indirectly via labelling procedures such as verbal estimation or directly by reproduction/production tasks. Production tasks involve asking the observer to produce a specific duration by pressing a button for the prescribed length of time. Reproduction tasks require the observation of a stimulus followed by executing a motor action (typically a button press) for a period that matches the subjective impression of the stimulus' duration (e.g. Jones et al. 2008; Lewis and Miall 2009).

Comparing these reproduced durations with physical duration allows measurement of any perceptual bias: either under or overestimation of stimulus duration. Reproduction has an advantage over duration discrimination in that bias estimates derived from the latter could reflect an overestimation of one stimulus or an underestimation of the other – relative judgments cannot easily distinguish between the two. Reproduction on the other hand gives an absolute measure of duration making it easier to assess the nature of any bias specific to the stimulus of interest.

A variation of the reproduction task is the fractionation task, where observers are requested to reproduce a proportion of the duration, for example by pressing the button for what they perceive to be half of the original duration (e.g. Warm et al. 1975). In all of these methods a large number of trials are often performed and the data averaged together. The resulting mean value gives a measure of perceptual bias and the variance (i.e. spread of responses) gives a measure of observer performance.

Verbal estimation tasks might require observers to verbally estimate the duration of a stimulus, for example ‘2 seconds’ (e.g. Penton-Voak et al. 1996; Gil and Droit-Volet 2011), or assign a value relating to the magnitude of the duration which is based upon their own internal scale (e.g. a multipoint labelling scale ranging from ‘very long’ to ‘very short’ - see Bobko et al. 1977 and Tse et al. 2004 for examples).

Whilst reasonably quick to perform, magnitude estimation tasks are inherently subjective, and as a result can be highly criterion dependent. In addition, these tasks have been criticised for conforming to Vierordt’s Law, where short durations have a tendency to be overestimated and long durations underestimated (Wearden and Lejeune 2008b; Lejeune and Wearden 2009).This phenomenon occurs when a task includes a range of standard durations and on any given trial the observer must reproduce (or produce, estimate, categorise etc…) a duration chosen from within this range. It is suggested that this represents a tendency to regress towards the mean duration within the range, as the crossover from overestimation to underestimation (the “indifference point” – at which observers are veridical) often occurs close to this value (Vierordt 1868). Conformity to Vierordt’s Law subsequently produces a violation of scalar timing, which posits that the mean estimate for each standard duration should increase linearly (see Chapter 2 Section 2.2).

3.4.4 The method of single stimulus

In this procedure, a series of stimuli differing along a given dimension (e.g. duration) are presented one at a time in a sequence of trials. Whilst several single stimulus methods exist, two popular paradigms are the method of temporal bisection and the method of temporal generalisation, which will be discussed in further detail below.

3.4.4.1 Temporal bisection

The method of temporal bisection originated from studies on animal timing (e.g. Church and Deluty 1977), and was later adapted for humans (Allan and Gibbon 1991). Initially observers are trained to recognise two reference durations, T_{short} and T_{long}. On each trial a single test stimulus is presented, and the observer must decide if it is closer to the remembered values of T_{short} or T_{long} by responding 'short' or 'long' (e.g. Allan 2002). This method therefore utilises a forced choice procedure, where the observer is required to make a decision on each presentation of the stimulus which cannot be answered 'yes' or 'no'. Even if the observer is maximally uncertain about whether the test stimulus is closer to T_{short} or T_{long}, they are still required to guess, which helps to reduce the influence of observer criterion.

At the end of the test phase the proportion of 'long' responses is plotted against the physical duration of the test stimuli (see Figure 3.4). The 'bisection point' (BP) of the function is identified as the x-axis value (i.e. the physical duration) which corresponds to 50% 'long' responses. The duration discrimination threshold is determined from the slope of the function. If the

observer is veridical, the BP will correspond to the physical midpoint between the T_{short} and T_{long} durations (Figure 3.4, red data). Perceptual bias will shift the BP to a value which does not correspond to the midpoint (Figure 3.4, blue data). A variation of this method occurs when the reference durations

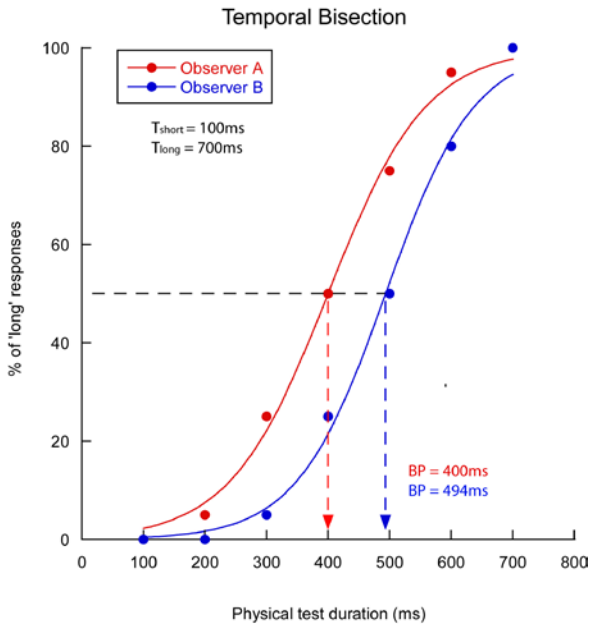

Figure 3.4: Demonstrating the procedure for extracting the bisection point (BP) in a temporal bisection task. In this example observer A (red data) is veridical, as the BP lies at the midpoint of T_{short} and T_{long}. However the function for observer B (blue data) is laterally shifted to the right representing a perceptual bias. This observer is giving fewer 'long' responses, signifying a temporal compression. The BP value of 494ms suggests that this duration now feels perceptually halfway between 100ms and 700ms, and thus must feel 'short' compared to its true physical duration.

are not identified at the beginning of the task, and is referred to as the 'partition method' e.g. (Wearden and Ferrara 1995). Here, the observer must classify test durations as short or long by building up an internal mean from all of the test stimuli, and comparing the stimulus on each trial to this value. Initially, discrimination judgments will be at chance performance levels until

sufficient test stimuli have been presented for the observer to build up an internal representation of the mean duration.

The method of bisection rests on several assumptions, including that an observer will give equal numbers of ‘short’ and ‘long’ responses when the duration of a stimulus lies at the perceptual midpoint. However, if an observer decides arbitrarily to respond ‘long’ each time they are uncertain, this would cause a shift in the BP that mimics a perceptual bias, but in fact reflects a response bias. It also assumes that the neural representation of the reference durations (learnt during the training phase) remains stable through the test phase, and that the subjective midpoint (used for deciding ‘long’ or ‘short’) remains stable at the arithmetic mean. If these assumptions are violated it would complicate the interpretation of temporal bisection data, and yet to obtain evidence either way has been deemed “impossible” (Garcia-Perez 2014).

3.4.4.2 Temporal generalisation

The method of temporal generalisation again originated from the field of animal timing (Church and Gibbon 1982) and was adapted for use in humans (Wearden 1991a; Wearden 1992).The technique involves an initial learning phase, in which the observer is repeatedly shown a reference stimulus of fixed duration. On each trial, a test stimulus is presented (from a range of different durations centred on the reference), and the observer must decide whether the test stimulus was the same as the reference duration (“yes”

response) or not the same (“no” response). The number of ‘yes’ responses are then plotted against the physical test durations. Trial to trial variation in timing, memory and decision processes, combined with an appropriate range of test durations results in a spread of ”yes” responses (see Figure 3.5), where the peak of this function will correspond to the standard duration if the observer is veridical, or be shifted laterally if they are biased.

The position of this peak may be compared across experimental conditions, for example by altering a non-temporal aspect of the test stimuli (e.g. auditory pitch) whilst keeping the reference the same. Performance on

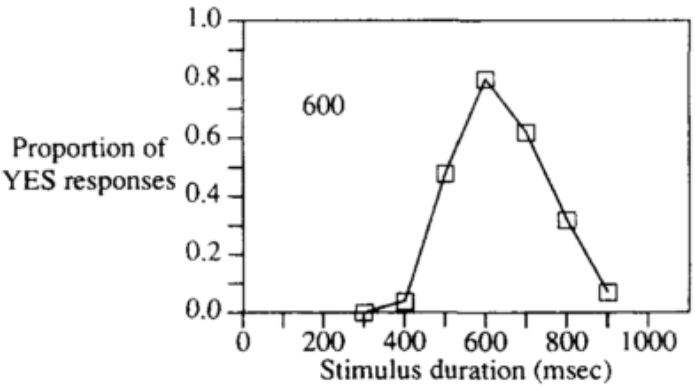

Figure 3.5: The results of a temporal generalisation task where the reference duration was 600ms (Wearden 1992). The peak of the function in this example corresponds to 600ms, suggesting that in this instance the observer was behaving veridically. Figure reprinted with permission from the American Psychological Association.

temporal generalisation tasks is extracted from the spread of the function, yet this may be dependent upon both the observer’s sensory threshold and their criterion for responding. A large JND and a high criterion for responding ‘different’ would both result in a wide function, compared to a small JND or

low criterion for responding ‘different’, and it is particularly difficult to separate out these two underlying factors.

For example, superior observer performance (i.e. smaller JNDs) has been reported with increased task difficulty, and this unexpected finding was attributed to changes in observer criterion (Ferrara et al. 1997). Two groups of 20 observers performed the same temporal generalisation task, but the step size between test durations was 75ms in one group (resulting in a narrower range) and 150ms in the second group (resulting in a broader range). The group with the narrower range of test durations were shown to have higher temporal sensitivity, which was suggested to reflect a more conservative approach to responding “same”. This highlights the difficulty in extracting accurate measures of sensory thresholds using criterion dependent measures. In addition, the temporal generalisation method assumes: 1) that the probability of responding ‘yes’ is symmetrically distributed around the standard duration, which may not be the case due to the scalar property (as shorter durations have a lower associated variance), and 2) that the representation of the standard duration remains stable throughout the task, and both of these assumptions are very difficult to test experimentally (Garcia-Perez 2014).

3.4.5 Method of constant stimuli

The method of constant stimuli (MOCS) is used to measure an observer's responses over a fixed range of stimuli (e.g. Getty 1975), and is the method chosen for all of the experiments described in this thesis. It uses a pre-determined range of test stimuli which are presented to the observer with equal probability. A stimulus with a fixed duration serves as a reference point, and the test stimuli consist of a range of durations (often equally spaced in magnitude) centred around this reference stimulus. On each trial the chosen test duration and the reference duration are both presented and the observer must make a comparative judgement between the two, usually via a 'two alternative forced choice' (2AFC) procedure (see Figure 3.6). Measures of the sensory threshold and/or perceptual bias are often achieved

Figure 3.6: Schematic of the MOCS combined with a 2AFC forced choice decision. Observers must make a comparative judgement between a fixed duration reference stimulus (500ms in the above example) and a variable duration test stimulus (varying around the reference duration). Specific to MOCS, the presentation order of the different test stimuli is randomised such that all have an equal presentation probability on each trial.

by plotting the proportion of 'test longer' responses against the physical test durations, in what is called a 'psychometric function' (see Section 3.5). Multiple presentations of each reference/test pair are often performed to reduce the influence of observer error (e.g. accidently choosing the wrong

response). However this does have the disadvantage of making the paradigm time consuming, which additionally runs the risk of observers losing concentration. This is often counteracted by splitting data collection into shorter blocks, with breaks between sessions.

Randomising the presentation order of reference and test stimuli during the MOCS is also advantageous as it makes it difficult for observers to adopt a method of single stimulus approach, in which they might ignore the reference stimulus and choose to compare the test stimulus on any given trial to an 'internal mean' of the previously presented test stimuli.

3.5 The psychometric function

The probability of the observer reporting that the 'test was longer than the reference' can be plotted against physical test duration as a psychometric function (see Figure 3.7). When the test is obviously shorter than the reference, the observer should rarely report that the test has a greater magnitude, and so the proportion of the responses here will be close to zero. Similarly, when the test is obviously longer than the reference the observer should report that it has the greater magnitude on most of the trials. If a curve is fitted through the data points we assume that the test duration corresponding to the 50% response level is perceptually equal to the reference stimulus, since the observer is performing at chance. This is known as the Point of Subjective Equality (PSE). When physically matched durations induce perceptual equivalence, observers can be said to be

'veridical' (perception matches reality). In this situation, the PSE value will coincide with the physical duration of the reference stimulus. Alternatively, when a perceptual match is induced by physically discrepant stimuli, the PSE will be a value which is longer or shorter than the reference duration and the observer's perception of the test and/or reference duration can be said to be non-veridical or 'biased'.

The discrimination threshold (JND) can also be measured from the curve, and is taken as the value where observers can reliably report that the test stimulus is longer or shorter than the reference. Depending on the caution the experimenter applies to assuming exactly which performance levels

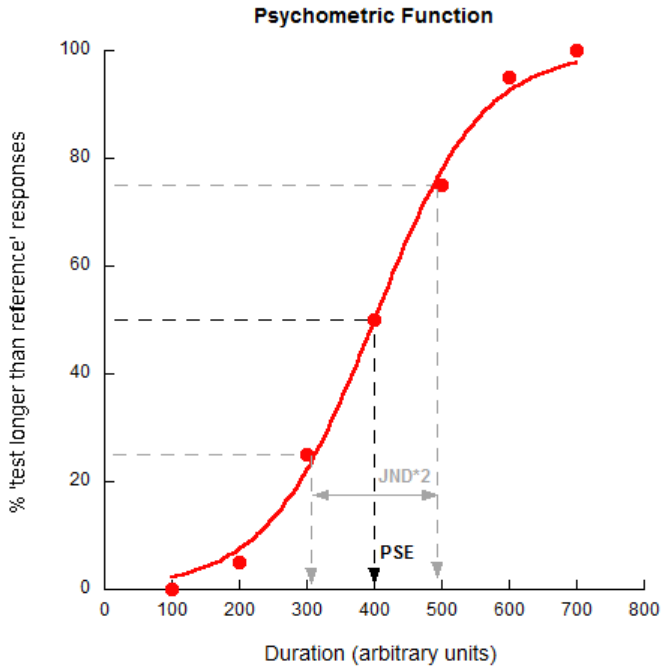

Figure 3.7: An idealised psychometric function. The PSE corresponds to the 50% test longer responses, and the JND may be taken as half of the difference between the 25% and 75% test longer responses.

represent ‘reliably reporting’ a change in the stimulus from PSE (50% performance), the values used to calculate JND might correspond to say, 25% and 75% ‘test longer than reference’ responses. However, a more conservative approach might see an experimenter selecting say, 10% and 90% performance levels. Whichever parameters are chosen, the JND is typically taken to be half the difference between the physical duration values corresponding to the selected (e.g. 25% and 75%) performance levels. This does however assume that an observer’s sensitivity to increases and decreases in duration is equal either side of the PSE.

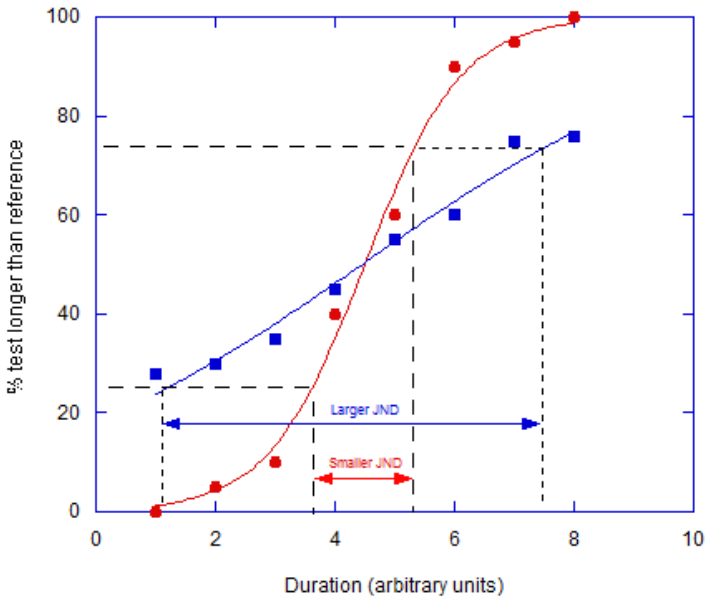

Figure 3.8: The slope of the psychometric function is related to the JND. The red curve has a steeper slope and consequently a smaller JND, and the blue curve is much flatter resulting in a large JND (as the observer is more uncertain of their decisions).

In addition to perceptual biases, the shape of the psychometric function can also reflect the variability of the observer’s responses. Specifically, the slope is directly linked to an observer’s sensitivity to stimulus change (see Figure

3.8). Low sensitivity is associated with a large spread of responses around the PSE and vice versa. The degree of this spread is an innate feature of observer performance. In order for an experimenter to capture this spread via a series of perceptual decisions, a range of stimulus values must be chosen which adequately samples performance across a range of uncertainly levels from, for example, ‘the first stimulus is obviously longer than the second’ and vice versa. If relatively large stimulus differences are presented when the observer has a relatively low discrimination threshold the observer’s innate performance level will be sampled too coarsely (Figure 3.9, right panel) where too few data points will contribute to the fit of the function and thus its parameters such as PSE and JND values. Equally, a small range of test durations presented to an observers with a relatively low discrimination threshold (Figure 3.9 left panel) will sample performance across a range where all response levels (i.e. degrees of perceptual certainty) are too close to chance performance. Either approach provides spurious and/or uncertain measures of performance.

In order to generate a reliable estimate of threshold, ideally the responses should not be saturated at either end of the function, and also they should extend beyond the 25% and 75% ‘test longer than reference’ response rate (as per Figure 3.7).

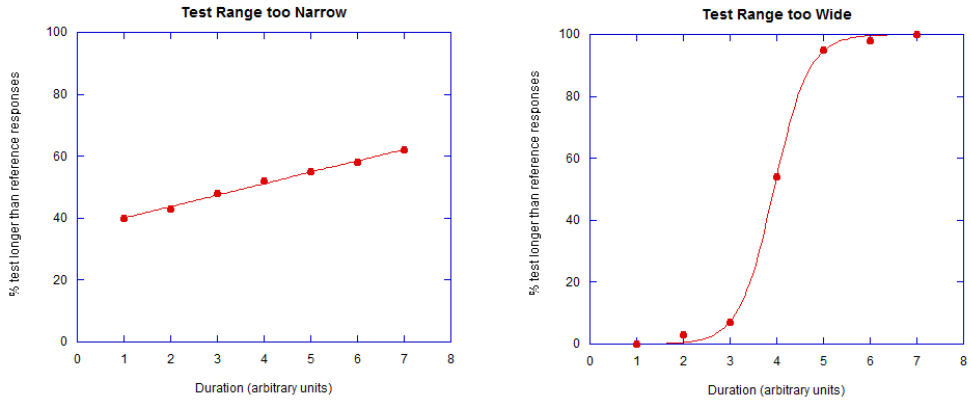

Figure 3.9: Showing example psychometric functions for a narrow range (flatter function) and a wide range (steeper function) of test durations.

3.6 Curve fitting

All experiments described in this thesis employed the two-alternative forced choice (2AFC) in combination with the method of constant stimuli. Psychometric functions were constructed by plotting the proportion of 'long' responses against test stimulus duration. Using the software package Kaleidagraph (Synergy Software: Reading, PA) each psychometric function was then fit with the logistic curve of the form:

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

Where μ is the test duration corresponding to the PSE and θ is an estimate of the discrimination threshold (JND).

The best fitting curve was determined by the software via a Levenberg-Marquardt algorithm, which is a technique used to apply a curve fit to non-linear data. The technique aims to minimise the sum of the squared error between each data point and the curve (represented by the "Chi Square"

value). The algorithm will work through numerous iterations of the curve fit based on some initial predictions for the PSE and JND values, which are entered into the software by the experimenter. With each new iteration the PSE and JND values are varied slightly and the Chi Square value is re-calculated, giving a measure of the 'error' between the data points and the curve (sometimes referred to as the "goodness of fit"). The algorithm re-evaluates and improves the curve fit parameters to find the smallest possible Chi Square value (and hence the 'best' fitting curve). Once the optimum fit is reached, the corresponding curve fit parameters of interest (PSE, JND) are recorded by the experimenter. Kaleidagraph also reports the standard error of the PSE and JND, which give a measure of the accuracy of each parameter.

It is important to be aware of the detrimental effect that outliers can have on this curve fit. Largely discrepant data points cause a disproportionately large effect on the calculation compared to other data points leading to inaccurate predictions. As a result, careful examination of the raw data with a specific policy for removing outliers, as well as ensuring that each data point represents a large number of responses will reduce this source of error.

Following extraction of the PSE and JND values, some of the experiments described in this thesis required further data analysis and additional curve fitting. These more specific curve fits are described within the methods section of each relevant chapter.

3.7 Bootstrapping

The experiments presented in Chapter 5 utilised a bootstrapping procedure to evaluate the accuracy of any experimental parameters obtained.

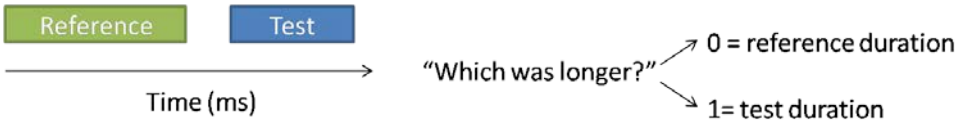

Figure 3.10: Schematic showing one experimental trial, in which the observer is presented with a reference – test duration pair and must decide "which duration was longer?" Choosing the reference duration gives a binary response of "0" and choosing the test duration gives a response of "1".

Commonly, bootstrapping procedures use a method known as 'sampling with replacement', in which a new data set is generated by randomly sampling the existing data. An alternative method, and the one used in this thesis, is to generate a new data set by randomly sampling values from a series of binomial distributions drawn from the original data.

Test Durations (ms)	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10	Total	Proportion
200	0	0	0	0	0	0	0	0	0	0	0	0
240	0	1	0	0	0	0	0	0	0	0	1	0.1
280	0	0	0	0	1	0	0	0	0	0	1	0.1
320	1	0	1	0	1	0	1	1	0	1	6	0.6
360	1	1	1	0	1	1	1	1	1	1	9	0.9
400	1	1	1	1	1	1	0	1	1	1	9	0.9
440	1	1	1	1	1	1	1	1	1	1	10	1

Figure 3.11: A table showing invented, representative raw data for a single observer. In this example only ten repetitions of each reference-test pair are shown for ease of explanation (labelled Run 1 – 10). The binary responses represent negative "test longer" responses (0) and positive "test longer" responses (1). The columns to the far right show the summed responses as both the total number of "test longer than reference" responses and the proportion of responses.

To begin, the raw data for each observer was stored in a table, which contained the 30 binary responses for each possible reference-test duration pair (see Figure 3.10), where ‘0’ represented a negative “test longer” response (i.e. the observer responded that the reference duration was longer than the test duration) and ‘1’ represented a positive “test longer” response (see Figure 3.11, where 10 repetitions are shown for simplicity). The total number of ‘test longer’ responses was then summed, and divided by the total number of repetitions to give each value as a proportion (see Figure 3.11, far right column). Using custom software running in Matlab (Mathworks, USA), a

Test Duration (ms)	Total number of “test longer” responses	“test longer” responses as a proportion
200	0	0
240	4	0.13
280	9	0.3
320	15	0.5
360	18	0.6
400	24	0.8
440	30	1

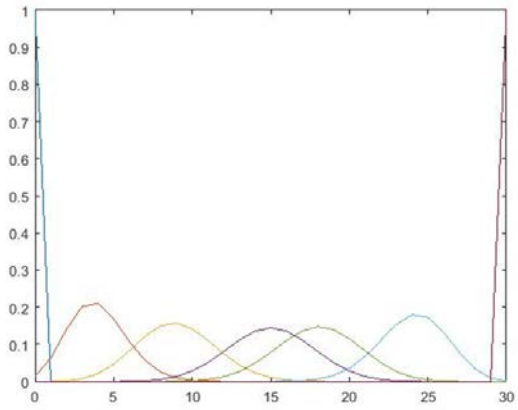

Figure 3.12: Colour coded binomial probability distributions for a representative data set, giving the probability that each value would be generated again if the task was repeated an infinite number of times. From each distribution a value is randomly chosen, converted into an integer and used to form a new, randomly sampled data set.

binomial probability distribution is calculated for each proportion (see Figure 3.12). This gives the probability of generating this value again if the task was repeated an infinite number of times, given that there are only two possible outcomes (0 or 1) on any given trial. It also takes into account the number of repetitions that were performed to generate the original data, which provides

minima (zero) and maxima (maximum number of repetitions per condition) for the corresponding binomial function's x-axis.

The bootstrapping programme then randomly selects an integer value from each probability distribution to generate a new set of resampled y values (probability of test longer responses). These values are then fitted with a logistic curve to generate new (resampled) PSE and JND estimates. The entire process is then repeated 1000 times (although 1000 iterations are commonly performed, this number is arbitrary and can be chosen by the experimenter), giving 1000 different randomly sampled PSE values. The non-parametric "percentile method" is then used to determine the 95% confidence intervals from the data (as the data is not assumed to be normally distributed). These intervals represent the range of values between which 95% of the bootstrapped PSE values fall, and can be used to represent the error for each observer's true PSE value (i.e. the PSE determined from the raw, non-bootstrapped data). The program ranks the PSE values from lowest to highest, grouping them into 5ms bins and then plotting them as a frequency distribution (see Figure 3.13). It then selects the 25^{th} and 976^{th} values (giving the lower and upper confidence intervals respectively).

If the observer's responses were highly variable (i.e., close to chance performance on most/all trials) there will be a greater spread in the PSE values calculated from each resampling procedure. Conversely, if the original responses were highly consistent across stimulus repetitions, then there will be greater agreement between the resampled PSE values (i.e. a narrow spread). Also, since the data is not necessarily normally distributed it should

be noted that the upper and lower confidence intervals may not be symmetrical either side of the mean PSE value.

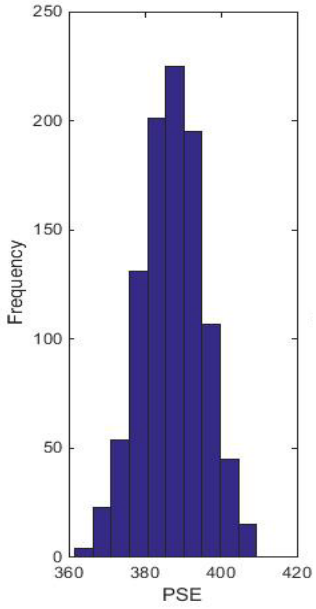

Figure 3.13: An example of a frequency distribution plot of PSE values (each bar represents a PSE bin – a range of PSE values spanning 5ms) based on 1000 samples. Upper and lower confidence intervals (CI) give the range between which 95% of the PSE values fall.

3.8 Apparatus

The experiments outlined in this thesis were controlled by a desktop computer running custom software written in Matlab (Mathworks; USA), with the Psychtoolbox extension (Brainard 1997). Visual stimuli were presented via a computer monitor, and auditory stimuli were presented using Sennheiser HD 280 headphones. Responses were collected using the computer keyboard. All experiments were conducted in a quiet, darkened room. Details specific to each experiment will be discussed in the methods section of relevant experimental chapters.

3.9 Experimental calibration

3.9.1 Gamma correction

Prior to each experiment, or the use of any new experimental equipment, each monitor was calibrated via the process of gamma correction. Because the relationship between the input voltage and the output luminance of a monitor is often non-linear, the requested luminance of a stimulus may deviate from that which the monitor delivers. Whilst the graphics card will send an output voltage to each pixel on the monitor corresponding to the requested grey level, the delivered grey level will be a power transform of the requested intensity.

To correct for this non-linearity via gamma correction, the output luminance of the monitor must be measured for a series of known voltage inputs (a series of chosen grey levels). Grey is used because the required intensity from each of the red, green and blue outputs of the monitor is equal. Greyscale images are often represented by intensity values which range from 0 to 255, where 0 is black and 255 is white. During gamma correction, requested grey levels are chosen by dividing the total number of grey levels (256) into equal linear steps then presenting each one sequentially through the monitor. These levels are sometimes normalised by the computer software to values between 0 and 1 by dividing each number by 255.

At each presented grey level, the monitor's output luminance is measured with a photometer. Luminance (cd/m^2) can then be plotted against the grey level values, and fit with a power curve of the form: $y = x^g$. The value of *g* is known as the gamma correction factor. Plotting the inverse curve ($y = x^{1/g}$) allows a new table of requested voltage values to be generated (for each of the 256 grey levels), which when presented, results in a linear luminance output.

Grey Level	Luminance reading prior to gamma correction (cd/m^2)	Luminance reading following gamma correction (cd/m^2)
0	0.05	0.05
0.1255	1.68	37.3
0.2510	9.45	73.1
0.3765	25.4	107
0.5020	50.8	142
0.6275	87.5	176
0.7529	137	211
0.8784	202	245
1	279	276

Table 1: An example of the luminance values taken both before and after gamma correction of an Eizo FG2421 monitor. Grey levels represent the voltage input sent to the monitor (normalised to range from 0 - 1), and the luminance readings give the monitor output values. Before gamma correction, the relationship is non-linear.

For the experiments described in this thesis, gamma correction was achieved via an in-built gamma calibration tool from Psychtoolbox called "CalibrateMonitorPhotometer". This program automatically selects nine grey

levels of 0, 0.13, 0.25, 0.38, 0.5, 0.63, 0.75, 0.88 and 1, which are presented sequentially on the monitor. Luminance readings were taken in a darkened room using a Minolta Chroma Meter CS-100 photometer mounted on a tripod. This was focused on the centre of the screen at a distance of one metre. Individual readings were entered into the program, and one example set of readings collected from the left Eizo FG2421 monitor prior to gamma correction are shown in Table 1 (middle column) alongside the corresponding grey levels.

These values were then plotted, and fit with a power function using Psychtoolbox (see Figure 3.14, left plot), where, for the readings shown in Table 1 the resulting gamma correction factor (*g*) was 2.59. The program then automatically generated a new series of voltage inputs using the inverse power fit, and stored these in a look up table (LUT). This LUT could then be called by Matlab before sending stimulus generation instructions to the computer's graphics card, thereby off-setting the monitor's inherent non-linearity.

To ensure that the gamma correction had been successful (i.e. its output was now linear), the linearised LUT values were implemented, and luminance recordings were repeated for all nine grey levels. The new, post gamma correction luminance values are shown in Table 1 (right column) and corresponding data points, fitted with a linear function, are shown in Figure 3.14 (right plot). The output luminance of the Eizo monitor is now linear with a peak luminance of 276cd/m^2. In experiments requiring the use of two monitors simultaneously (see Chapter 5), it was also vital to ensure that the

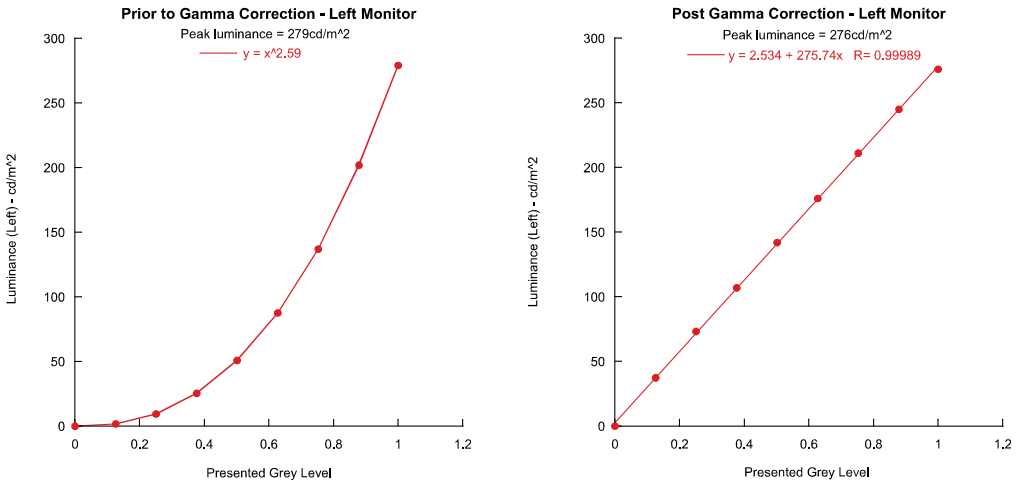

Figure 3.14: An example of the physical luminance output for a series of requested grey levels, using the same Eizo monitor readings as described by Table 1. The left plot shows the luminance levels prior to gamma correction, where *g* = 2.59. The right plot shows the luminance levels following gamma correction, where the output is now linear.

peak luminance values were matched following gamma correction. This was achieved by comparing both linear functions and manually adjusting the brightness of one monitor until the optimal superimposition of the two functions was found.

3.9.2 Verifying stimulus timing

When designing temporal experiments it is important to be aware that the presented durations of the stimuli may not always correspond with those that have been requested. This could arise from errors within the experimental code, or problems with the hardware (i.e. computer, monitor, headphones etc…) not faithfully reproducing the durations requested by the code. In order to ensure that the duration and relative temporal relationship of all stimuli was accurate during the experiments outlined in this thesis, all durations

were verified using either a Picoscope 2204A (2200 series) or a Gould 160420 msec/sec oscilloscope prior to the start of any data collection.

An oscilloscope transforms electrical signals into graphical form, typically displaying voltage against time. To measure visual stimuli, a photodiode was positioned directly in front of the monitor so that the face of the diode was approximately 5mm from the centre of the visual stimulus. Presentations of

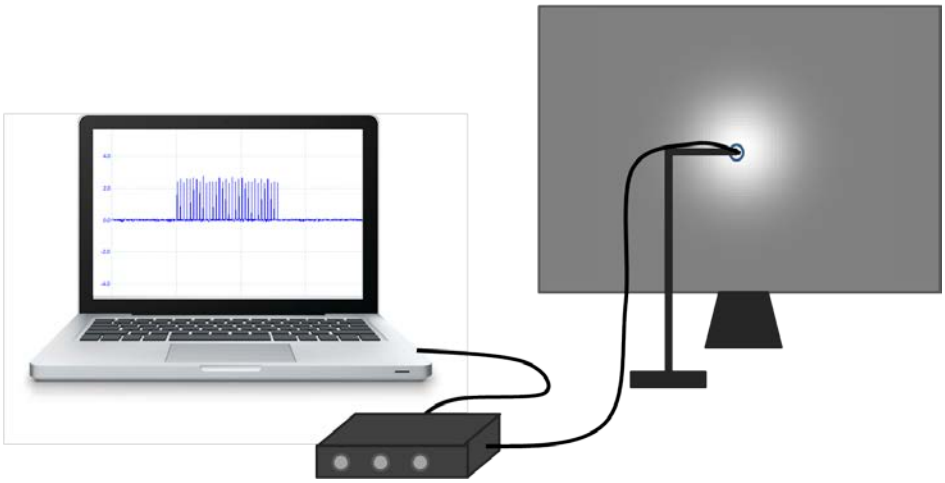

Figure 3.15: Schematic of the verification of a visual stimulus duration using the Picoscope oscilloscope. A photodiode is positioned directly in front of the visual stimulus, and connected to the oscilloscope, where the electrical signal is then displayed on a laptop screen using custom software. The amplitude of this signal increases during stimulus presentation, and the duration can be verified by counting the individual spikes corresponding to the monitor refresh rate (see Figure 3.16 for further detail).

the stimulus were then initiated, with a predetermined, fixed duration. During periods of stimulus presentation the step change in luminance was detected by the photodiode and converted into an electrical signal. This was registered by the oscilloscope as a change in signal strength and displayed graphically on either a laptop screen (for the Picoscope - using custom software) (see Figure 3.15), or on the Gould oscilloscope's screen (see Figure 3.16) as an increase in the amplitude of the signal.

Auditory durations were measured by connecting the oscilloscope directly to the soundcard through the 'aux out' port. Stimuli and inter-stimulus intervals were verified individually for a range of durations spanning those required in each experiment.

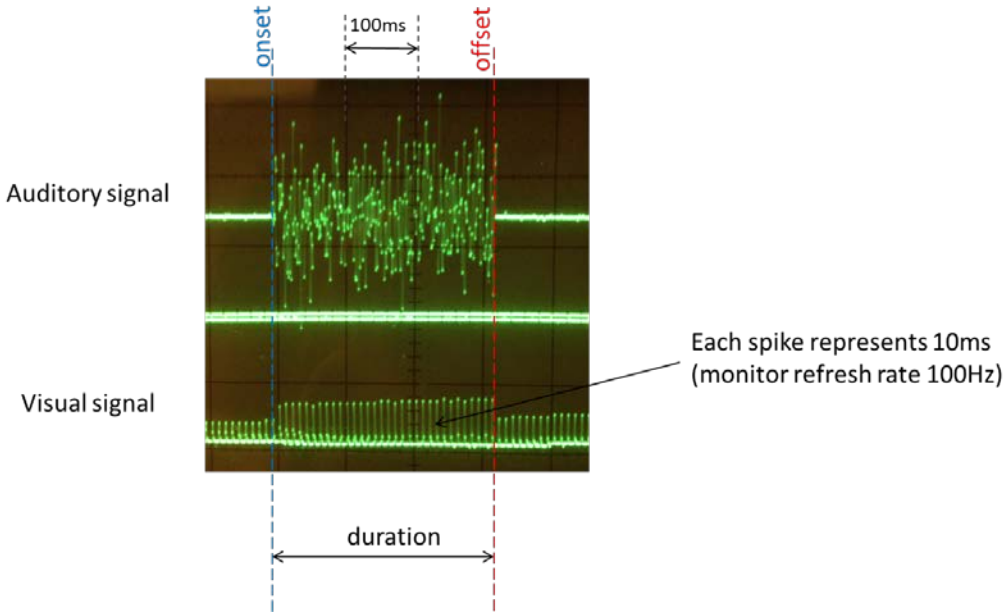

Figure 3.16: An example of timing verification using the Gould oscilloscope when multiple durations are required. In this example both the auditory stimulus and visual stimulus had requested durations of 320ms. The temporal alignment is verified by comparing the onset (blue dashed line) and offset (red dashed line) of each stimulus, to see whether these signals align on the oscilloscope (aided by the black grid on the screen background). The duration is then verified by measuring the spatial interval over which amplitude of the signal remains increased. In the above example, each oscilloscope display square (black grid lines) represents 100ms. Since the refresh rate of this particular monitor was 100Hz, each spike of the visual signal represents one refresh (10ms).

For the experiments outlined in Chapter 7 it was also necessary to compare the temporal alignment of the stimuli. This was achieved by simultaneously comparing the auditory and visual signals using dual channel input. Simultaneous capture of the signals on the oscilloscope display allowed both the duration and temporal alignment to be measured (see Figure 3.16).

Chapter 4: An investigation into the spatial tuning of the duration aftereffect

The work presented in this chapter was published as:

Fulcher, C., McGraw, P.V., Roach, N.W., Whitaker, D. and Heron, J., (2016). Object size determines the spatial spread of visual time. *Proceedings of the Royal Society B* Vol. 283 (No. 1835), doi:10.1098/rspb.2016.1024

Introduction

Although sub-second timing information is critical to the accuracy of most sensory and motor processing, human receptor surfaces do not appear to encode time directly in the way they initiate the analysis of non-temporal features such as pitch, location or temperature. Even at less peripheral locations within the nervous system, evidence remains sparse for any neural structures whose primary function relates to the encoding of temporal information. Despite this, we are capable of formulating temporal estimates that, although noisy (Gibbon et al. 1997; Morgan et al. 2008) are made seemingly without conscious effort and form one of the only perceptual metrics that transcends all sensory modalities (Gorea 2011). This 'supramodal' quality has contributed to the dominance of dedicated, modular mechanisms for time perception such as the pacemaker-accumulator (Creelman 1962; Treisman 1963; Gibbon and Church 1984), oscillator/coincidence-detector (Miall 1989; Matell and Meck 2004) or memory decay (Staddon and Higa 1999) systems. To varying degrees, all of these systems facilitate temporal perception by monitoring ongoing background neural activity around the time of stimulus presentation.

In computational terms, centralised models have the attraction of economy in that they avoid the potentially superfluous proliferation of independent, localised timing mechanisms across primary sensory areas. However, the convergence of sensory inputs onto specialised processing modules necessitates an *a priori* pooling of information across these inputs. It therefore follows that stimulus-specific time perception of any kind presents non-trivial challenges to centralised timing processes. For sub-second duration perception, the possibility of multiple localised timing mechanisms is given credence by reports of sensory-specific distortions of perceived duration. For example, perceived visual (but not auditory) duration is compressed around the time of a saccade (Morrone et al. 2005) or via repeated presentation of identical images (Pariyadath and Eagleman 2007). More generally, estimates of auditory duration are expanded relative to those for visual stimuli, as well as being significantly less variable (Walker and Scott 1981; Wearden et al. 1998; Westheimer 1999; Burr et al. 2009b), inconsistent with a singular central mechanism for the two sensory modalities.

Further examples of sensory-specificity have been revealed by adaptation experiments where exposure to consistent duration information leads to a 'duration aftereffect' (DAE): adaptation to relatively short/long auditory or visual durations induces perceptual expansion/compression of subsequently viewed/heard intermediate duration stimuli. These repulsion-type aftereffects are bidirectional, limited to the adapting stimulus modality and tuned around the adapting duration (Walker et al. 1981; Becker and Rasmussen 2007; Heron et al. 2012; Heron et al. 2013). The neural basis of these effects

remains unclear. One possibility is that they reflect a human analogue of the ‘channel-based’ analysis predicted by neurons with bandwidth-limited duration tuning found in a range of neural structures across several amphibian and mammalian species (for a recent review see Aubie et al. 2012). In the visual domain, the activity of these neurons could form a relatively late-stage, “dedicated”, duration-encoding mechanism (Ivry and Schlerf 2008) that - whilst sensory specific - could operate at a level where basic stimulus features have been pooled to allow selectivity for more complex, object-based analysis (Cox 2014). Alternatively, if visual event duration forms part of a ‘primal sketch’ (Marr 1982), duration-tuned neurons would extract duration information alongside low-level stimulus features, prior to any pooling.

Here this question is addressed by utilising the orderly relationship between spatial selectivity and visual cortical hierarchy (Felleman and Van Essen 1991). Specifically, neurons located in extrastriate visual cortex, which perform more complex forms of visual analysis, often inherit pooled inputs from lower level structures (Kohn and Movshon 2003; Freeman and Simoncelli 2011). This pooling of information over larger spatial regions supports the analysis of more global image properties, produces receptive fields that are necessarily larger than their inputs and exhibit correspondingly coarser spatial selectivity. Conversely, primary sensory (or even pre-cortical) areas are more closely associated with high degrees of spatial selectivity (Allman and Kaas 1974; Gattass et al. 1981; Van Essen et al. 1984; Gattass et al. 1988; Xu et al. 2001; Dumoulin and Wandell 2008; Lee et al. 2010; Cheong et al. 2013). By measuring the spatial tuning of DAEs, the following

series of experiments aim to constrain candidate neural loci for duration processing within the visual processing hierarchy.

Experiment 4.1

4.1.1 Methods

4.1.1.1 Observers

Six observers (three naive) took part in the experiment. All observers gave their informed, written consent to participate, and had normal or corrected to normal vision and hearing at the time of the experiment.

4.1.1.2 Stimuli and Apparatus

All visual stimuli were presented on a gamma-corrected Compaq P1220 CRT monitor with a refresh rate of 100Hz and a resolution of 1280x1024. This was connected to a 2x2.26GHz Quad-core Apple Mac Pro desktop computer running Mac OS 10.6.8. All stimuli were generated using Matlab 7.9.0 (Mathworks, USA) running the Psychtoolbox Extension version 3.0 (Brainard and Pelli, 1997,www.psychtoolbox.org). The physical durations of all auditory and visual stimuli were verified using a dual-channel oscilloscope. The auditory stimulus was a 500Hz tone presented through Sennheiser HD 280 headphones. Visual stimuli were isotropic, luminance defined Gaussian blobs (mean luminance 77cd/m^2) presented against a uniform grey background of 37cd/m^2, whose luminance profile was defined as follows:

$$y = \mu\left(1 + e^{-\left(\frac{x^2}{2\sigma_{stim}{}^2}\right)}\right)$$

Where μ is the mean luminance value of the grey background and σ_{Stim} is the standard deviation of the Gaussian.

In this experiment (see Figures 4.2a-c and 4.3b) σ_{Stim} was set to 1°.

4.1.1.3 Procedure

Observers viewed the visual stimuli binocularly in a quiet, darkened room whilst maintaining fixation on a white 0.07° circular fixation marker presented 5.33° to the left of the centre of the screen. Viewing distance was controlled (via chin rest) to ensure one pixel subtended one arc minute. A block of trials began with an initial adapting phase consisting of 100 serially presented visual stimuli. Within a block the duration of these stimuli was fixed at either 160ms or 640ms. The inter-stimulus interval (ISI) was randomly jittered between 500 -1000ms. The adaptation phase was followed by a further four 'top up' adapting stimuli and a subsequent test phase (see Figure 4.1) consisting of a fixed (320ms) duration auditory reference stimulus followed by a variable duration visual test stimulus. Observers then made a two alternative forced choice (2AFC) duration discrimination judgment as to "which was longer, flash or beep?" Visual test stimuli varied in seven approximately logarithmic steps: 240, 260, 290, 320, 350, 390 and 430ms, which were randomly interleaved within a method of constant stimuli.

Observers responded via key press which triggered the next top-up and test cycle, until all test durations had been presented ten times per block of trials. The adapting stimulus was presented at fixation, 5° or 10° to the right of fixation. Test stimuli were either presented at the adapting location or locations providing 5° or 10° adapt-test spatial intervals (see Figure 4.1). This

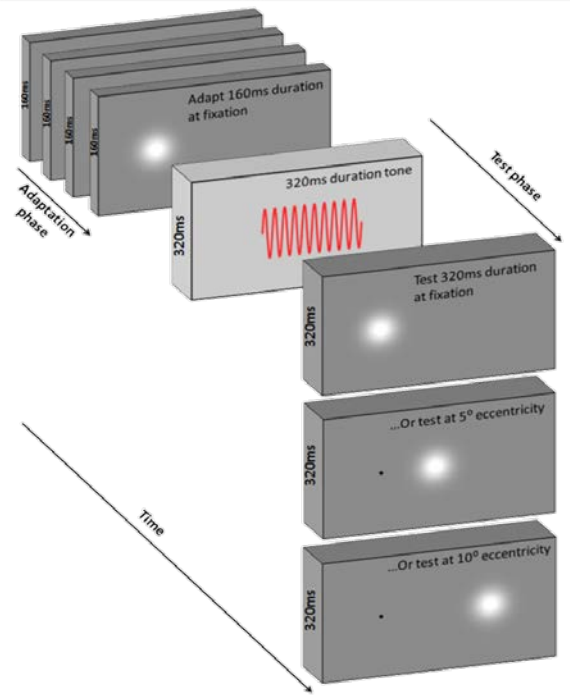

Figure 4.1: A schematic showing the adapt-test paradigm. In the adaptation phase observers view a series of visual stimuli of fixed duration (160ms in this example) at one of three possible adapt locations (fixation in this example). In the following test phase, observers make a duration discrimination judgement between a 320ms auditory reference duration, and a variable visual test duration (320ms in this example). The test stimulus may occur at fixation, at 5° eccentricity or at 10° eccentricity (constant within a block), forming nine possible adapt-test spatial configurations.

provided nine adapt-test spatial configurations (3 adapt locations x 3 test locations), each of which remained constant within a block of trials. Each adapt-test spatial configuration was repeated for both adapting durations giving a total of 18 conditions. Blocks pertaining to each condition were

completed in a random order. Each observer completed three blocks per condition to give 30 repetitions per data point, per observer. In total, data collection lasted approximately 9 hours per observer.

The resulting psychometric functions were fitted with a logistic function of the form:

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

Where PSE represents the point of subjective equality, corresponding to the physical test duration that was perceptually equivalent to the 320ms auditory reference stimulus and θ is an estimate of the observer's duration discrimination threshold (half the difference between the values corresponding to 27% and 73% test longer responses). From these functions, PSE values were extracted for each observer for both the 160ms and 640ms adaptation conditions, across each of the nine adapt-test spatial configurations.

4.1.1.4 Modelling

To aid in making inferences regarding the spatial scale of duration coding mechanisms, a simple filtering model was developed. The neural representation (*Rep*) of each stimulus across retinotopic cortex was simulated by convolving its horizontal contrast envelope with a Gaussian spatial filter:

$$Rep = e^{-\left(\frac{x^2}{2\sigma_{stim}^2}\right)} \otimes e^{-\left(\frac{x^2}{2\sigma_{filt}^2}\right)}$$

Where σ_{stim} and σ_{filt} are the standard deviations of the stimulus and filter respectively, and *x* indicates the spatial distance from the centre of the stimulus/filter (all in degrees of visual angle).

Because both stimulus and filter are Gaussians, *Rep* is itself a Gaussian centred at the location of the stimulus, with a standard deviation σ_{rep} given by:

$$\sigma_{rep} = \sqrt{\sigma_{stim}^2 + \sigma_{filt}^2}$$

The proportional overlap *O* between adapting and test neural representations was calculated by:

$$O = 2\int_{-\infty}^{0} \frac{1}{\sigma_{rep}^2\sqrt{2\pi}} e^{-\left(\frac{(x-\frac{d}{2})^2}{2\sigma_{rep}^2}\right)} dx$$

Where *d* is the centre to centre distance between adapting and test stimuli.

The expected duration aftereffect (DAE) was assumed to be a linear function of this overlap:

$$DAE = kO$$

Where *k* is the peak DAE obtained with identical adapting and test stimuli.

For each stimulus size, the spatial filter model was fitted to the tuning function relating DAE magnitude to separation, finding the values of σ_{rep} and

k that minimised the sum of squared residual errors between expected and measured aftereffect magnitudes.

4.1.2 Results and discussion

Figure 4.2a shows a sample psychometric function for representative observer CF. The proportion of responses where the visual test was perceived as longer than the auditory reference is plotted as a function of visual test duration for the condition where both the adapting stimulus and test stimuli were presented at 10° from fixation (i.e. with no spatial separation). Repeated presentations of the 640ms adapting stimulus (solid black curve, black squares) depresses the number of 'test longer than reference' responses, which reflects a perceived compression in the duration of the test stimulus: a physical test duration of 377ms is judged as perceptually equivalent to a physical auditory reference duration of 320ms. Conversely, the function relating to the 160ms adaptation condition (dashed curve, black circles) is shifted leftwards, reflecting an expansion of the perceived duration of the test stimulus: a physical test stimulus of 315ms now has perceptual equivalence with the reference stimulus. These temporal distortions are consistent with previous reports of bi-directional, repulsive duration aftereffects (Heron et al. 2012; Heron et al. 2013).

The extent of the lateral separation between the two functions provides a measure of DAE magnitude (DAM) and can be expressed as the arithmetic difference between PSE values for the two adapting duration conditions:

$$DAM = PSE_{640} - PSE_{160}$$

Where PSE_{640} is the PSE value obtained from the 640ms adapting duration and PSE_{160} is the PSE value obtained from the 160ms adapting duration. For the observer shown in Figure 4.2a, DAM = 62ms when adapting and test durations are both presented at the same location. Of particular interest in the current study was to establish how DAM varied during manipulation of the adapt-test spatial interval.

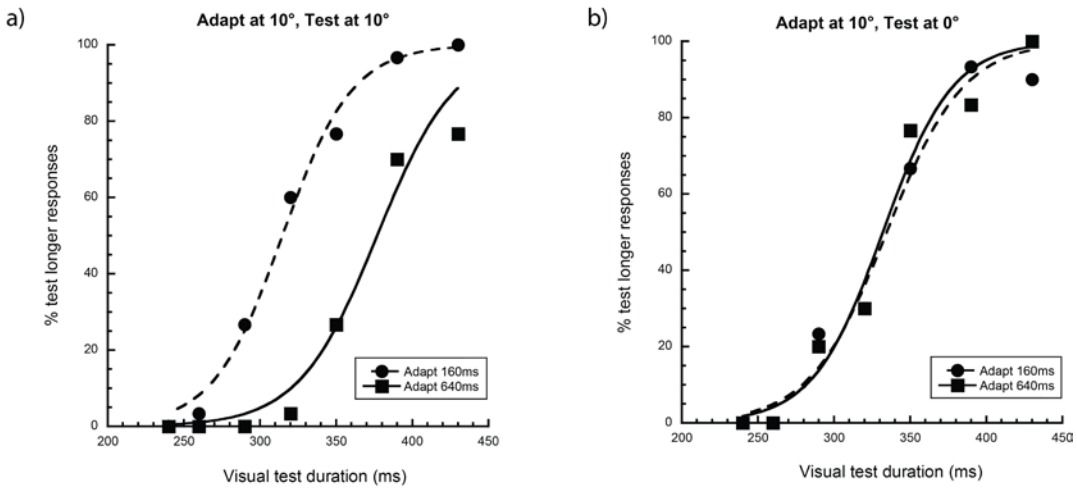

Figure 4.2: a) Psychometric functions for a single representative observer making duration discrimination judgments following duration adaptation. Circles refer to the 160ms adaptation condition and the squares show the 640ms adaptation condition. In this condition, adapting and test duration were presented at 10° temporal to fixation. **b)** Data from the same observer under identical conditions except for the introduction of a 10° spatial interval between adapting and testing locations.

Figure 4.2b shows psychometric functions for the same observer when the adapting and test stimuli were separated by 10° ('Adapt at 10°, test at fixation'). The superimposition of the two functions is in stark contrast to the lateral separation shown in Figure 4.2a. This represents a reduction in the

effectivity of the adapting stimuli: the perceived duration of the test stimulus shows negligible variation across both adapting durations.

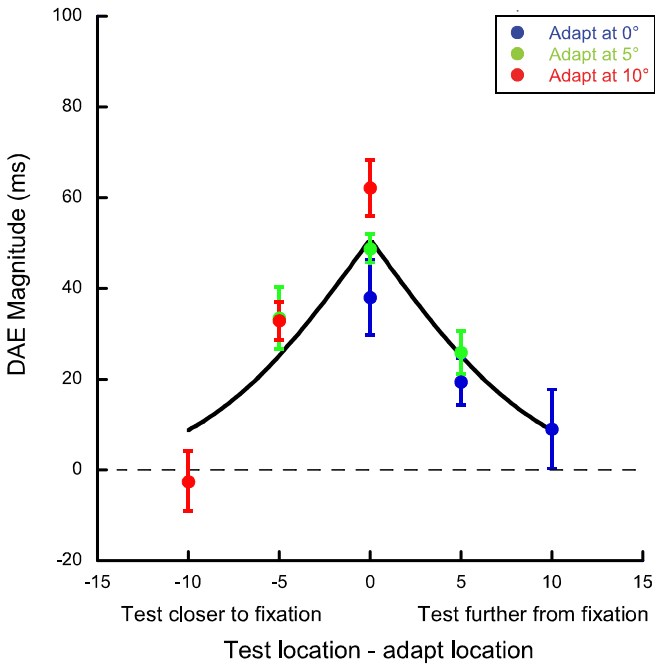

Figure 4.3: A spatial tuning plot showing the variation in duration aftereffect (DAE) magnitude across a range of adapt-test spatial configurations for a single representative observer (see Methods and Figure 4.4 for details). An x-axis value of zero represents conditions where adapt and test duration were presented at the same spatial location. Positive (negative) x-axis values represent conditions in which the test stimulus was presented further from (closer to) fixation than the adapting stimulus. Blue circles represent conditions where the adapting stimuli were presented at fixation, green circles represent conditions where the adapt location was 5° eccentricity and red circles represent conditions where the adapt location was 10° eccentricity. The black curve represents the best fitting neural overlap function to the individual data points, as calculated from the spatial filter model. Error bars represent the SE.

Figure 4.3 shows data from the same observer where DAM is plotted as a function of all nine adapt-test spatial configurations. For all three adapting locations, robust duration aftereffects are generated by presenting adapt and test stimuli at the same spatial location (Figure 4.3 - central data points). As

the adapt-test spatial interval is increased, DAM shows a progressive decrease, indicating a reduction in the perceptual bias induced by adaptation. This pattern of spatial tuning is manifest for all three adapting locations, as demonstrated by the red, green and blue data points forming a single function.

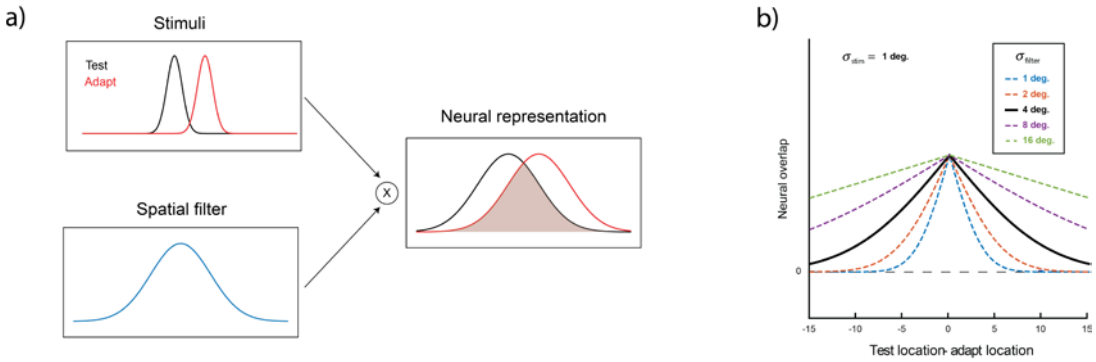

Figure 4.4: a) Schematic of a simple spatial filtering model, showing how a neural representation of the adapt and test stimuli could be generating by convolving the horizontal contrast profiles of both stimuli with a Gaussian filter. This allowed the area of overlap between the neural representations to be calculated. **b)** Showing examples of neural overlap functions obtained from the model using different sized spatial filters.

Spatially tuned duration aftereffects are evidence that – at some level – event timing must be segregated into distinct regions of visual space, a finding that could signal the presence of neurons that are selective for both the duration *and* spatial location of a visual event. But what is the spatial scale of duration coding mechanisms? To address this question quantitatively, a simple spatial filtering model was developed based on the assumption that duration aftereffects occur when (and only when) adapting and test stimuli stimulate overlapping neural populations (see methods for details). As illustrated in Figure 4.4a, the horizontal contrast profiles of the stimuli were first convolved with a Gaussian filter corresponding to neural

blur, and the proportional overlap between the resulting neural representations of the adapt and test stimuli was calculated. The proportion of overlap was then determined for a range of different adapt-test spatial separations. Figure 4.4b shows the resulting spatial tuning functions obtained with a range of neural representation sizes. Application of the model to the individual data shown in Figure 4.3 is shown by the black curve, whose peaked shape is driven by the area of overlap between neural representations of the adapt and test stimuli (see Figure 4.4a). The model revealed a best fitting σ_{rep} of 3.67° for this representative observer. When averaged across all six observers the mean σ_{rep} value was calculated as 4.04°, which is several multiples of σ_{stim} (the spatial spread of the stimulus). In other words, duration adaptation extends into spatial regions well beyond the physical confines of the adapting stimuli themselves. A relatively large aftereffect spread across space could be consistent with late-stage processing subserved by a coarse, fixed-scale of spatial filtering (Maunsell and Newsome 1987). If this scale (σ_{filt}) is larger than the stimulus (σ_{stim} - as depicted in Figure 4.4a) the degree of overlap between adapting and test neural representations (σ_{rep}) would be similar across modest changes in stimulus sizes above and below 1°. This possibility was examined by repeating Experiment 4.1 using smaller (0.5°) and larger (1.5°) Gaussian stimuli.

Experiment 4.2

4.2.1 Methods

4.2.1.1 Observers

The same six observers (three naïve) from Experiment 4.1 took part in Experiment 4.2.

4.2.1.2 Stimuli and apparatus

The stimuli and apparatus used were identical to Experiment 4.1, except that stimulus size was modified by increasing (σ_{stim} = 1.5°) or decreasing (σ_{stim} = 0.5°) the value of σ_{stim}.

4.2.1.3 Procedure

The procedure was identical to Experiment 4.1. Stimulus size was held constant within, but varied between blocks. Each observer completed three blocks per condition, to give 30 repetitions per data point. In total, data collection lasted approximately 18 hours per observer.

4.2.2 Results and discussion

For each observer, psychometric functions were constructed using the method described in Experiment 4.1, and PSE values were extracted for each condition. From these, the duration aftereffect magnitude (DAM) values were calculated, and plotted as a function of adapt-test spatial configuration (as described in Section 4.1.2). Group averaged results for each of the three

size conditions are shown in Figure 4.5a-c, where irrespective of stimulus size, DAM can be seen to decline systematically with adapt-test spatial interval. Black curve fits represent the best fitting neural overlap functions for each stimulus size as calculated from the model (see Section 4.1.1.4).

The rate of DAM decline across increases in adapt-test spatial interval also notably varies with stimulus size. This progressive broadening of spatial

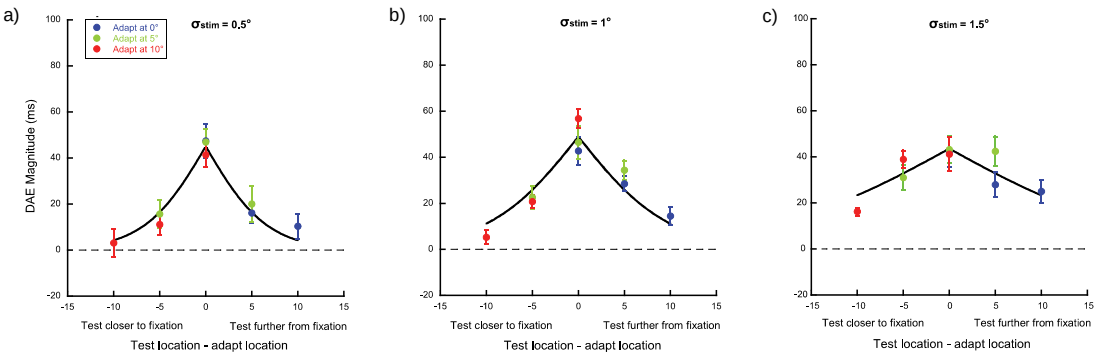

Figure 4.5: (a-c) Mean spatial tuning plots for the three stimulus sizes (σ_{stim}= 0.5°, 1° and 1.5°), showing DAE magnitude as a function of the spatial separation between adapt and test locations. Blue circles represent conditions where adaption occurred at 0°, green circles represent conditions where adaptation occurred at 5° and red circles represent conditions where adaptation occurred at 10°. For each adapt-test spatial configuration, stimulus size was held constant between adapting and test phases. The black curves represent the best fitting neural overlap functions for each stimulus size, as calculated from the spatial filter model. Error bars represent the SEM (n=6).

tuning with increasing stimulus size is summarised in Figure 4.6a, where best fitting σ_{rep} values are plotted as a function of σ_{stim}. For comparison, the dotted lines show a family of model predictions for different levels of neural blur. Clearly, changes in the spatial tuning of the DAE with stimulus size are not consistent with any fixed scale of spatial filtering.

From the best-fitting σ_{rep}values the neural blur of the filter σ_{filt} which would have produced this pattern of results can be calculated from the model. The data predicts filter sizes of 2.76°, 3.91° and 7.86° for our three stimulus sizes of 0.5°, 1° and 1.5°. Rather than a fixed level of coarse spatial filtering, this suggests a 'self-scaled' relationship in which the spatial scale of the filter determining aftereffect tuning forms a multiple of the spatial scale of the stimulus. Simulations based on this principle are shown in Figure 4.6b where the best fitting scaled filter is 5.2*σ_{stim} (Figure 6b – black line).

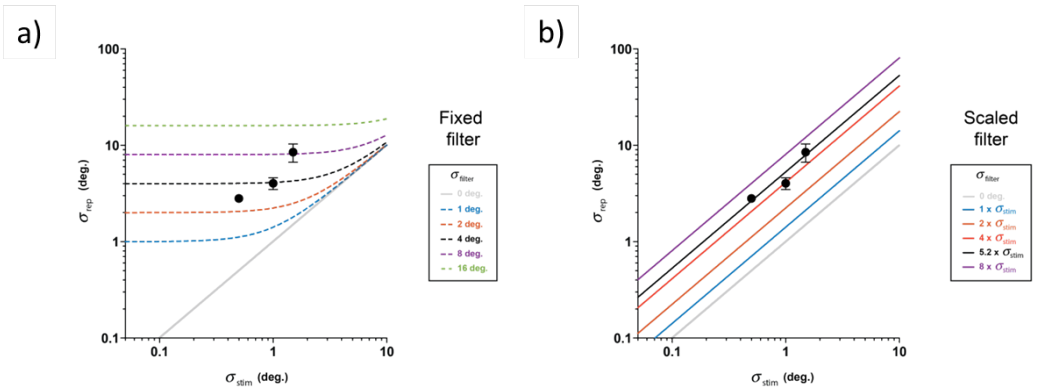

Figure 4.6: Best fitting σ_{rep}values plotted as a function of σ_{stim}(black data points) against **a)** a series of model predictions (dotted lines) based on fixed scale spatial filters of varying size and **b)** a series of model predictions (solid lines) based on a scaled filter that varies proportionally with stimulus size.

To determine how well the overlap model describes the spatial tuning of the aftereffects, it was compared to an alternative model in which the DAM at each of the five spatial separations between adapt and test location shown in Figure 4.5 (see abscissa values) was taken as the average of all of the DAM values corresponding to each location. For example, location "-10" would refer to the condition 'adapt at 10° and test at 0°', and location "5" would refer to the average of the two values from the 'adapt at 0°, test at 5°' and 'adapt

at 5°, test at 10°' conditions. This "saturated" model therefore uses five parameters (the average DAM at each location) to predict the relationship between aftereffect magnitude and adapt-test spatial separation. An F-test was then conducted to compare the overlap model with the saturated model, for each of the three σ_{stim} values. The results showed no significant differences between the overlap and saturated model for σ_{stim}= 0.5° (F(52,49) = 0.99, p=0.49), for σ_{stim}= 1° (F(52,49) = 1.06, p=0.59) and for σ_{stim}= 1.5° (F(52,49) = 0.98, p=0.47). This indicates that both models can adequately explain the spatial tuning data. However, since the overlap model provides a simpler explanation for the spatial tuning of the DAE and is only based on two parameters ("k" and "O", see equation on p156), it is therefore the preferred choice for the present data.

The spatial tuning of the DAE indicates that it must be linked to the receptive field characteristics of underlying visually responsive neurons (e.g. spatial frequency tuning). Sufficiently large changes in adapt-test position would stimulate neural populations with non-overlapping receptive fields, therefore limiting the transfer of the aftereffect across space. A further prediction that arises from receptive field involvement is that the DAE may show selectivity for changes in stimulus size at a fixed spatial location, as neurons with different sized receptive fields will respond optimally to different spatial frequencies (i.e. stimulus sizes). If independent neural populations are stimulated by different sized adapt and test stimuli, it is possible that these neural representations would not interact (overlap) and thus limit DAE transfer. On the other hand, if there is any commonality between the neural populations stimulated by different sized adapt and test stimuli, overlap

between the resulting neural representations may facilitate transfer of the DAE. We tested this hypothesis using stimuli whose spatial extent was varied across adapt and test phases.

Experiment 4.3

4.3.1 Methods

4.3.1.1 Observers

Seven observers (four naïve) took part in Experiment 4.3, five of whom had previously participated in Experiments 4.1 and 4.2. All observers gave their informed, written consent to participate, and had normal or corrected to normal vision and hearing at the time of the experiment.

4.3.1.2 Stimuli and Apparatus

The stimuli and apparatus were identical to Experiment 4.2. The size of the visual stimuli were again σ_{stim} = 0.5° and σ_{stim} = 1.5°.

4.3.1.3 Procedure

The procedure was identical to Experiment 4.1, except that all stimuli were presented at fixation. Observers performed two conditions in which the size of the visual stimuli were constant across adapt and test phases, “adapt 0.5°, test 0.5°” and “adapt 1.5°, test 1.5°” and two conditions in which size varied between adapt and test conditions, “adapt 0.5°, test 1.5°” and “adapt 1.5°,

test 0.5°". For ease of discussion, these conditions will be labelled $A_{0.5}T_{0.5}$ / $A_{1.5}T_{1.5}$ and $A_{0.5}T_{1.5}$ / $A_{1.5}T_{0.5}$ from this point forward. Each observer completed 24 blocks, and data collection lasted approximately four hours per observer.

4.3.2 Results and discussion

For each observer, psychometric functions were constructed using the method described in Experiment 4.1, allowing extraction of PSE values for each condition. From these, the duration aftereffect magnitude (DAM) values were calculated. The DAM values for all four conditions, averaged across

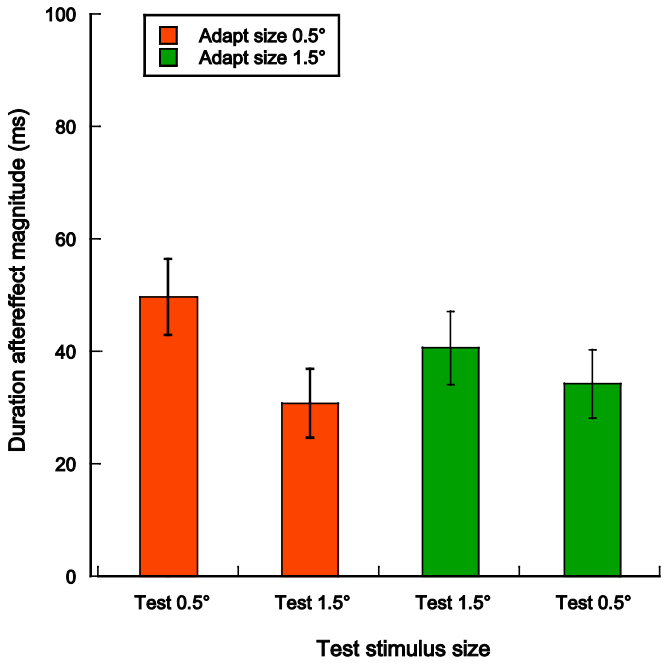

Figure 4.7: Duration aftereffect (DAE) magnitude averaged across observers (n=7), following adaptation to σ_{stim}= 0.5° (orange bars) or σ_{stim}= 1.5° (green bars). Error bars represent the SEM.

observers (n=7), are shown in Figure 4.7. In the two conditions where stimulus size was kept constant across adapt and test phases, the resulting

DAM values were $A_{0.5}T_{0.5}$ = 49.7ms and $A_{1.5}T_{1.5}$ = 40.6ms. These conditions replicate the adapt 0° and test 0° conditions from Experiment 4.2, and show strong agreement in the overall magnitude of the aftereffect (Experiment 4.2 data: 0.5° = 47.4ms and 1.5° = 42.4ms). In conditions where stimulus size was changed between the adapt and test phases, the resulting DAM values were $A_{0.5}T_{1.5}$ = 30.8ms and $A_{1.5}T_{0.5}$ = 34.2ms. A paired samples t-test found no significant differences between any of the experimental conditions, suggesting that the mechanism driving the DAE is not selective for stimulus size at a fixed spatial location. This would suggest that the overlapping neural representations of the adapt and test stimuli that arise at a single spatial location facilitate transfer of the duration signal across changes in stimulus size.

However, whilst not significant, there was a trend towards size selectivity: the strength of the aftereffect showed modest reduction when size was varied between the adapt and test phases. This could indicate that the magnitude of the size change was simply not large enough to demonstrate selectivity at a fixed spatial location. An alternative possibility is that the lack of significance resulted from the broadband nature of the spatial frequency content of the Gaussian blobs. Consequently, each blob may have stimulated a range of receptive field sizes within a single spatial location. Although the smaller 0.5° blob would have stimulated a range of receptive fields tuned to relatively higher spatial frequencies than the 1.5° blob, both may have stimulated common mid-range tuned receptive fields facilitating partial transfer of the DAE between the different blob sizes.

General discussion

The experiments outlined in this chapter sought to investigate the interaction between spatial information, recent sensory history and the perception of duration. Adaptation techniques were used to generate bidirectional, repulsive DAEs, which were tested for their sensitivity to adapt-test changes in spatial location. Whilst the results of Experiment 4.3 investigating size selectivity at a single fixed location were inconclusive, Experiments 4.1 and 4.2 demonstrated that sensitivity to adapt-test changes in spatial location was found to be coarse: the effects of adaptation spread into a region considerably larger than the adapting stimulus itself (Figure 4.3 and 4.5b). Additionally, the size of this region was proportional to the size of the adapting stimulus (Figure 4.5a-c). Our model simulations allowed us to assess our spatial tuning data alongside predictions based on a range of fixed, coarse-scale spatial filters (Figure 4.6a), versus scaled filtering which forms a multiple of stimulus size (Figure 4.6b). Fixed-scale filters were unable to capture the relationship between stimulus size and aftereffect spread. Instead, our data are better described by modelling based on the principle that DAEs are generated by a mechanism with self-scaled filtering properties. The effect of this self-scaling is to spread DAEs across an area that is approximately five-times larger than the adapting stimulus.

Broad spatial tuning has practical implications for how adaptation-induced biases are measured. Because duration adaptation does not transfer between sensory modalities (Heron et al. 2012), our observers judged the perceived duration of a visual test stimulus relative to an auditory reference. An alternative is to use a visual reference that is presented at an unadapted

spatial location. However, our data show that it is critical to sufficiently separate the stimulus (particularly if the stimuli themselves are large), otherwise adaptation will influence both the reference and test stimuli during the 2AFC judgment. This provides a possible explanation for why robust DAEs have not been reported in experiments using large visual test and reference stimuli presented in relatively close spatial proximity (Curran et al. 2016).

The spatial tuning reported here appears to contradict the findings of a very recent study where aftereffects were generated in one hemisphere (e.g. 10° left of fixation) and then tested in the opposite hemisphere (e.g. 10° right of fixation) (Li et al. 2015a). In the Li et al. study, adapting and test stimuli were always presented at 10° either side of fixation. This raises the possibility that inter-hemispheric communication between corresponding areas of cortical eccentricity (e.g. Rochefort et al. 2009) could facilitate transfer of the DAEs around an iso-eccentric annulus centred on fixation. This scenario would produce spatial tuning across the annulus' diameter (as per this study) but not around its circumference (as per the Li et al. study). To investigate this possibility, a control experiment was designed. The stimuli and procedure were identical to Experiment 4.1, except that only a 0.5° sized stimulus was employed and all adapting stimuli were presented 10° to the right of fixation. Test stimuli were either presented in the same location as the adapting stimuli (Adapt R, Test R), or 10° to the left of fixation (Adapt R, Test L) so as to create a 20° adapt-test spatial interval spanning either side of fixation (tested in separate blocks). Three observers (two naïve) took part in the control experiment, each performing three blocks per condition (totalling 30

repetitions per observer, per data point). DAM values were then calculated for each observer (see Section 4.1.2 for details), the results of which are shown in Figure 4.8.

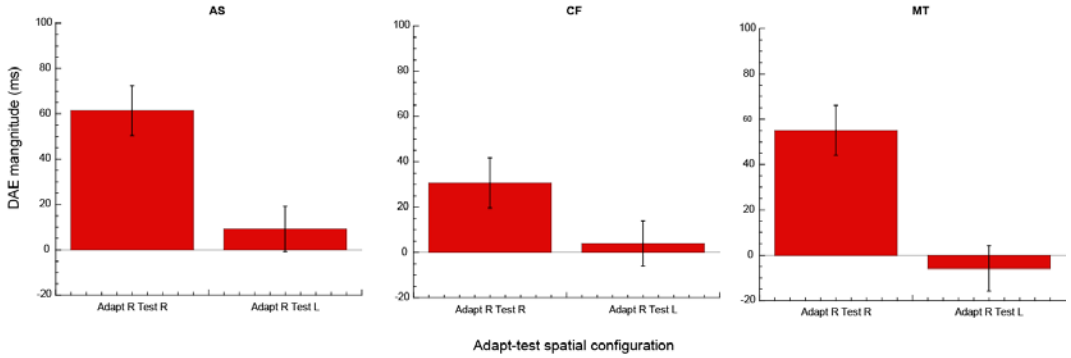

Figure 4.8: Data from a cross-hemifield control experiment where three observers (AS, CF and MT) adapted to a 0.5° sized stimulus with a fixed duration (either 160ms or 640ms) stimuli located 10° to the right of fixation. Test stimuli were then presented at the same location ('Adapt R, Test R') or 10° left of fixation (i.e. in the opposite hemifield) ('Adapt R, Test L'). The height of the bars represents duration aftereffect magnitude (see main text for details). Error bars represent the within-observer SE.

In keeping with earlier experiments, (Figure 4.5a-c) all observers show robust DAEs when adapting and test stimuli were both presented 10° right of fixation. However, no significant aftereffects were generated when adapting stimuli were presented at 10° right of fixation and test stimuli were presented 10° left of fixation, despite matching eccentricity across hemispheres. This finding is consistent with a spatial filtering account of our 'within-hemisphere' data (Figure 4.5a), which predicts a negligible aftereffect magnitude for the 0.5° sized stimulus across a 20° adapt-test spatial interval. It is possible that this discrepancy between the present data and that of Li et al. could have arisen due to large differences in the number of observations contributing to each data point between the two studies. For example, each data point in the

Li et al. study resulted from 80 observations, compared to the 180 observations in the present study. Given that Li et al. did find a trend towards spatial selectivity in their first experiment (albeit non-significant), it is possible that with greater statistical power their conclusions may have shown greater agreement with those presented here.

At the opposite extremes to position-invariant accounts of temporal processing, aftereffects are generated when observers view continuous periods of temporally dynamic (flickering or drifting) visual patterns. Subsequently viewed test stimuli typically undergo perceptual *compression,* (but see Ortega et al. 2012) within the same region of the visual field (Johnston et al. 2006; Burr et al. 2007). These aftereffects show very narrow (~1°) spatial tuning (Ayhan et al. 2009) and no interocular transfer (Bruno et al. 2010), leading some to propose an adaptation locus within the magnocellular layers of the LGN (Ayhan et al. 2011), (but see Burr et al. 2011). Similarly, repetition suppression paradigms show that the presentation of two or more identical visual stimuli in close temporal proximity leads to the underestimation of the second stimulus' duration (Pariyadath and Eagleman 2008). This effect is exaggerated when the two stimuli share the same orientation and are presented within ~2° of each other. Again, these effects have been attributed to mechanisms driven by early striate visual neurons (Zhou et al. 2014).

This group of duration phenomena appear to share some common features: unidirectional (mostly compressive) perceptual distortion which is tightly tuned to low-level stimulus characteristics such as spatial location. These features contrast sharply with the DAEs reported here which could suggest

that the two types of aftereffect (unidirectional, narrowly tuned versus bidirectional, broadly tuned) might be signatures of distinct temporal processing mechanisms.

However, recent advances in our understanding of visual spatial adaptation offer an alternative interpretation. Adaptation to stimulus features such as contrast (Baccus and Meister 2002; Solomon et al. 2004), temporal frequency (Dhruv et al. 2011), motion (Priebe et al. 2002; Kohn and Movshon 2003; Lee and Lee 2012; Larsson and Harrison 2015) and orientation (Larsson et al. 2006; Montaser-Kouhsari et al. 2007; Larsson and Harrison 2015) modulates neural activity across a wide range of areas from retina, striate, extrastriate and association cortex. Simultaneously recording activity from different stages of the neural hierarchy reveals an adaptation cascade where the neural activity at any given site is a product of adaptation intrinsic to neurons at that site *and* adaptation inherited from earlier visual areas (Dhruv and Carandini 2014; Larsson and Harrison 2015). In some cases (Dhruv and Carandini 2014; Patterson et al. 2014), the 'downstream' recipients of 'upstream' adaptation are unable to distinguish between adapted and non-adapted inputs, leading to a cumulative superimposition of distinct adaptation effects (Seriès et al. 2009; Stocker and Simoncelli 2009).

Could adaptation effects from different levels of neural processing also occur for temporal information? Because receptive field size increases systematically throughout pre-cortical, striate and extrastriate visual areas (Allman and Kaas 1974; Gattass et al. 1981; Van Essen et al. 1984; Gattass et al. 1988; Xu et al. 2001; Lee et al. 2010; Cheong et al. 2013) the broad spatial tuning reported here dictates that bidirectional, repulsive DAEs must

originate at a cortical location beyond that responsible for the narrowly tuned, unidirectional effects discussed above. Whatever the relationship between these two aftereffects, simple inheritance of earlier adaptation would predict that repulsive DAEs should display similarly narrow spatial tuning (Kohn and Movshon 2003; Xu et al. 2008). Instead, the tuning profiles shown here suggest repulsive DAEs are generated by a subsequent phase of adaptation that is embodied with the spatial selectivity of neurons whose larger receptive field size reflects their downstream location (Priebe and Lisberger 2002; Larsson et al. 2006; Larsson and Harrison 2015). In this context, the output duration signal from early mechanisms (Johnston et al. 2006; Pariyadath and Eagleman 2008; Zhou et al. 2014) would feed forward to form the (compressed) input signal for a downstream mechanism responsible for the repulsion-type aftereffects reported here. Hybrid paradigms would permit quantification of each mechanism's relative contribution to an observer's duration estimate by adapting to a series of stimuli with a consistent duration (as per the current study) *and* temporal frequency (Johnston et al. 2006; Burr et al. 2007). For example, adapting to relatively short duration bursts of 20Hz flicker predicts that moderate duration flickering test stimuli will undergo perceptual expansion via repulsion-type aftereffects but compression via temporal frequency-based mechanisms.

As argued elsewhere (Heron et al. 2012), channel-based duration encoding by neurons with bandwidth-limited sensitivity to a range of durations (Ivry 1996) is consistent with repulsion-type aftereffects. In the visual domain, a relevant example is the duration tuning seen across the millisecond range in 'off response' neurons within areas 17 and 18 of cat visual cortex (Duysens

et al. 1996). Within these regions (and their primate homologues V1 and V2), individual neurons show tuning for a raft of stimulus features such as orientation, spatial frequency, contrast and motion (Kohn 2007; Webster 2011). Neurons with bandpass duration selectivity have also been documented in the auditory systems of a wide range of species including cat auditory cortex (He et al. 1997), and the auditory midbrain nuclei of bats (Casseday et al. 1994), guinea pigs (Wang et al. 2006), rats (Perez-Gonzalez et al. 2006) and mice (Brand et al. 2000). In addition to stimulus duration, these same neurons invariably show selectivity for auditory pitch (Aubie et al. 2012) and, in some cases, spatial location (Macías et al. 2013). This cross-species and cross-modality generality points towards duration being a generic feature, to which a wide variety of neurons can show tuning.

Which neurons might be responsible for mediating channel-based processing of duration in humans? Recent neuroimaging (Hayashi et al. 2015), transcranial magnetic stimulation (Wiener et al. 2012) and single-unit recording (Leon and Shadlen 2003; Jazayeri and Shadlen 2015) evidence suggests a duration processing role for sub-regions within inferior parietal lobule. However, visually responsive parietal areas have large, often bilateral receptive fields (Motter and Mountcastle 1981), the vast majority of which are at least 5° in diameter (Blatt et al. 1990; Ben Hamed et al. 2001; Quraishi et al. 2007). It therefore seems likely that the adaptation-induced perceptual distortions described here and elsewhere (Johnston et al. 2006; Pariyadath and Eagleman 2008; Ortega et al. 2012; Zhou et al. 2014) reflect intrinsic adaptation in upstream visual areas which undergo subsequent duration encoding in extrastriate/association areas such as lateral intraparietal cortex

(LIP) and supramarginal gyrus (SMG). Motor, premotor and supplementary motor cortices are also reported to show duration-dependent patterns of neural activity (Lebedev et al. 2008; Mita et al. 2009; Merchant et al. 2013b) but again, how intrinsic duration adaptation within these areas could facilitate even broadly-tuned spatial specificity (or indeed perceptual distortions in the absence of any motor action) remains unclear.

When considering the underlying mechanism and its neural underpinnings it is important to acknowledge the relationship between stimulus size and spatial tuning (Figure 4.5). This size dependency is incompatible with the uniformly broad tuning predicted by a large fixed-scale spatial filter that encodes duration across a range of stimulus sizes (see horizontal sections of dashed lines in Figure 4.6a). Is there any evidence for a visual processing stage which not only summates low-level information across a moderate spatial extent, but also whose scale is fundamentally linked to the scale of the low-level information it receives? A prime example of exactly this relationship is provided by the interdependency between mechanisms encoding spatial variations in luminance (first-order) and those encoding variations in texture/contrast (second-order). It is widely accepted that the rectified output of small, linear first-order filters form the input to subsequent, larger second-order filters (for a recent review see Graham 2011). To extract contrast/texture modulations each second-order filter performs ‘spatial pooling’ by combining the outputs of several neighbouring first-order filters (Chubb and Sperling 1988; Cavanagh and Mather 1989). As a result, second-order perceptual phenomena are more spatially diffuse than their first-order counterparts (Ellemberg et al. 2004; Sukumar and Waugh 2007;

Hutchinson and Ledgeway 2010). This is thought to be driven by the fact that neurons tuned to second-order attributes have larger receptive field sizes than those tuned to first-order attributes. Typically, the former are observed in cat area 18 and primate area V2 (Foster et al. 1985; Levitt et al. 1994; Mareschal and Baker 1998; Smith 2007; Li and Baker 2012), whereas the latter are located in striate area 17/V1 (Baker and Mareschal 2001).

Critically, second-order pooling of first-order inputs creates spatial scale-dependency between the two stages: second-order filter size forms a multiple of its first-order input (Bergen 1991). Psychophysical estimates place this multiple at between 3-4x (Sukumar and Waugh 2007), 8-16x (Sagi 1990; Sutter et al. 1995) or 40-50x (Kingdom and Keeble 1996), dependant on the stimulus and task (Westrick and Landy 2013). Single-unit recordings have demonstrated that this relationship is underpinned by neurons whose spatial frequency tuning for contrast or texture-defined information is between 5-30x lower than for luminance-defined information (Zhou and Baker 1996; Mareschal and Baker 1999; Li et al. 2014).

If DAEs are indeed a product of duration-tuning within neurons also selective for second-order image statistics then two clear predictions follow: (1) aftereffects should propagate into a region larger than that predicted by first-order filtering (i.e. the borders of the stimulus itself), and (2) the size of this region will be a fixed multiple of adapting stimulus' size, reflecting the proportionality between first- and second-order size tuning. The data and model simulations reported here show precisely this effect.

In summary, the results presented in this chapter are suggestive of a mid-level form of duration encoding by visual neurons that are selective for a stimulus' spatial characteristics and its duration. Although such a mechanism has the apparent disadvantage of relatively coarse spatial resolution, it would in fact provide duration estimates that avoid some of the ambiguities associated with the earliest stages of visual processing. For example, using first-order luminance alone during object identification can yield spurious results that are corrupted by shadows and shading gradients (Baker and Mareschal 2001). By pooling across a larger spatial area it is possible to disambiguate object-background borders via second-order changes in texture or contrast. Relatedly, changes in viewing distance alter absolute first and second-order spatial scale but, for any given object, the size ratio between these cues does not change. This 'scale invariance' (Johnston 1987; Kingdom and Keeble 1999; Vakrou et al. 2007) ensures that our ability to detect and discriminate between stimulus features defined by second-order cues remains constant across distances in a way that does not hold for first-order cues (Howell and Hess 1978). Therefore, if duration selectivity *were* a feature of neurons tasked with more complex image attributes it would afford perceived duration a degree of object specificity that could be robust enough to cope with occasions where lower-level information is less reliable.

Chapter 5: The binocularity of visual time

5.1 Introduction

Chapter 4 of this thesis provides evidence that visual duration aftereffects (DAEs) show broad, scale-dependent spatial tuning – a pattern which could reflect a mid-level mechanism that pools spatial information from smaller, presumably earlier inputs.

In addition to scale-dependent spatial selectivity, Chapter 4's findings also coincide with earlier reports of selectivity for the adapting sensory modality (e.g. Walker et al. 1981; Becker and Rasmussen 2007; Heron et al. 2012). The DAE must therefore be mediated by neurons responsive to visual information alone, whose locus must lie prior to the site of audio-visual integration. In cortical terms, this type of integration was traditionally considered to be a function of multisensory neurons found in temporal and parietal extrastriate areas (Calvert 2001). However, anatomical studies now suggest direct cortico-cortical integration pathways between primary visual and auditory areas (Falchier et al. 2002; Rockland and Ojima 2003). Furthermore, abundant neurophysiological support for subcortical multisensory neurons suggests that audio-visual integration could begin at even more peripheral sections of the afferent sensory pathways (Driver and Noesselt 2008; Stein and Stanford 2008; Seilheimer et al. 2014). Thus, the modality-specificity of duration aftereffects does not *a priori* allow differentiation between neuronal selectivity for duration at early, versus late stages of the processing hierarchy.

In the visual system, a defining feature of this hierarchy's lower reaches is the variation in the proportion of its binocular neurons. Signals arising from the nasal retina of one eye join those from the temporal retina of the fellow eye at the lateral geniculate nucleus (LGN). Despite their close proximity within this structure, monocular information originating from each eye remains segregated within its monocular layers. Upon reaching the striate cortex, projections from the LGN to layer IV a&c of V1 maintain their monocularity within ocular dominance columns (Hubel and Wiesel 1962; Hubel and Wiesel 1972; Blasdel and Fitzpatrick 1984; Menon et al. 1997; Adams et al. 2007). However, neurons in neighbouring layers (e.g. layers 2, 3, 5 and 6) largely respond to signals originating from either eye (Hubel and Wiesel 1968). V1 therefore forms a composite of monocular and binocular neurons. Conversely, extrastriate areas are comprised of a markedly higher proportion of binocularly driven neurons, to the extent that most are thought to be exclusively binocular (Zeki 1978; Maunsell and Vanessen 1983; Felleman and Vanessen 1987; Janssen et al. 2000; Uka et al. 2000; Kaskan et al. 2009).

In addition to being binocularly driven, neurons as early as V1 can also demonstrate selectivity for retinal disparity (Barlow et al. 1967). Disparity arises due to the lateral displacement of each eye within the head, resulting in non-identical images falling on each retina. If the object is placed in front or behind the plane of fixation, corresponding features from the same physical object will stimulate non-corresponding retinal locations (see Figure 5.1). The difference between these locations constitutes "retinal disparity". Since the magnitude of disparity also differs with the physical distance of the

object, neurons tuned to retinal disparity can encode information about depth from the visual scene.

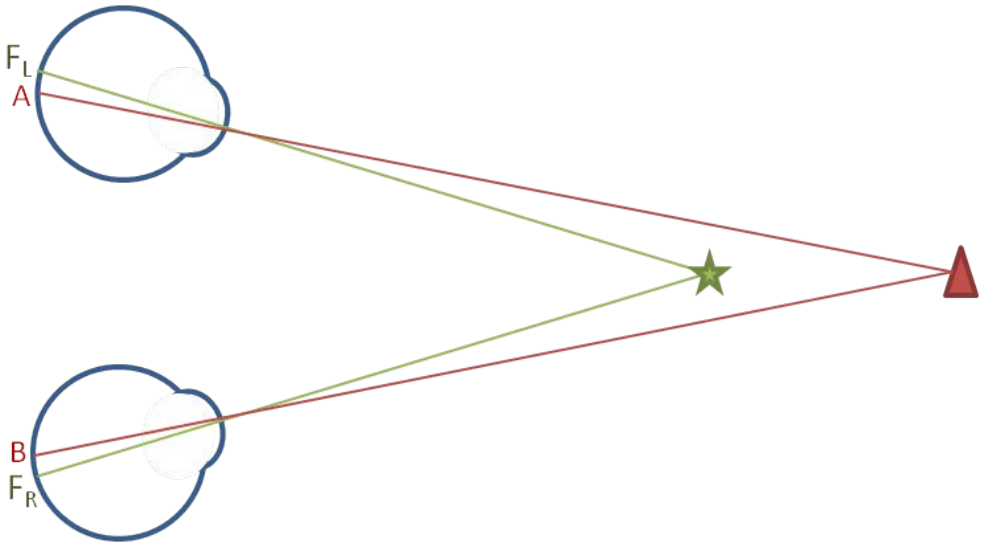

Figure 5.1: Schematic showing an example of retinal disparity. If the observer is fixating on the star, this image will fall on the fovea (F) of the right (R) and left (L) eye. As the triangle is situated behind the star (i.e. behind the plane of fixation), the image of the triangle falls on non-corresponding retinal locations in each eye (points A and B).

Studying the binocularity of visual aftereffects can be a useful tool for differentiating mechanisms with neural loci before or after monocular inputs converge at binocular neurons. Traditionally, the binocularity of aftereffects from the spatial vision literature has been investigated by studying their ability to exhibit inter-ocular transfer (IOT) (e.g. Gibson 1933). Typically an IOT paradigm involves the presentation of an adapting stimulus to one eye, after which the test stimulus is presented to the fellow eye. The magnitude of any resulting aftereffect is then measured. Successful IOT is considered indicative that the neurons driving the aftereffect must receive binocular input, so as to facilitate transfer of the adapted signal. Conversely, absence of IOT would suggest that the aftereffect is driven by monocular neurons,

such that perception of the test stimulus is unaffected by the adapted signal from the fellow eye.

In spatial vision, IOT has been demonstrated for the tilt aftereffect (Campbell and Maffei 1971; Ware and Mitchell 1974), the motion aftereffect (Favreau 1976; Lehmkuhle and Fox 1976; Nishida et al. 1994), shape aftereffects (Gheorghiu et al. 2009), the size aftereffect (Blakemore and Sutton 1969; Murch 1972) and the contrast threshold elevation aftereffect (Blakemore and Campbell 1969; Ware and Mitchell 1974). In some cases, the extent of the IOT has been shown to correlate with an individual observer's stereoacuity (Mitchell and Ware 1974), suggesting that these aftereffects are driven primarily by binocular neurons with a cortical locus. Conversely, the colour motion aftereffect (Mayhew and Anstis 1972) and McCollough illusion (McCollough 1965; Allan et al. 1991) are abolished by presenting the adapt and test stimuli to different eyes leading to the suggestion that these phenomena could originate at a pre-cortical site, or be driven by primarily monocular neurons in the striate cortex.

Thus, a neural locus for DAEs in the striate cortex would see partial or complete inter-ocular transfer (IOT) when adapting and test durations are presented to different eyes. By the same logic, an extra striate locus is more likely to be consistent with complete transfer. Conversely, a pre-cortical mechanism would predict the opposite: that the aftereffect would be limited to test stimuli presented to the same eye as the adapter.

To date, investigations of interocular transfer of aftereffects in the temporal domain have been limited to studies utilising continuous viewing of dynamic (either flickering or drifting) visual adapting stimuli. When subsequently viewed stimuli are similarly dynamic, their perceived duration is compressed within a spatially localised region (Johnston et al, 2006). It has variously been reported that these 'compressive' effects are abolished (Bruno et al, 2010) or persevered (Burr et al. 2007) when adapting and testing stimuli are presented to opposite eyes. However, given that bidirectional repulsive DAEs require adaptation to a sequence of fixed durations (as opposed to continuously presented dynamic stimuli), it is likely that they form a signature of a different temporal mechanism, selective for duration itself rather than temporal frequency. Whether these bidirectional, repulsive DAEs would demonstrate IOT is presently unknown.

Another – as yet unexplored – strategy for investigating the binocularity of DAEs would be to utilise a visual stimulus defined by depth. A popular stimulus in the study of human depth perception is one constructed from random dot stereograms (Julesz 1971). These consist of two images filled with random dots, which are identical except for a spatial region (often near the centre), where the dots in one image are spatially displaced relative to the corresponding dots in the opposite image. When cross-fused or viewed dichoptically, the region containing the right/left relative spatial displacement provides a retinal disparity which stimulates depth-selective neurons. The output signal of these neurons imposes a difference in the perceived depth planes of the image region(s) containing the disparity, relative to those regions with no disparity (i.e. the background – see Figure 5.2).

Dynamic luminance noise provides a uniform background to either eye, consisting only of dynamic first-order information. Any stimulus defined by retinal disparity across left and right images therefore provides no monocular cues to the stimulus' presence. In other words, when retinal disparity is introduced to such a background, it effectively isolates mechanisms that are both binocular and selective for disparity-defined depth information. This is

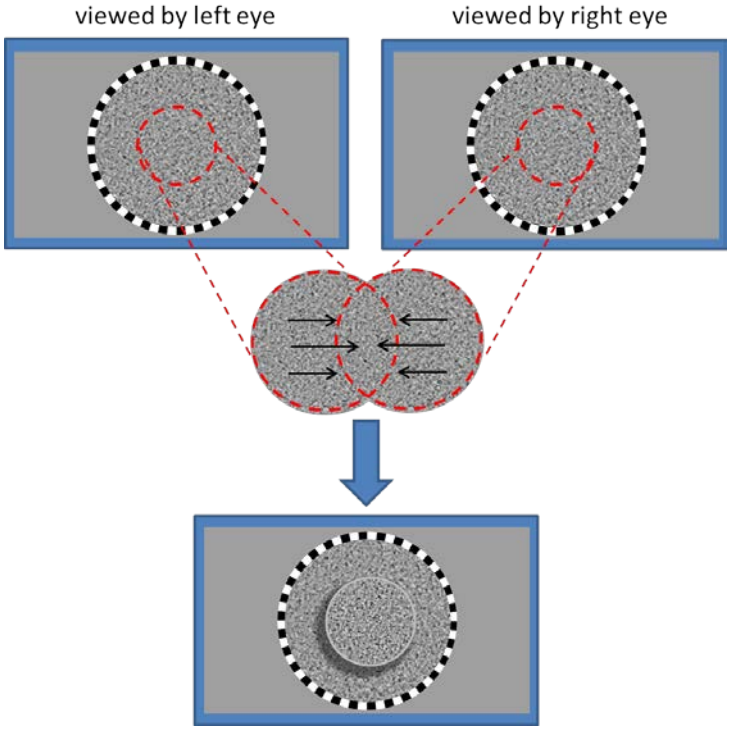

Figure 5.2: Schematic showing the generation of a disparity defined stimulus. The dynamic luminance noise within the red 'target zone' (top row) is laterally shifted (middle row) to induce a crossed retinal disparity. When viewed binocularly, this disparity caused the target region to appear in a raised depth plane relative to the background. The retinal disparity was removed after a period matching the required adapting or testing stimulus duration.

not always the case for other stimuli (e.g. a drifting Gabor) which can be designed to give binocular retinal disparity, but are also visible to monocular mechanisms.

If DAEs can be generated following adaptation to disparity-defined duration signals this would confirm a cortical locus, as depth-selective neurons reside no earlier than the primary visual cortex (Barlow et al. 1967; Xue et al. 1987). However, it should be noted that the absence of a DAE would not necessarily indicate that the underlying mechanism resides at a pre-cortical locus, as the neurons could be binocularly driven but not tuned for retinal disparity (Hubel and Wiesel 1970; Poggio and Fischer 1977).

As such, the experiments described in this chapter sought to test the contribution of binocular neurons to the duration aftereffect via two complimentary methods: 1) measuring the aftereffect's degree of IOT ("IOT task"), and 2) using a duration signal that was only visible to disparity selective neurons ("disparity task").

5.2 Methods

5.2.1 Observers

Ten observers (seven naïve) took part in the IOT task and seven observers (three naïve) took part in the disparity task. All observers gave their informed, written consent to participate, and had normal or corrected to normal vision, stereoacuity and hearing at the time of the experiment.

5.2.2 Stimuli and apparatus

Visual stimuli were presented on two gamma-corrected Eizo FG2421 LCD monitors with a refresh rate of 120Hz and a resolution of 1920x1080. For the IOT task these were connected to a 2x2.8GHz Quad-core Apple Mac Pro

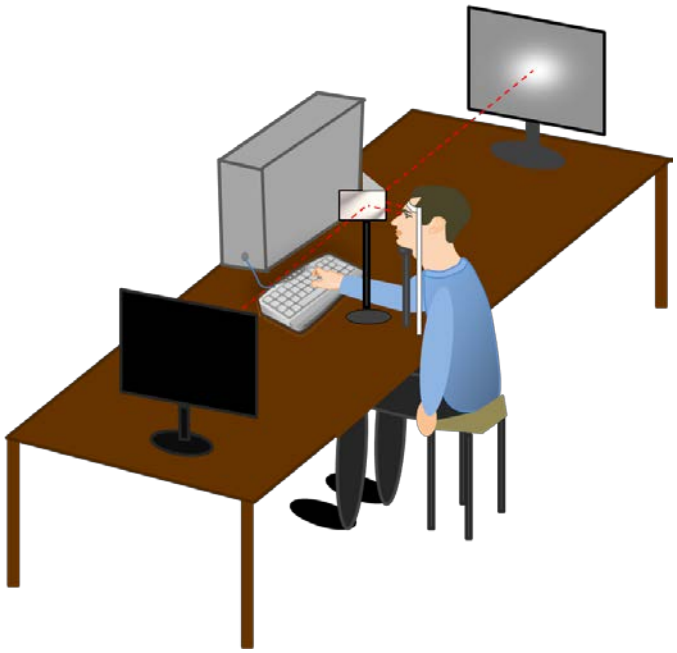

Figure 5.3: Schematic showing the dichoptic arrangement used in both the IOT task and disparity task. Observers viewed the left and right monitors through a two mirror stereoscope. In the above example, the Gaussian blob used in the IOT task (see text for details) is shown on the right monitor.

desktop computer running Mac OS 10.5.8, and for the disparity task they were connected to a 3GHz E5-1660v3 8-Core HP Z440 desktop computer running Windows 8.1 Pro. All stimuli were generated using Matlab (version 7.7.0.471 or 8.4.0, Mathworks, USA) running the Psychtoolbox Extension (version 3.0.8, Brainard 1997: www.psychtoolbox.org).

The physical durations and simultaneity of all auditory and visual stimuli were verified using a dual-channel oscilloscope. Visual stimuli were viewed dichoptically through a two mirror stereoscope which allowed the left and right monitors to be monocularly viewed by left and right eyes respectively (see Figure 5.3). The viewing distance of 2.3 metres was maintained using a chin and forehead rest to ensure that one pixel subtended 0.4 arc minutes.

Auditory stimuli were 500Hz tones presented through Sennheiser HD 280 headphones. Details of the specific visual stimuli pertaining to each task will be described below.

5.2.2.1 IOT task

Visual stimuli were isotropic, luminance defined Gaussian blobs (mean luminance 77cd/m^2) presented at fixation against a uniform grey background of 37cd/m^2, whose luminance (L) profile was defined as follows:

$$L = L_{mean}\,(1 + e^{-\left(\frac{x^2}{2\sigma_{stim}{}^2}\right)})$$

Where L_{mean} is the mean luminance of the Gaussian and σ_{Stim} is the standard deviation.

Throughout the IOT experiment the size of the stimulus, defined by σ_{Stim}, was set to 1°.

5.2.2.2 Disparity task

Visual stimuli were constructed from dynamic random dot stereograms viewed through the dual-mirror stereoscope. They consisted of two dichoptically presented left and right eye circular regions within which each pixel was randomly assigned a binary luminance value (either black or white), which updated at 24Hz to create dynamic noise. Left and right images were identical except for a central 'target zone', which consisted of a smaller circular region of 500 pixels diameter (subtending 3.33° of visual angle when

viewed at 2.3m), in which the noise patch was shifted laterally (creating crossed or ‘positive’ disparity). When viewed binocularly, the central ‘target zone’ appeared at a raised depth plane relative to the surround, corresponding to a disparity of 360 seconds of arc (see Figure 5.2).

The region containing the dynamic noise was bordered by a static checkered annulus which was presented to both eyes, and thus aided binocular fusion. Beyond this border, the luminance of the background was set to mid-grey, which was equal to the average luminance of the dynamic noise.

5.2.3 Procedure

Observers were seated in a quiet, darkened room, and positioned equidistant between the two monitors. A forehead and chin rest were used to maintain head position such that the right and left eyes were centred in front of the stereoscope’s right and left mirrors.

Each block of trials began with observers viewing a pair of vertical and horizontal nonius lines, separately presented to each eye. For example, the right eye might view nonius lines orientated at 3 o’clock and 6 o’clock while the left viewed their counterparts at 12 o’clock and 9 o’clock (see Figure 5.4). A surrounding checkered annulus was presented to both eyes to promote binocular fusion. The tip and tilt of the mirrors was then adjusted independently until the four lines appear horizontally and vertically aligned. This calibration procedure allowed the mirror system to compensate for any observer specific oculomotor anomalies and therefore allow stable fusion throughout the experiment.

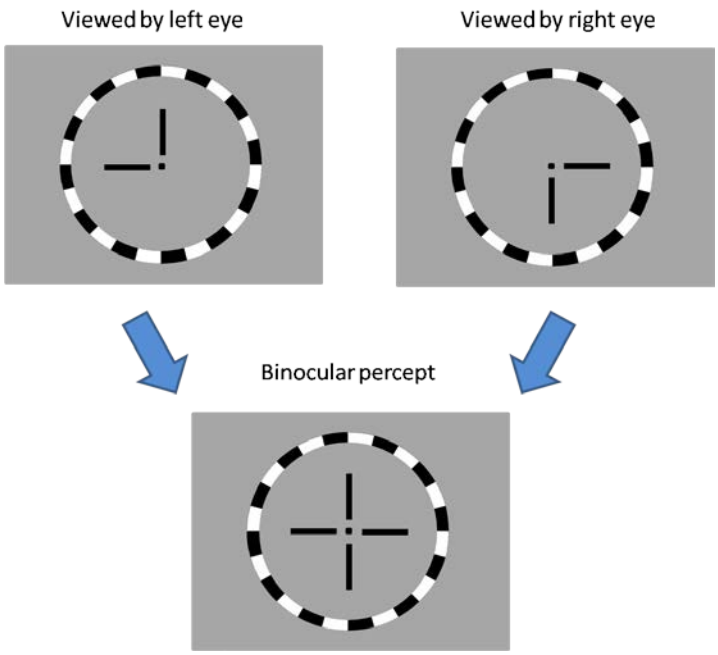

Figure 5.4: Schematic showing the nonius lines viewed by the observer at the start of the experiment. A pair of vertical and a pair of horizontal lines are presented such that each eye sees a different image (i.e. one half of each pair). When viewed binocularly via a dual-mirror stereoscope the observer aims to align the nonius lines by independently adjusting the mirrors, until the fused binocular percept matches the alignment shown in the bottom panel.

5.2.3.1 IOT task

Observer's maintained fixation on a white 0.07° circular fixation marker presented to both eyes at the monitor's centre. They then adapted to relatively long (666ms) or short (166ms) sequences of fixed duration visual stimuli before making 2AFC duration discrimination judgments about the relative durations of a 333ms auditory stimulus and a visual test stimulus which varied in seven approximately logarithmic steps around 333ms. The adapt-test experimental design used to gather this data was identical to Experiment 4.1 (see Chapter 4, p.152), with the exception that visual adapting stimuli were presented to one eye only. In the "same" condition, the adapting stimulus was presented to the same eye as the visual test stimulus,

and in the “different” condition the observer adapted with the opposite eye (see Figure 5.5). The choice of test eye was randomly assigned to each observer at the start of the experiment so that half of the observers used their right eye and half used their left eye.

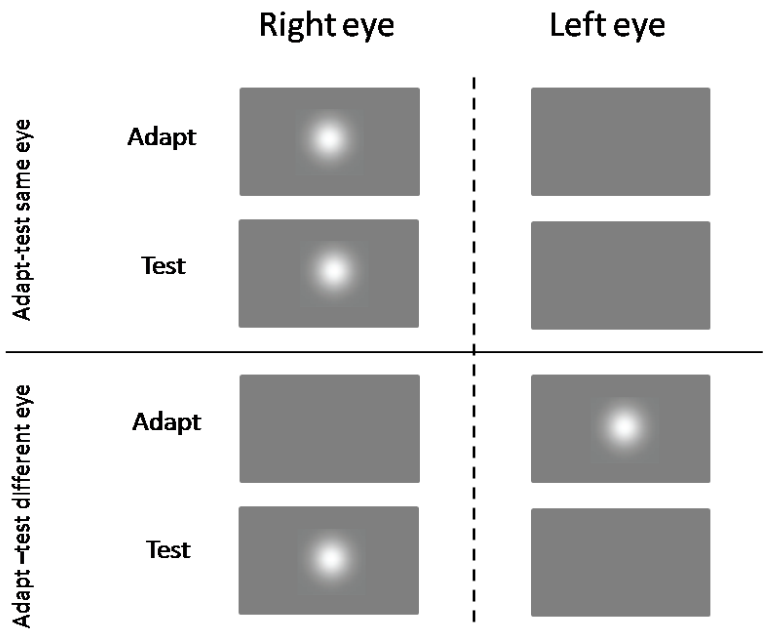

Figure 5.5: Schematic showing the ‘same’ and ‘different’ conditions used in the IOT task. In this example the observer was assigned the right eye as the test eye.

Three blocks were performed for each of the two adapting durations (totalling 30 observations per observer, per data point), and also for each of the two adapting conditions. The total experiment therefore comprised of twelve experimental blocks, totalling approximately 2 hours per observer.

Half of the observers also performed additional conditions to ensure that the choice of test eye did not affect the magnitude of the duration aftereffect. These five observers performed all four permutations of adapt/test eye.

5.2.3.2 Disparity task

The procedure was the same as that of the IOT task, except all visual stimuli were now defined by retinal disparity alone and viewed binocularly. The retinal disparity was only presented during the period coinciding with the

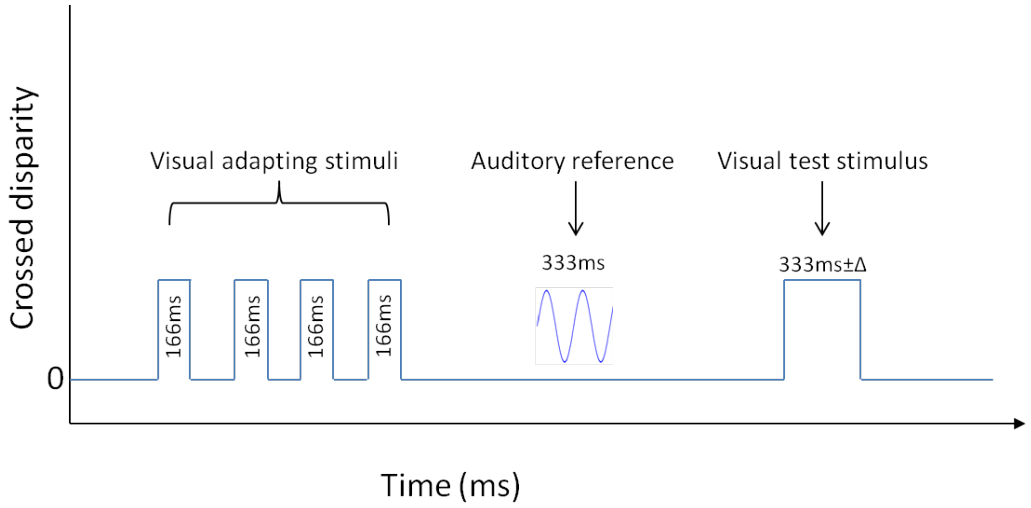

Figure 5.6: Schematic showing the variation in crossed disparity during the course of an adapt-test trial sequence. When the stimulus is presented the disparity changes from zero to 360", returning to zero when the stimulus is not visible.

adapting or test durations ensuring a uniform depth plane across the display during inter-stimulus interval (ISI) periods (see Figure 5.6). In addition, the number of adapting stimuli was reduced to 50 serial presentations at the start of a block, and the number of repetitions within a block was also reduced to five. Observers completed six blocks for each adapting duration (totalling 30 observations per data point).

5.3 Results

For both the IOT and disparity task the proportion of 'test longer than reference' responses were plotted against the physical visual test durations for each adapting duration (166ms and 666ms), and each experimental condition. The psychometric functions were then fitted with a logistic function of the form:

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

Where μ is the test duration that is perceptually equivalent to the 333ms auditory reference duration, and θ is an estimate of the discrimination threshold. This allowed extraction of the point of subjective equality (PSE) values.

For the five observers who performed all four permutations of the adapt/test eye in the IOT task, a paired samples t-test revealed that there was no significant difference in the size of the aftereffect generated in either of the "same" test eye conditions (i.e. adapt and test right eye versus adapt and test left eye, p=0.13), or the "different" conditions (i.e. adapt right, test left versus adapt left, test right, p=0.54). Consequently, for the remaining analysis their data was combined across equivalent conditions.

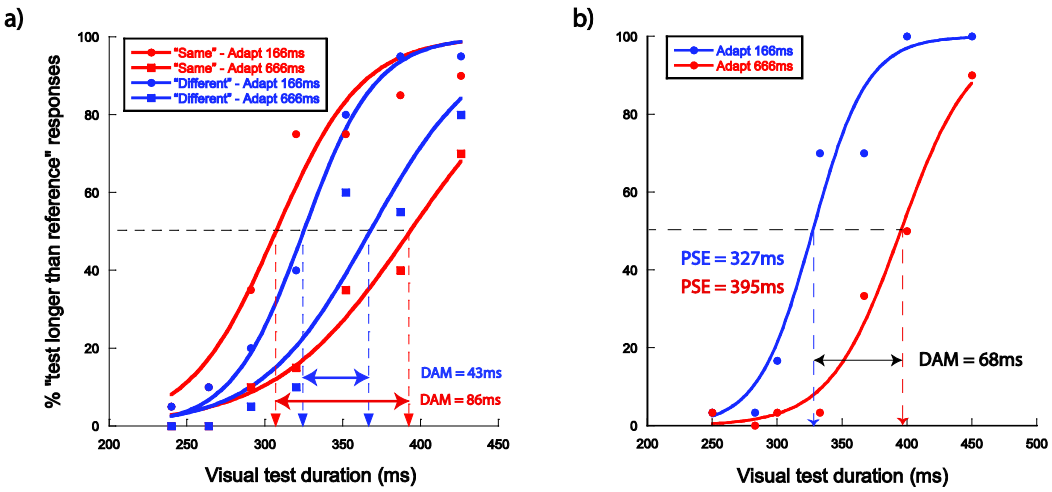

Figure 5.7: Example psychometric functions for **a)** the IOT task and **b)** the disparity task, showing duration discrimination judgements as a function of visual test duration. **a)** Data for representative observer JEM following adaptation to 166ms (circles) or 666ms (squares) duration stimuli for both the “same” condition (red data) and the “different” condition (blue data). **b)** Data for representative observer AR following adaptation to 166ms (blue data) or 666ms (red data) disparity-defined durations. In all cases dashed vertical lines represent extraction of the point of subjective equality (PSE). Horizontal arrows show the DAM values extracted from the arithmetic difference between PSEs in each condition/task.

Psychometric functions for the IOT task and disparity task are shown in Figure 5.7, for representative observers JEM and AR respectively. Figure 5.7a shows functions for both the “same” (red data) and “different” (blue data) adapt-test eye conditions in the IOT task. In relative terms, the proportion of “visual test longer than auditory reference” responses is higher following adaptation to 666ms duration visual stimuli (Figure 5.7a - squares) and lower following adaptation to 166ms duration stimuli (Figure 5.7a - circles). This pattern of responses reflects adaptation-induced perceptual expansion and compression of perceived visual test duration and therefore replicates the repulsive duration aftereffects reported in Chapter 4 of this thesis.

When adapting and testing durations are presented to the same eye (Figure 5.7a – red curves) the relative rightward (adapt 666ms condition) and leftward (adapt 166ms) shifts of the psychometric functions produces disparate PSE values of 393ms and 307ms, respectively. This pattern can also be observed when adapting and testing durations are presented to opposite eyes (Figure 5.7a - blue curves), albeit with a smaller shift in PSE values between 368ms and 325ms.

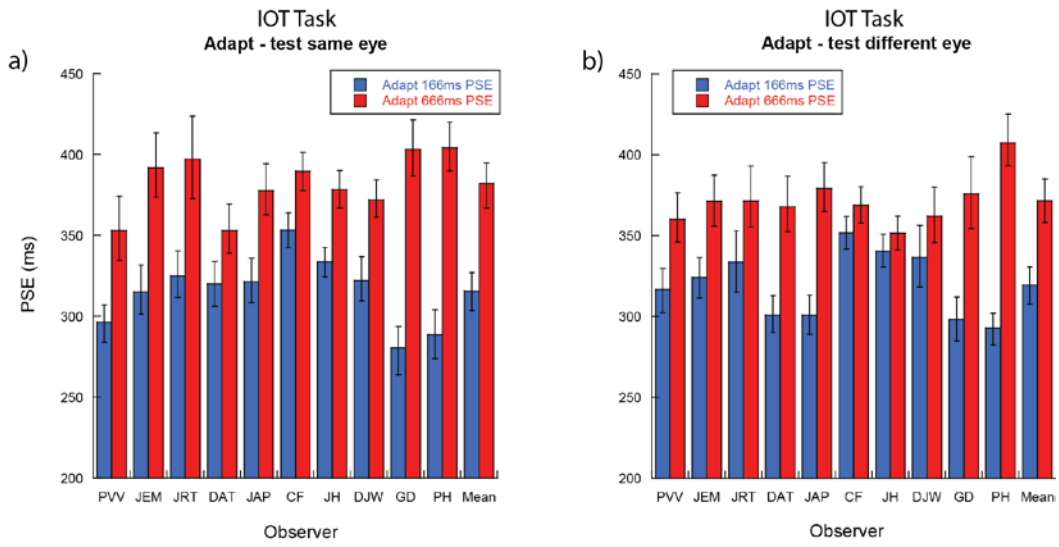

Figure 5.8: Individual PSE data for both the **(a)** "same" and **(b)** "different" conditions of the IOT task, following adaptation to 166ms (blue data) and 666ms (red data) durations. Mean PSEs are shown on the far right, and error bars represent 95% confidence intervals of the bootstrapped PSE distributions.

Figure 5.7b shows that similar (relative) rightward/leftward shifts in the psychometric functions are found following adaptation to 666ms/166ms disparity-defined durations. The PSE value of 395ms following adaptation to 666ms durations (red data) represents a temporal compression of perceived visual test duration. The opposite pattern of perceptual distortion is seen

following adaptation to 166ms durations (blue data), where the PSE value is reduced to 327ms.

In addition, all of the functions shown in Figure 5.7 are slightly rightwards shifted relative to veridical (PSE = 333ms). This is consistent with each observer's tendency to perceive auditory durations as slightly longer then their (physically identical) visual counterparts, irrespective of adaptation (see Chapter 2, Section 2.4.3).

Individual PSE values for each observer are shown in Figure 5.8 for the IOT task and Figure 5.9 for the disparity task. Despite inter-observer differences in absolute PSE values, all observers show perceptual compression/expansion (higher/lower PSE values) following adaptation to relatively long/short durations across both tasks. Importantly, the pattern holds for all observers irrespective of whether adapting and testing durations are presented to the same (Figure 5.8a) or opposite (Figure 5.8b) eyes, and whether the visual stimuli were luminance defined (Figure 5.8) or disparity defined (Figure 5.9). For both tasks, confidence intervals were determined for each PSE value using a non-parametric bootstrapping method, which was based on 1000 samples (see Chapter 3 Section 3.7).

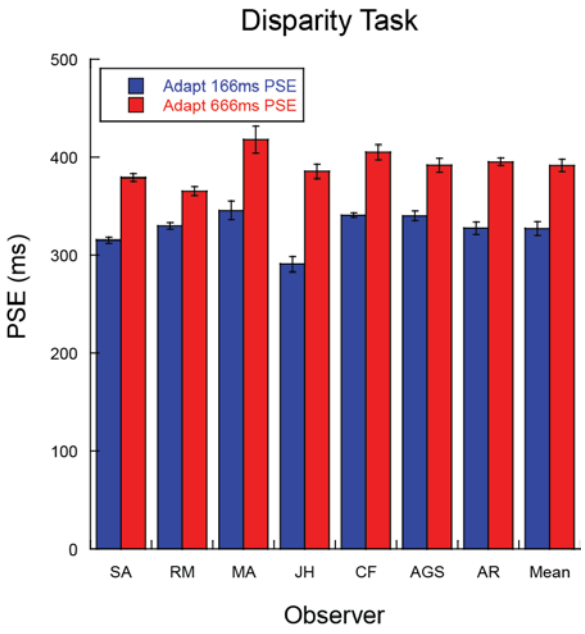

Figure 5.9: Individual PSE data from the disparity task, following adaptation to 166ms (blue data) and 666ms (red data) durations. Mean PSE values are shown on the far right. Error bars reflect 95% confidence intervals of the bootstrapped PSE distributions.

Taking the arithmetic difference between the PSEs for each adapting duration (166ms or 666ms) provides a measure of duration aftereffect magnitude (DAM) (see Chapter 4, Section 4.1.2). DAM values were first calculated for each individual observer, and then averaged across observers for each task (i.e. IOT and disparity). Confidence intervals were determined by calculating the DAM for each pair of bootstrapped PSEs (1000 samples) and extracting the 5% and 95% values from this distribution (see Chapter 3, Section 3.7). For the IOT task, DAM values were 66ms in the “same” condition (Figure 5.10, light blue data) and 52ms for the “different” condition (Figure 5.10, dark blue data), indicating 79% interocular transfer of the DAM when the adapt and test stimuli were presented to different eyes. A paired samples t-test indicated that this reduction was not significant (p=0.11), although the DAM was also significantly greater than zero in both conditions

(p<0.01). This suggests complete interocular transfer of the duration aftereffect.

Averaged across observers, the DAM in the disparity task was 62ms, which was significantly greater than 0ms as determined by a one-sample t-test (p<0.001). This clearly demonstrates that retinal disparity alone generates a duration signal that is accessible to the duration selective mechanism driving the duration aftereffect. By definition, this mechanism must therefore be cortical and binocular in nature (Barlow et al. 1967; Poggio and Fischer 1977).

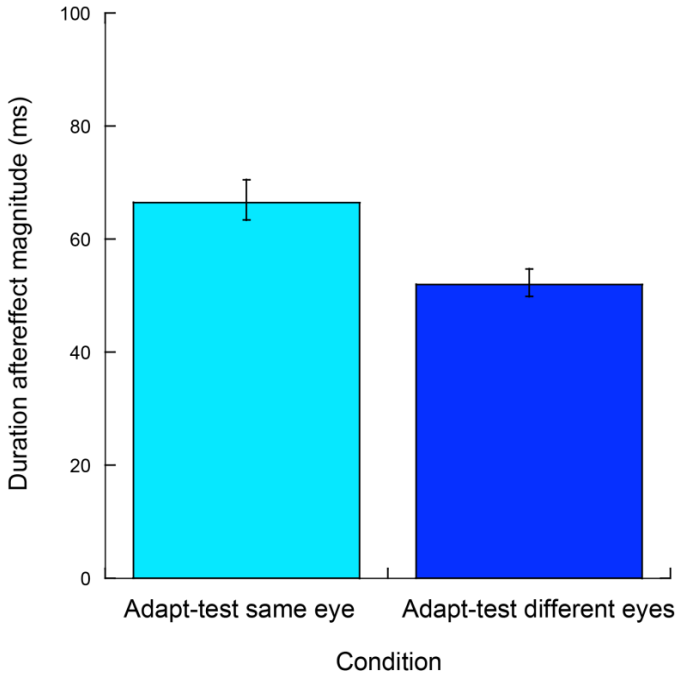

Figure 5.10: Duration aftereffect magnitude (DAM) averaged across all ten observers for the “same” condition and the “different” condition in the IOT task. Error bars represent 5% and 95% confidence intervals of bootstrapped DAM distributions.

5.4 Discussion

The results of the IOT task demonstrate that adapting to relatively short or long visual durations induced repulsive, bidirectional DAEs irrespective of whether adapting and test durations were presented to the same or opposite eyes. Although the magnitude of the aftereffect was lower for the latter, this failed to reach significance, suggestive of full interocular transfer. The disparity task provided evidence that a duration signal generated by retinal disparity alone is sufficient to generate a duration aftereffect.

The IOT findings have parallels with earlier work showing that adapting to auditory temporal frequency induces a bidirectional repulsive rate aftereffect which shows interaural transfer (Becker and Rasmussen 2007). The authors argue that this is indicative of a mechanism which receives input from both ears, and is therefore likely to reside in a central neural location rather than in the peripheral neurons of each cochlea.

Duration aftereffects derived from depth-defined durations confirms that duration-selective neurons responsible for the aftereffect must be either depth-selective themselves (Barlow et al. 1967; Poggio and Fischer 1977) or receive their input from earlier depth selective neurons at the level of the striate cortex or above. This provides compelling evidence that the processing of visual durations occurs at a cortical location. Further useful speculation about the location of duration-selective neurons is limited by the fact that disparity tuned neurons exist in almost every stage of the cortical visual processing hierarchy (Burkhalter and Van Essen 1986; Poggio et al. 1988; Roy et al. 1992; DeAngelis and Newsome 1999; Janssen et al. 1999; Janssen et al. 2003).

Whilst neurons in V1 respond to disparity presented in dynamic random dot stereograms (Prince et al. 2002), the ability to extract shape information from a stimulus may require the more global processing of disparity across a wider region of space. This could be achieved by neurons at a downstream location pooling the inputs across several low level filters to extract the relatively large, depth-defined circular target. For example in primates, neurons tuned to disparity in extrastriate area V2 have been proposed to act as stereoscopic edge detectors (Von der Heydt et al. 2000). These signals might then be integrated by neurons downstream in inferior temporal cortex (IT), which have been shown to respond selectively to two-dimensional disparity-defined shapes (Tanaka et al. 2001). Nevertheless, it is also possible that observers could extract the depth (and the duration) signal from disparity occurring within a much smaller region of space (e.g. 0.25°), without the need to extract global shape by integrating depth inputs across the display. This local processing could be achieved by linear, depth-selective luminance filters in V1, although the DAE's broad, scale-dependant spatial tuning (see Chapter 4) is perhaps more consistent with neurons in extrastriate cortex. Explicit testing of the DAE's selectivity to mid-level stimulus characteristics (e.g. adapting to durations defined by contrast modulations that are vertically oriented in the adapting phase and horizontal in the test phase) may aid differentiation between the two.

Although the extraction of depth-defined durations clearly requires the function of depth-selective neurons, it may be the case that the duration adaptation mechanism is not selective for depth itself. Depth signals may in fact be pooled by a late-stage duration mechanism in order to allow a

consistent duration percept across changes in an object's viewing angle and/or distance. To determine whether the duration mechanism is in fact selective for depth, it would be necessary to look for evidence of DAE tuning for retinal disparity. For example, adapting to durations presented in crossed disparity and testing with durations presented in uncrossed disparity would test this possibility.

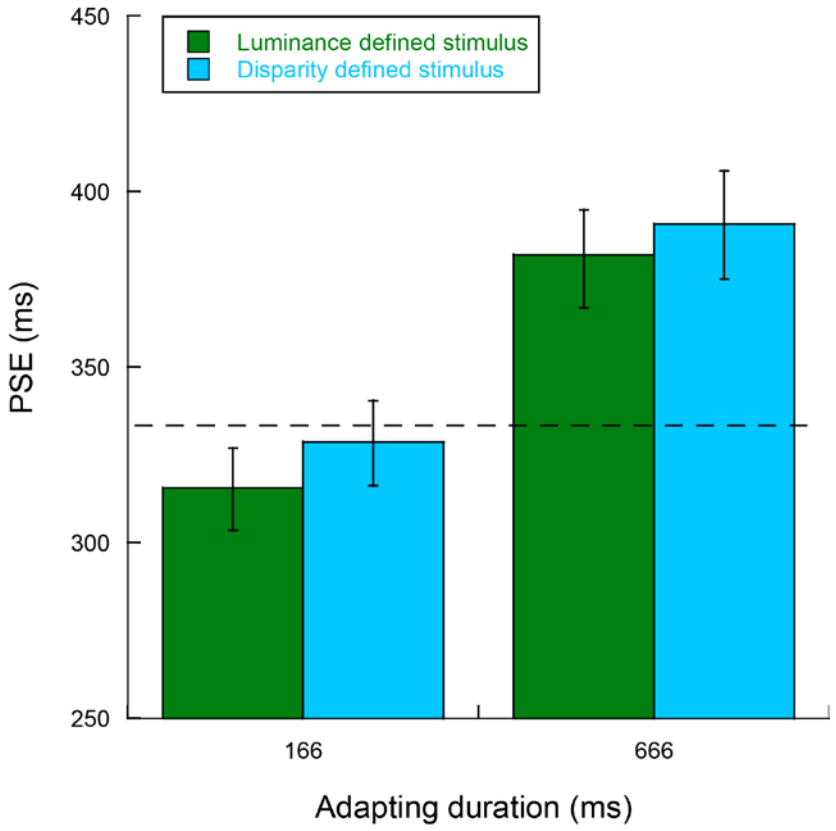

Figure 5.11: Comparison of mean PSEs as a function of adapting duration for the "adapt-test same eye" condition of the IOT task (luminance defined – green data) and the disparity task (disparity defined – blue data). The horizontal black dashed line represents veridical duration perception. Error bars represent 95% confidence intervals derived from 1000 bootstrapped samples.

The use of a dynamic stimulus in the disparity task does have some parallels with the compressive aftereffects induced following adaptation to 20Hz visual flicker (Johnston et al. 2006; Johnston et al. 2008: see Chapter 2 Section

2.4.4). In the present study, the 24Hz dynamic luminance noise utilised in the disparity task could have resulted in generalised temporal compression of perceived test duration relative to the IOT task (which used only static visual stimuli). This would potentially be superimposed on top of DAEs driven by adaptation to stimuli of fixed duration, and would manifest as higher overall PSE values in the disparity task, irrespective of adapting duration. On inspection, little difference was seen between the averaged 166ms PSEs or 666ms PSEs in either task, although a slight trend towards higher PSEs in the disparity task was observed (see Figure 5.11). This suggests that dynamic luminance noise may have caused a small compressive-type effect on perceived duration relative to the static stimuli used in the IOT task. However, two key differences exist between the experiments of Johnston et al. (2006; 2008) and the present study. Firstly, Johnston et al. used an adapting stimulus that was continuously presented (≈15 – 20s), whereas in the present study observers adapted to intermittent presentations of fixed duration stimuli. Secondly, Johnston et al. required their observers to attend to the dynamic adapting stimuli (and hence attend to temporal rate), yet in the present study observers were instructed to attend to the *duration* of the adapting stimulus and *not* to the dynamic luminance noise. Thus the compressive-type effect described by Johnston et al. may have been limited in the present study by methodological differences.

In summary, two complementary tests of the duration aftereffect's binocularity provide compelling evidence that the underlying mechanism has a cortical neural locus subserved by binocular neurons that can encode durations that are invisible to monocular mechanisms.

Chapter 6: High versus low-level stimulus specificity of duration aftereffects

Introduction

The results from Chapter 4 and 5 of this thesis indicate that the neurons responsible for driving the duration aftereffect (DAE) demonstrate scale dependent spatial tuning and are binocular in nature. DAEs can also be elicited using a stimulus defined only by retinal disparity, a stimulus attribute extracted no earlier than primary visual cortex (V1). Collectively, these findings point to a cortical locus for duration processing, yet the relative position of duration selective neurons within the visual processing hierarchy is still unclear.

It is well established that the processing of spatial information in the visual cortex occurs via an interconnected network of functionally specialised areas (Zeki et al. 1991). Whilst a vast number of feedforward, lateral and feedbackward connections exist between these areas, it is generally accepted that the system processes visual signals in a hierarchical fashion (Felleman and Van Essen 1991; Grill-Spector and Malach 2004). This concept is schematised in Figure 6.1: early visual areas process low-level features such as lines and edges, whilst later stages of processing collate information from preceding stages to process the visual scene more holistically, resulting in more complex, high-level pattern analysis.

V1 is largely tasked with encoding local changes in luminance intensity, in particular detecting lines and edges to form a ‘primal sketch’ of the visual scene / object (Marr 1982). The retinotopic arrangement of V1, in which

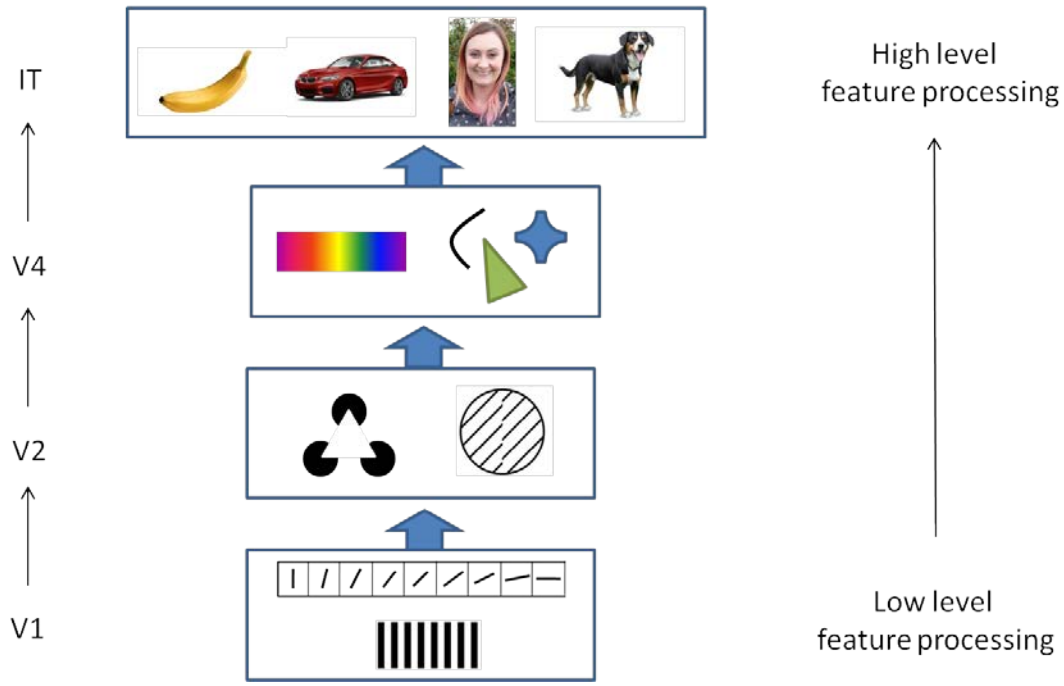

Figure 6.1: Schematic showing hierarchical feature processing within a subset of functionally specialised areas of the human visual cortex. Initially low-level stimulus characteristics such as orientation and spatial frequency are extracted (V1). Progressively, as successive stages of processing are encountered (see blue arrows), neurons respond to increasingly complex features including illusory contours and texture (V2), colour and shape (V4), and objects and faces (IT).

adjacent neurons have receptive fields that correspond to neighbouring regions of visual space, creates a ‘map’ of the visual world within the cortex. For each section of the map, neurons with overlapping receptive fields show an orderly progression of orientation and spatial frequency tuning (Hubel and Wiesel 1962; Hubel and Wiesel 1968; Tootell et al. 1981; De Valois et al.

1982a), such that V1 is often described as a network of spatially localised, two dimensional band-pass filters.

Retinotopic mapping is also evident in extrastriate areas V2 and V4. Neurons in both of these regions also respond to luminance bars and edges, but in addition show tuning preferences to increasingly complex image properties. For example, V2 appears to have a role in detecting illusory contours or texture defined boundaries (Von Der Heydt et al. 1984; Von der Heydt and Peterhans 1989: - see Figure 6.2), and neurons in V4 may be tuned to specific colours or shapes (Zeki 1973; Desimone and Schein 1987; Kobatake et al. 1994; Pasupathy and Connor 1999).

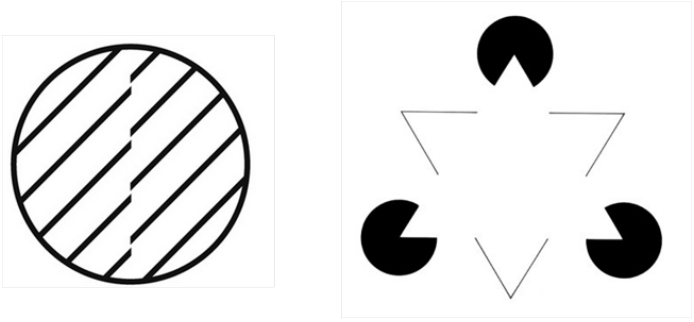

Figure 6.2: Examples of illusory contours, in which edges are perceived despite the absence of a luminance boundary defining the contours.

In higher visual areas such as inferior temporal cortex (IT), neurons show tuning preferences for specific objects, shapes or faces, but respond weakly to simple luminance bars or edges (Gross et al. 1972; Perrett et al. 1982; Desimone et al. 1984; Kanwisher et al. 1997). In addition, the responses of these neurons are invariant to transformations in stimulus size, colour, contrast, spatial position, and (to a lesser extent) viewing angle (Perrett et al.

1982; Schwartz et al. 1983; Desimone et al. 1984; Rolls and Baylis 1986; Sáry et al. 1993). Models of object recognition have proposed that these characteristics reflect pooling of afferent signals from previous stages of processing (Riesenhuber and Poggio 1999; Rolls et al. 2000; Freeman and Simoncelli 2011), a notion that is supported by the progressive increases in receptive field size from V1 to IT (Hubel and Wiesel 1968; Desimone and

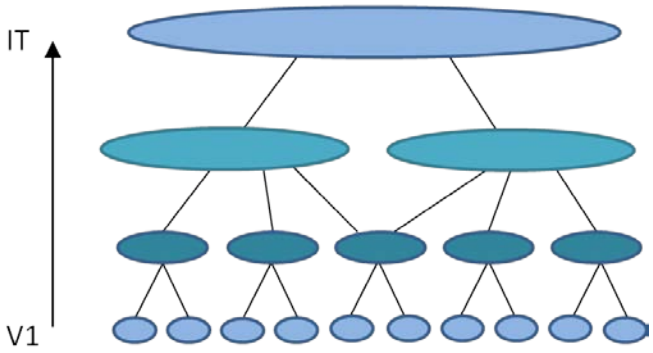

Figure 6.3: Schematic demonstrating how increases in receptive field occur by pooling across afferent signals originating from each previous processing stage.

Gross 1979; Gattass et al. 1981; Gattass et al. 1988). This pooling could explain why fewer IT neurons demonstrate selectivity to low-level features such as orientation (Desimone et al. 1984), if signals are averaged across multiple lower level inputs (see Figure 6.3).

This chapter presents a series of experiments designed to test the selectivity of duration encoding to a range of visual characteristics of varying complexity, with the aim of inferring potential loci for duration encoding within the visual processing hierarchy. Given the prevalence for orientation tuning in early visual areas such as V1 (Hubel and Wiesel 1968), we begin by examining selectivity to orientation.

Experiment 6.1

The majority of neurons in V1 show bandwidth limited response selectivity for the orientation of a luminance bar, grating or edge (Hubel and Wiesel 1962; Hubel and Wiesel 1968; De Valois et al. 1982b). Thus, different neural populations are likely stimulated by luminance gratings that differ in orientation by 90°, hence the lack of transfer of the "tilt aftereffect" (see Chapter 2, Section 2.3.3) across gross changes in orientation between adapt and test phases (Gibson and Radner 1937; Mitchell and Muir 1976). If duration is extracted by orientation tuned neurons at a relatively early processing stage, it is predicted that the duration aftereffect will also be abolished when orientation is varied by 90° between adapting and test phases. Conversely, if the aftereffect transfers across gross changes in orientation, this would indicate that the underlying duration selective neurons pool across this stimulus feature.

6.1.1 Methods

6.1.1.1 Observers

Seven observers (four naïve) took part in the experiment. All observers gave their informed, written consent to participate, and had normal or corrected to normal vision and hearing at the time of the experiment.

6.1.1.2 Stimuli and apparatus

Visual stimuli were presented on a gamma-corrected Eizo FG2421 LCD monitor with a refresh rate of 120Hz and a resolution of 1920x1080. This was connected to a 3GHz E5-1660v3 8-Core HP Z440 desktop computer running Windows 8.1 Pro. All stimuli were generated using Matlab 8.4.0 (Mathworks, USA) running Psychtoolbox extension version 3.0.11 (Brainard and Pelli, 1997, www.psychtoolbox.org). The auditory stimulus was a 500Hz tone presented through Sennheiser HD 280 headphones. Visual stimuli were Gabor patches comprising a sinusoidal luminance carrier with a Gaussian contrast envelope. The spatial frequency of the carrier was 2 cycles per degree (cpd) and the Gaussian envelope size, defined by the standard

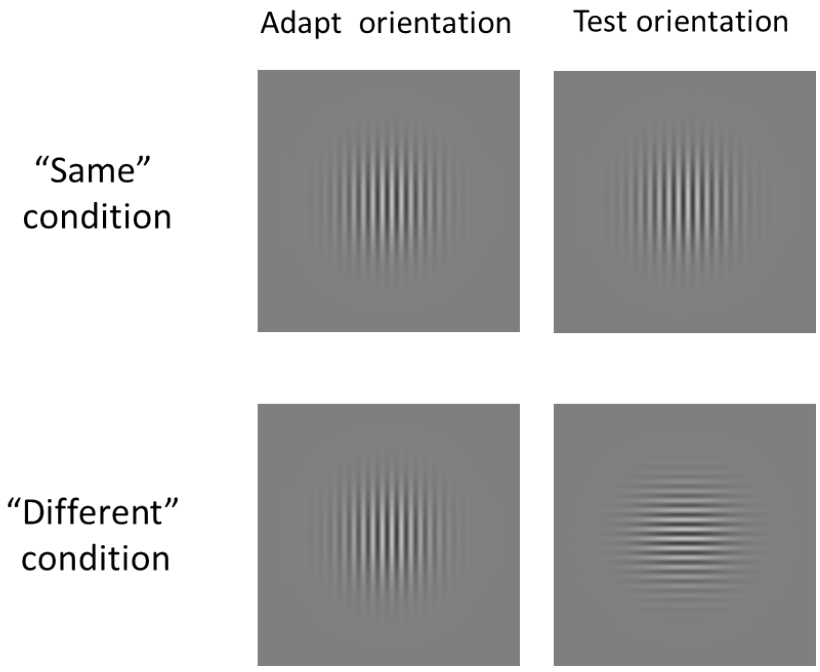

Figure 6.4: Schematic showing the conditions used in Experiment 6.1. In the "same" condition both the adapt and test stimuli were orientated at 90°, and in the "different" condition the adapt stimulus was orientated at 90° and the test stimulus was orientated at 180°.

deviation of the envelope (σ_{Stim}) was 1.5°. The phase of the carrier varied randomly on each presentation of the visual stimulus. In all cases, the adapting stimulus' carrier orientation was set to 90°. In the "same" condition the test stimulus' carrier orientation was also 90°, and in the "different" condition it was set to 180° (see Figure 6.4).

The physical durations of all auditory and visual stimuli were verified using a dual-channel oscilloscope. The viewing distance of 93cm was maintained using a chin and forehead rest to ensure that one pixel subtended 1 arc minute.

6.1.1.3 Procedure

The adapt-test experimental design used to gather this data was identical to Experiment 4.1 (see Chapter 4, p.152), with the exception that all stimuli were presented at fixation. Observers adapted to 100 sequential presentations of either relatively long (666ms) or short (166ms) fixed duration visual stimuli, during which the interstimulus interval (ISI) between each successive stimulus was randomly jittered between 500 – 1000ms. Following four additional 'top-up' stimuli, observers performed a 2AFC duration discrimination judgment to decide "which was longer?" between a 333ms auditory stimulus and a visual test stimulus which varied in seven approximately logarithmic steps around 333ms. Responses were made via a keypress, and this triggered the next "top-up / test" cycle.

Observers performed three blocks for each adapting duration condition (comprising 30 observations per data point), and also three for each

orientation condition. This totalled twelve blocks per observer (2 adapting durations x 2 test stimulus orientations x 3 repetitions), which equated to approximately two hours of observing.

6.1.2 Results and discussion

For both adapting durations (166ms and 666ms) and both orientation conditions ("same" and "different") the proportion of 'test longer than reference' responses were plotted against the physical visual test durations using the same procedure as described in Chapter 4. Fitting logistic functions to these data allowed extraction of the PSE values for each observer. Psychometric functions for representative observer BAA are shown in Figure 6.5 for both the "same" (red data) and "different" (blue data) conditions. Adaptation to 666ms durations in the "same" condition (red squares) led to a temporal compression of the visual test stimulus' perceived duration, such that it had to be expanded to 367ms to maintain perceptual equivalence with the 333ms auditory reference duration. Conversely, the 'adapt 166ms function' (red circles) is laterally shifted leftwards, representing a temporal expansion of perceived duration. Here the test stimulus needed to be shortened to 325ms to feel perceptually equivalent to the auditory reference. A similar pattern of temporal compression/expansion can also be seen following adaptation to long/short durations in the "different" condition (blue data points).

For each observer, the arithmetic difference between adapt 166ms and adapt 666ms PSE values was calculated to find the duration aftereffect

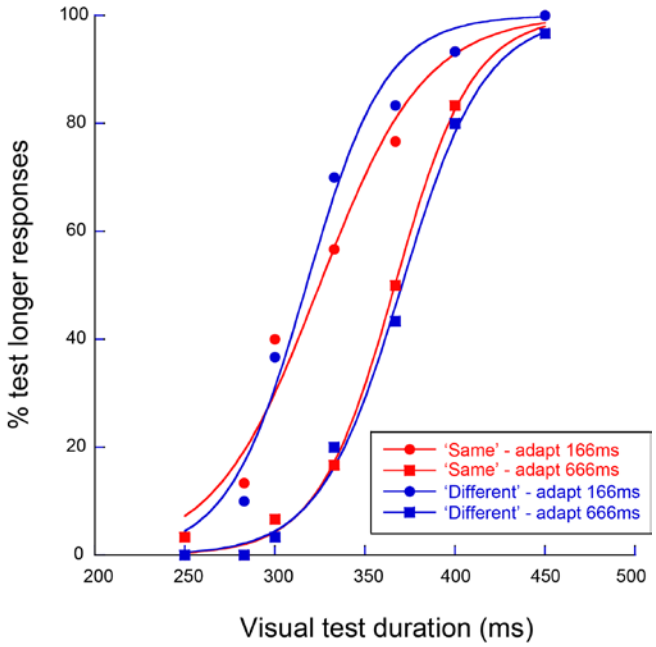

Figure 6.5: Psychometric functions for representative observer BBA, showing duration discrimination judgements as a function of visual test duration. Data are shown following adaptation to 166ms (circles) or 666ms (squares) duration stimuli for both the "same" condition (red data) and the "different" condition (blue data).

magnitude (DAM) for both orientation conditions. The DAM values averaged across all seven observers are shown in Figure 6.6, and were 51ms for the "same" condition and 47ms for the "different" condition. A paired samples t-test found no significant difference between orientation conditions ($p=0.584$), indicating full transfer of the duration aftereffect across 90° changes in carrier orientation.

These results indicate that the mechanism driving the duration aftereffect is not selective for carrier orientation. This finding is consistent with a recent study by Li et al. (2015), in which observers adapted to alternate

presentations of horizontal and vertical Gabor patches before performing a duration discrimination task between an auditory reference and a variable visual test duration (which varied between a vertical orientation and horizontal orientation on each trial, see Figure 6.7). In their “congruent”

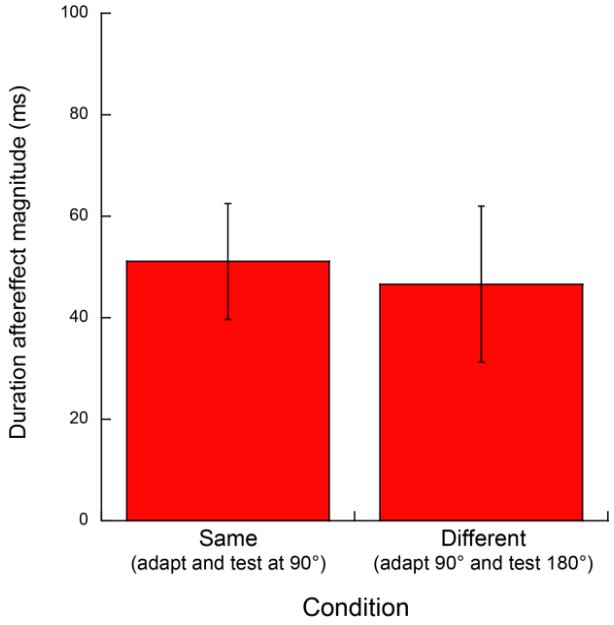

Figure 6.6: Group mean DAM data (n=7) showing duration aftereffect magnitude for adapting and testing at 90° (“same”) and adapting at 90°, testing at 180° (“different”). Error bars represent the SEM.

conditions, all visual stimuli had a fixed duration (either 160ms or 640ms) irrespective of orientation, whilst in their “incongruent” conditions the vertical stimuli were 160ms and horizontal stimuli were 640ms (or vice versa). Duration aftereffect magnitude was calculated from trials containing vertical test stimuli and trials containing horizontal test stimuli separately, giving two ‘congruent’ and two ‘incongruent’ values. The results showed the same pattern irrespective of test orientation: a significant duration aftereffect was evidenced only in the congruent condition. Any temporal expansion elicited by adaptation to the 160ms duration in the incongruent condition was

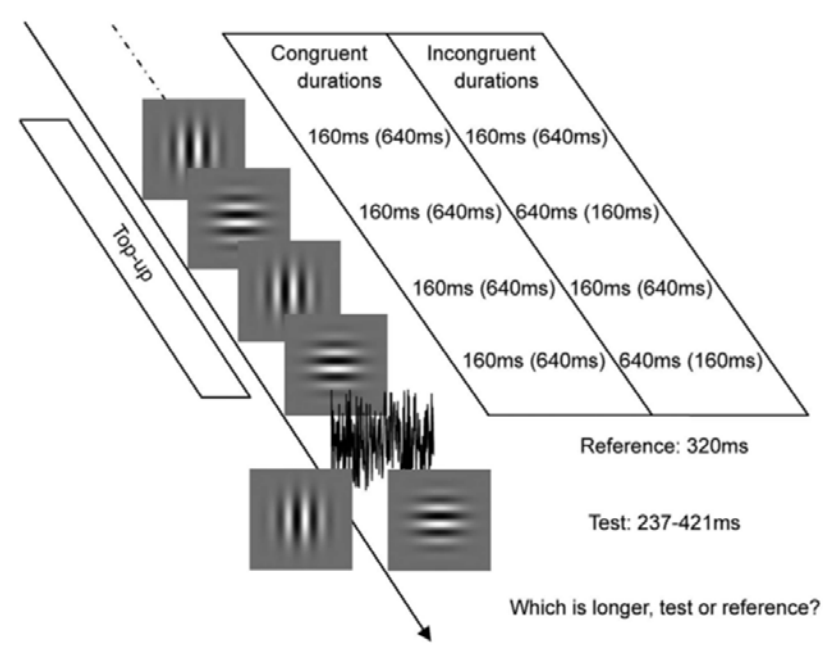

Figure 6.7: Schematic showing the experimental paradigm used by Li et al. (2015b) (see main text for details). The components of a single trial are shown above. Initially observers viewed four top-up adapting stimuli, before making a duration discrimination judgement between an auditory reference (shown as a black noise burst) and a variable visual test stimulus (which was either vertical or horizontal on different trials). Figure reprinted with permission from Macmillan Publishers.

assumed to have been cancelled out by co-occurring temporal compression following adaptation to the 640ms duration.

One explanation for the reported lack of orientation selectivity is that duration signals could be extracted at a pre-cortical location. Receptive fields within the retina and LGN have concentric excitatory and inhibitory regions (Kuffler 1953; Hubel 1960) and consequently these neurons possess no orientation tuning. However, such a scenario is made unlikely by the broad, scale dependent spatial tuning (see Chapter 4) and binocular characteristics of the duration aftereffect (see Chapter 5) reported in earlier chapters.

An alternative proposal is that duration is extracted at a later stage of processing, beyond that of low-level stimulus characteristics such as orientation. Neural areas attributed to high-level object/face identification are thought to achieve complex analysis (and viewpoint invariance) by pooling across the output of neurons tuned to low-level features at progressively earlier stages within the hierarchy. If duration is extracted at a relatively late stage, the merging of afferent signals could potentially facilitate transfer of the DAE across gross changes in orientation (see Figure 6.8). To probe the issue of higher-level involvement, it would therefore be necessary to utilise a complex visual stimulus whose identification relies on the pooling of low-level features.

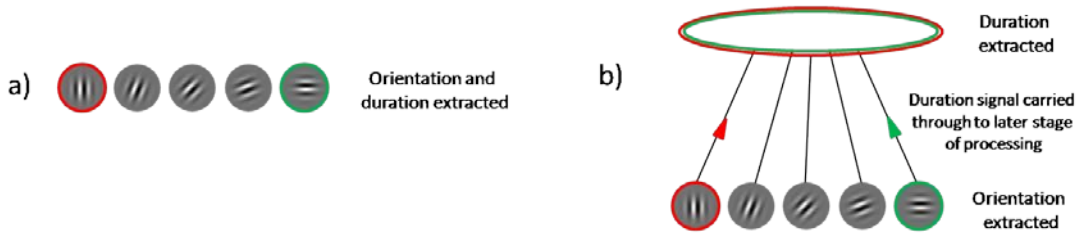

Figure 6.8: Schematic showing five neurons that are selective for stimulus orientation. **a)** If duration is extracted alongside orientation, there would be no transfer of the duration aftereffect across the adapting orientation (highlighted in red) and the test orientation (highlighted in green). **b)** If duration is extracted at a later stage of processing, where signals from multiple neurons are combined, the duration aftereffect may transfer across changes in orientation.

Whilst the processing of face information is generally ascribed to a small region of association cortex known as the fusiform face area (FFA) (Kanwisher et al. 1997), this area is likely to receive input from local edge or shape detectors at earlier stages in the hierarchy. A recent neuroimaging study found that increasing the size of a face stimulus resulted in increases

in neural activity which were highly correlated across both V1 and FFA (Yue et al. 2011). Additionally, manipulating the low-level features of the face stimulus (e.g. size and contrast) resulted in greater FFA activation in response to large, high contrast *non-face* objects than to the small, low-contrast faces. This suggests that processing in FFA is strongly influenced by neural activity at earlier stages of the processing hierarchy.

In addition, local adaptation to curves or tilted lines can alter perceived facial expression in photographs of real faces (Xu et al. 2008; Dickinson et al. 2010; Dickinson et al. 2012; Dickinson and Badcock 2013), and there is growing evidence that adaptation in early visual areas can cascade through successive processing stages (Kohn and Movshon 2003; Dhruv and Carandini 2014; Larsson and Harrison 2015). If manipulating neural responses to low and mid-level features can directly influence high-level image analysis, it is likely that the processing of faces involves the merging of afferent signals. Therefore, in Experiment 6.2 we examine DAE selectivity at higher stages of the hierarchy using stimuli comprised of differing facial identities. If event duration is encoded by neurons that also process facial identity, we predict that the DAE will be abolished by changes in facial identity between the adapt and test phases, as these will activate different neural populations with distinct identity tuning preferences.

Experiment 6.2

6.2.1 Methods

6.2.1.1 Observers

Nine observers (six naïve) took part in the experiment. All observers gave their informed, written consent to participate, and had normal or corrected to normal vision and hearing at the time of the experiment.

6.2.1.2 Stimuli and Apparatus

The apparatus and auditory stimulus used were identical to that described in Experiment 6.1. The visual stimuli were synthetic faces derived from grayscale photographs (see Figure 6.9, and for a full description of stimulus generation see Wilson et al. 2002). These faces were designed so that facial identity strength could be controlled as percentage deviation from an 'average' face (based on 40 male individuals).

With permission from the authors, two facial identities were chosen at random from the cohort of 40 male synthetic faces (see Figure 6.9). These were generated using the original custom software written in Matlab (Mathworks, USA) and converted to TIFF files for use during the experiment. Both faces differed in facial identity by 25% from the average face (n=40). This ensured that facial identity differences were around 5 multiples of typical identity discrimination thresholds (Logan et al. 2016). Unlike straightforward photographic face images these synthetic faces allowed greater control over the low-level features within the stimulus (e.g. luminance and spatial

frequency), yet retained sufficient geometric information to allow accurate identification of facial identity. They also allow precise control over the identity information signal strength and therefore the extent to which two

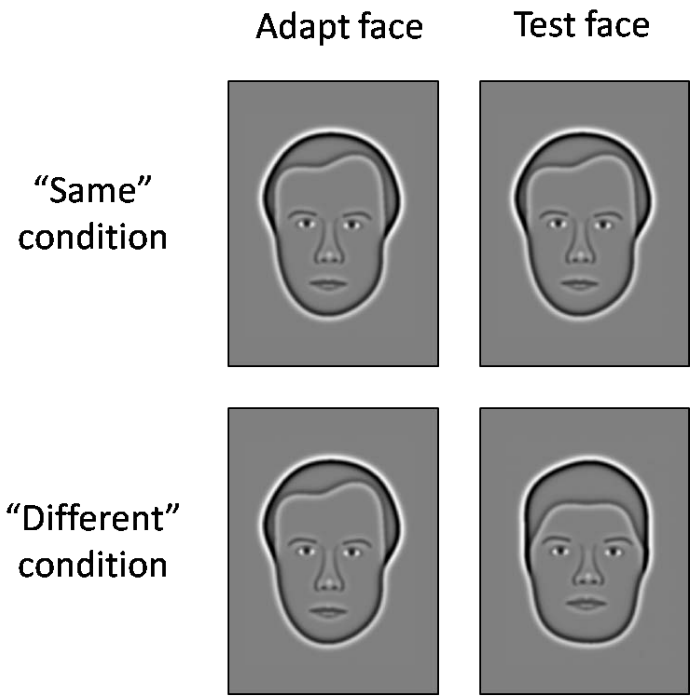

Figure 6.9: Schematic showing the synthetic face stimuli used in the "same" condition and the "different" condition of Experiment 6.2. Each of the two facial identities is 25% different from the average face.

faces can be discriminated from one another via identity alone. The stimuli measured 5cm by 6.5cm (subtending a visual angle of ≈3°) when presented on the screen at a distance of 93cm. This size was chosen so as to be approximately similar to visual stimuli deployed in earlier experiments (see Chapters 4 and 5).

6.2.1.3 Procedure

The procedure was the same as described in Experiment 6.1, albeit with face stimuli instead of Gabor patches. Observers performed 3 blocks for each condition (30 observations per data point), totalling 12 blocks overall (2 adapting durations x 2 test stimulus orientations x 3 repetitions). Data collection was approximately 2 hours per observer.

6.2.2 Results and discussion

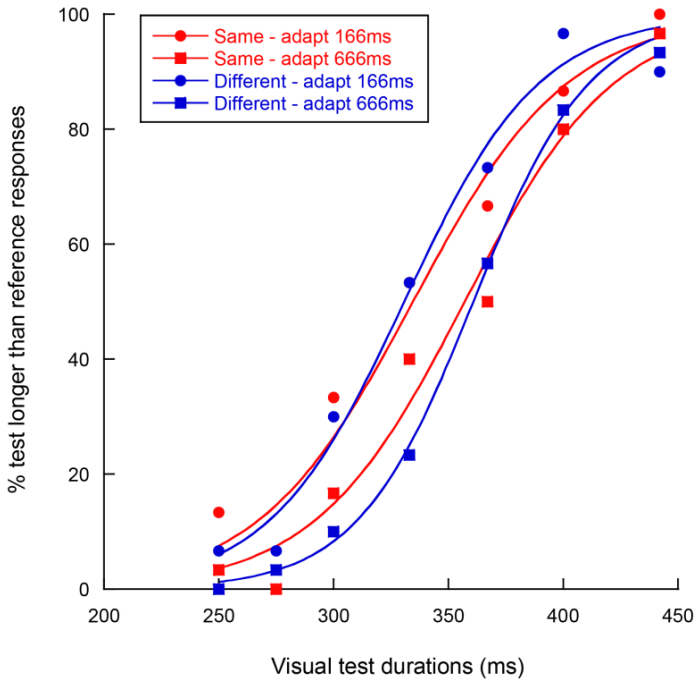

Figure 6.10: Psychometric functions for representative observer AGS showing the adapt-test "same" facial identity (red data) and adapt-test "different" facial identity (blue data) conditions following adaptation to 166ms (circles) and 666ms (squares) durations.

The proportion of "test longer than reference" responses was plotted against the visual test durations, and psychometric functions were again fitted with a logistic curve. Psychometric functions for representative observer AGS are

shown in Figure 6.10 for both the adapt-test “same” facial identity (red data) and adapt-test “different” facial identity (blue data) conditions. In the “same” condition, adaptation to 666ms durations (red squares) resulted in perceptual compression of the visual test stimulus’ duration, such that it had to be expanded to 357ms in order to maintain perceptual equivalence with the

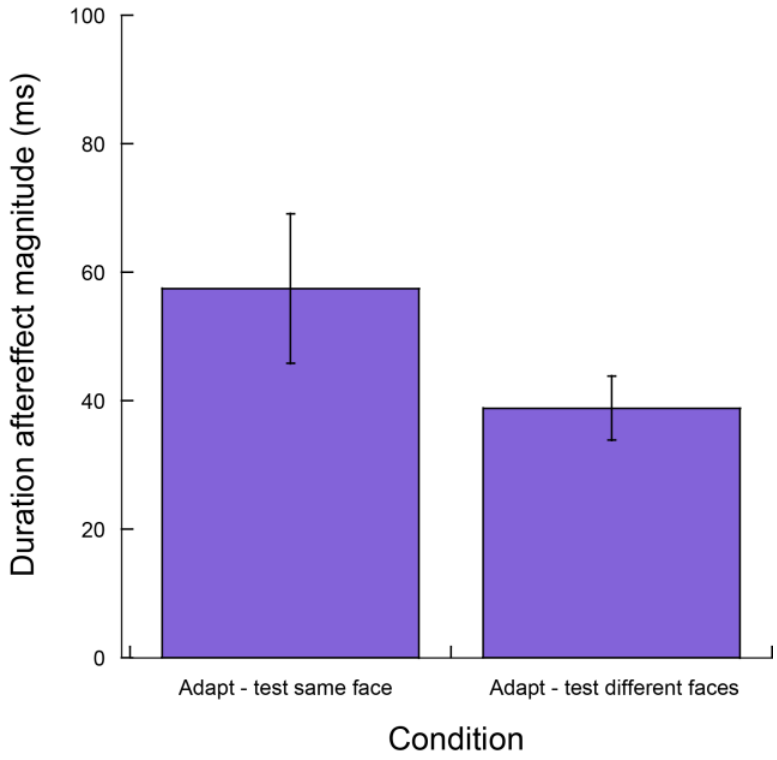

Figure 6.11: Mean facial identity data (n=9) showing duration aftereffect magnitude for the “same” (adapt and test face 1) and “different” (adapt face 1, test face 2) conditions. Error bars represent the SEM.

333ms auditory reference. In comparison, adapting to 166ms durations (red circles) resulted in a temporal expansion of perceived duration, as witnessed by the relative shift of the function leftwards. The PSE values were not symmetrical around 333ms due to the observer’s “sound longer than vision” bias (Goldstone and Lhamon 1974; Wearden et al. 1998: see Chapter 2, Section 2.4.3). The same pattern of compression/expansion was also evident

following adaptation to long/short durations in the “different” condition (blue squares / circles).

For each observer the duration aftereffect magnitude (DAM) was calculated for both the “same face” and “different face” conditions, by taking the arithmetic difference between the adapt 166ms and adapt 666ms PSE values.

DAM values averaged across observers (n=9) are shown in Figure 6.11. When both adapt and test stimuli contained the same facial identity (“same” condition) the DAM value was 57ms, and when facial identity was altered between adapt-test phases (“different” condition) this fell to 39ms.The trend indicates a reduction in aftereffect strength when adapting and testing faces were of distinct facial identities. However, a paired samples t-test showed that this reduction failed to reach significance ($p=0.156$). The results therefore indicate that the duration aftereffect is unlikely to be selective for facial identity.

The apparent lack of strong selectivity to both low-level (orientation) and high-level (facial identity) visual features could be an indication that the duration mechanism is not selective for any identifying visual features. However, in Experiment 6.2 the lack of selectivity could also have arisen due to the complex nature of facial identity processing. Whilst the processing of low-level features such as orientation is generally attributed to orientation-tuned neurons in V1, it is unlikely that every unique facial identity would have a corresponding neuron tuned to its specific combination of visual features. In this scenario, the number of possible facial identities (including differing

clothes, viewpoints, hairstyles, lighting etc...) would be enormous, rendering this biologically implausible. Instead it is proposed that identity is encoded by how much any given face differs from an 'average' face, which is determined from prior experience and exposure (Leopold et al. 2006). It is possible that on this spectrum, our two facial identities may have activated partially overlapping neural populations, facilitating the transfer of duration information. Furthermore, although the initial stages of facial feature encoding are likely to occur through hierarchical visual processing, the *perception* of facial identity may involve a network of interconnected cortical regions and may be mediated by semantic content and attention (Haxby 2000; Ishai et al. 2005). As images become more complex, it becomes very difficult to isolate specific neural regions experimentally. With increasing numbers of areas being recruited for facial processing it is possible that feedforward and feedback pathways between these areas facilitated transfer of the DAE across different identities.

Alternatively, the complexity difference between facial identity and luminance-defined orientation (Experiment 6.1) leaves open the possibility that duration may be extracted after luminance-defined orientation but prior to complex pattern analysis. Although it failed to reach significance, the trend towards reduced aftereffect transfer across changes in facial identity could reflect some selectivity to mid-level visual features (e.g. curves or shapes), which could not be fully equated across the two faces. If duration is extracted at a mid-level stage of processing, selectivity to mid-level features could have been masked by the subsequent pooling of facial identity mechanisms further up the hierarchy. The spatial tuning of the DAE described in Chapter

4 is also consistent with a relatively early neural locus for duration extraction, albeit one that pools signals across space, resulting in the proportional relationship between the size of the adapting stimulus and the spread of the aftereffect.

Whilst neurons in V1 act as linear filters, detecting edges defined by local changes in luminance (see Figure 6.12 solid white box), they cannot detect boundaries defined by changes in contrast, across which the average luminance remains constant (see Figure 6.12 dashed white box). This task

Figure cannot be displayed due to copyright law. To view the original figure please see Li et al. (2014), Figure 1 (p12082).

Figure 6.12: Photograph of a natural scene. The solid white box shows a region in which the tree can be distinguished from the grass through changes in first-order luminance. The dashed white box shows an area of the image in which the grass is reflected in the water. Here the average luminance is equal across both parts of the image, and the grass can only be distinguished from its reflection in the water through changes in contrast (requiring second-order analysis) (Li et al. 2014).

could be achieved by "second-order" filters, following a process of non-linear "rectification" (Chubb and Sperling 1988; Malik and Perona 1990; Wilson et al. 1992). Extracting spatial variations in contrast requires pooling inputs from several linear (first-order) filters, and subsequently can be considered a more global form of image analysis. Selectivity to second-order image statistics

(e.g. those in which edges or shapes are defined by texture or contrast) has been demonstrated in extrastriate area 18 of cats (an analogue of human V2) and V2 in primates, suggesting that this analysis may reflect processing subsequent to the earliest stage of visual processing (Von Der Heyclt et al. 1984; Zhou and Baker 1993; Mareschal and Baker 1999).

If the extraction of event duration occurs at a similar location within the processing hierarchy, the mechanism may be selective to image properties that also require a more global form of spatial analysis. Experiment 6.3 therefore examines selectivity of the DAE to changes in stimulus size defined only by spatial variations in contrast. The visual stimuli comprise of Gabor patches, which allow independent control over first-order luminance-defined spatial information and contrast-defined information. The former pertains to the Gabor's carrier and the latter to the contrast envelope. In Experiment 6.3 carrier spatial frequency and orientation are held constant across adapt and test phases, whilst the contrast-defined envelope size is either held constant, increased or decreased across adapt-test phases.

As receptive field size varies inversely with spatial frequency tuning in the primary visual cortex (Movshon et al. 1978; De Valois et al. 1982a; Foster et al. 1985), it is unlikely that the relatively high spatial frequency of the carrier and the predominantly low spatial frequency content of the envelopes could both be encoded by the same first-order receptive field. First-order neurons tuned to the spatial frequency of the envelope won't be activated because the net luminance across the Gabor's surface sums to zero. Any DAE selectivity to envelope size may therefore be attributed to the involvement of second-order mechanisms, which encode envelope size by pooling across

the rectified output of multiple, spatially-abutting first-order receptive fields to extract global changes in contrast.

Experiment 6.3

6.3.1 Methods

6.3.1.1 Observers

Eight observers (five naïve) took part in the experiment. All observers gave their informed, written consent to participate, and had normal or corrected to normal vision and hearing at the time of the experiment.

6.3.1.2 Stimuli and apparatus

The stimuli and apparatus were identical to those used in Experiment 6.1, with the exception that two different envelope sizes (defined by the standard deviation, σ_{stim}) were used for the Gabor patch. These were σ_{stim}= 0.5° ("small" Gabor) and σ_{stim} = 1.5° ("large" Gabor). In all experimental conditions the carrier had a spatial frequency of 2 cycles per degree (cpd), a randomly varying phase and an orientation of 90°, in order to keep first-order spatial information constant (see Figure 6.13).

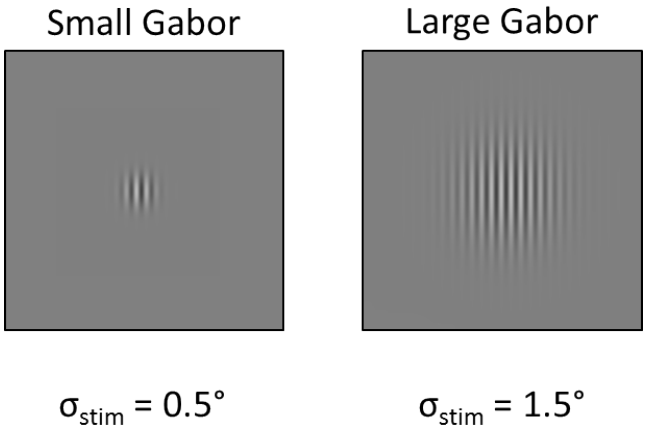

Figure 6.13: The Gabor stimuli used in Experiment 6.3. The small size Gabor had an envelope size (σ_{stim}) of 0.5° and the large size Gabor of 1.5°. In all experimental conditions the orientation of the carrier was 90° and spatial frequency of the carrier was 2 cycles per degree.

6.3.1.3 Procedure

The procedure used was identical to Experiment 6.1, except that carrier orientation and spatial frequency were constant across conditions and stimulus size was varied. All four permutations of adapt / test stimulus size were performed, such that there were two "same" conditions (adapt 0.5° / test 0.5° and adapt 1.5° / test 1.5°) and two "different" conditions (adapt 0.5° / test 1.5° and adapt 1.5° / test 0.5°). This enabled additional investigation into the dependency of overall strength of the aftereffect on stimulus size, and whether the direction of the size change was an important factor. For ease of discussion, these conditions will be labelled $A_{0.5}$ $T_{0.5}$ / $A_{1.5}$ $T_{1.5}$ and $A_{0.5}$ $T_{1.5}$ / $A_{1.5}$ $T_{0.5}$ from this point forward.

6.3.2 Results and discussion

The proportion of "test longer than reference" responses was plotted against the physical test durations and the data was fitted with a logistic curve as per

Chapter 4 of this thesis. Psychometric functions for representative observer CF are shown in Figure 6.14 following adaptation to the 0.5° Gabor (Figure 6.14a) and to the 1.5° Gabor (Figure 6.14b). In both cases the red data points represent the “same” condition (in which size was constant across adapt-test phases) and blue data points represent the “different” condition (in which size was varied between adapt-test phases). For both adapt sizes, there is a greater lateral separation of the psychometric functions for the “same” condition compared to the “different” condition, suggesting that adaptation had a greater distorting effect on the perceived duration of the test stimulus when size was held constant across adapt-test phases.

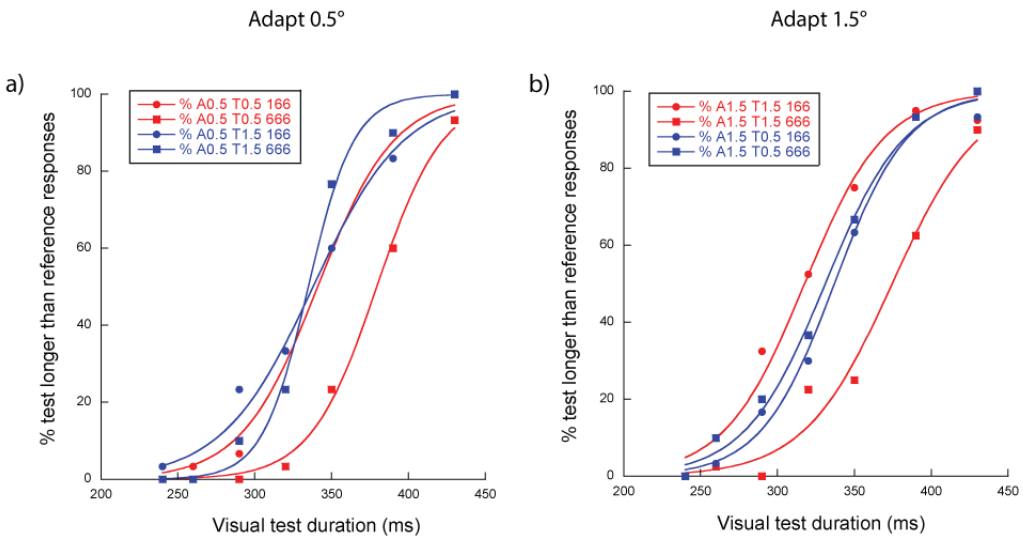

Figure 6.14: Psychometric functions for representative observer CAF for the “same” (red) and “different” (blue) conditions following adaptation to **a)** the 0.5° Gabor and **b)** the 1.5° Gabor.

For all observers, PSE values were extracted for each experimental condition and the duration aftereffect magnitude (DAM) was calculated. Mean DAM values (n=8) are shown in Figure 6.15, where blue data

represents adapting to the 0.5° Gabor and yellow data represents adapting to the 1.5° Gabor. The resulting DAM values in the "same" conditions were $A_{0.5}$ $T_{0.5}$= 49ms and $A_{1.5}$ $T_{1.5}$= 58ms, and in the "different" conditions were $A_{0.5}$ $T_{1.5}$ = 20ms and $A_{1.5}$ $T_{0.5}$= 35ms. A paired samples t-test found no significant difference between the two "same" conditions: $A_{0.5}$ $T_{0.5}$/$A_{1.5}$ $T_{1.5}$

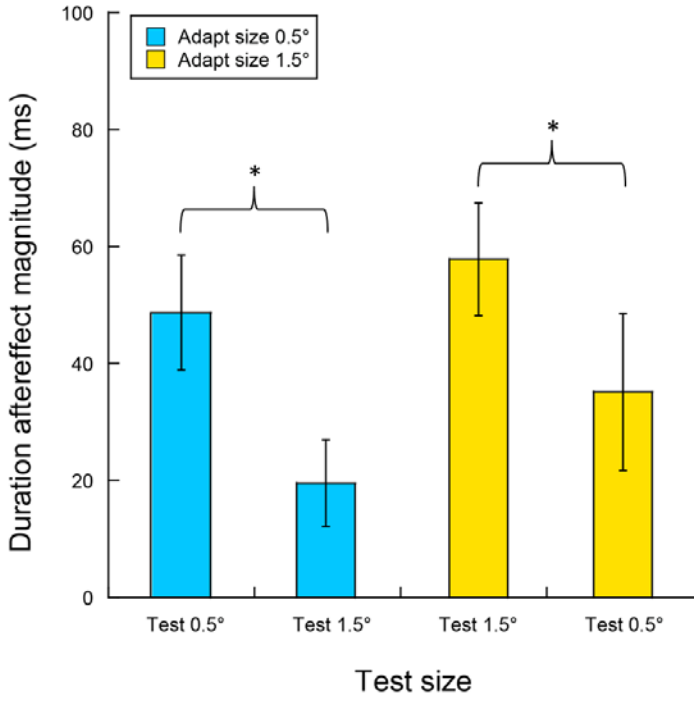

Figure 6.15: Duration aftereffect magnitude (DAM) averaged across observers (n=7) for the Gabor size experiment. Observers adapted to 0.5° (blue data) or 1.5° (yellow data) Gabors in separate experimental blocks. The size of the test Gabor was either the same as the size of the adapting Gabor ("same" conditions) or it varied in size ("different" condition). Asterisks denote significance of $p<0.05$. Error bars represent the SEM.

(p=0.122). The overall strength of the DAE was therefore not determined by the spatial extent of the adapting stimulus. There was also no significant difference between the two "different" conditions: $A_{0.5}$ $T_{1.5}$/$A_{1.5}$ $T_{0.5}$ (p=0.162). However, in both cases there was a significant reduction in DAM when stimulus size was varied between the adapt and test phases, irrespective of the adapt size: $A_{0.5}$ $T_{0.5}$/ $A_{0.5}$ $T_{1.5}$ ($p<0.05$) and $A_{1.5}$ $T_{1.5}$ / $A_{1.5}$ $T_{0.5}$ ($p<0.05$)

(see Figure 6.15 asterisks). This suggests that the mechanism encoding stimulus duration may also be involved in the encoding of global changes in stimulus contrast.

General discussion

The results of Experiment 6.1 demonstrated that repulsive, bidirectional duration aftereffects could be elicited despite gross changes in Gabor carrier orientation between adapt and test phases. This indicates that the neurons responsible for encoding event duration are not selective to luminance defined orientation, a relatively low-level feature. Experiment 6.2 then explored DAE selectivity to facial identity, a visual feature typically ascribed to high-level visual processing. Whilst a trend towards selectivity was noted, this failed to reach statistical significance, suggesting that neurons which are tuned for event duration do not show a preference for facial identity. Finally Experiment 6.3 examined DAE selectivity for stimulus size defined by only by contrast. The results revealed a moderate but significant reduction of the duration aftereffect when stimulus size was altered between the adapt and test phases. This size selectivity could be consistent with a mechanism that encodes both stimulus duration and global changes in contrast by pooling across multiple, first-order inputs to extract envelope size. Notably, selectivity for size was not dependent on the direction of the size change (small → big or big → small). This is likely to reflect the activation of two independent contrast filters, each with a receptive field size optimal for each envelope size.

A possible complication that may have arisen from utilising facial identity as a high-level stimulus in Experiment 6.2 is that some observers may not have perceived the change in facial identity between adapt and test conditions. Despite choosing two faces whose identity differences were grossly above the average discrimination threshold, a small percentage of the population are known to have difficulty in discriminating facial identity (for a review see Corrow et al. 2016). How this developmental anomaly arises is still unclear, as it may reflect a problem in extracting and encoding facial features at a structural level, or it may be a more generalised cognitive problem in *perceiving* or recognising facial features. If normal facial processing is disrupted, such that the brain treats different faces as a single invariant combination of features, this could potentially facilitate transfer of the DAE across different facial identities. Since we did not individually measure identity discrimination thresholds, we cannot rule out the possibility that our results may have been affected by observer difficulties in perceiving facial identity.

Furthermore, although selectivity for envelope size was significant, no adapt-test size change condition saw the aftereffect magnitude fall to zero in Experiment 6.3 (see Figure 6.15). This could have arisen from the use of Gabor stimuli, and may therefore represent a limitation of the present study. Whilst the Gabor does contain some contrast-defined properties whose extraction would involve second-order processes, it also contains first-order properties which remained invariable across conditions (e.g. carrier spatial frequency). It is therefore possible that second-order processing is tied to duration encoding, and could be responsible for the moderate DAE

selectivity to contrast-defined size, but that some transfer of the duration signal was also facilitated through the use of consistent first-order properties across both adapt and test stimuli. However, if an invariant carrier was to explain partial transfer of the DAE across contrast envelope sizes, the aftereffect would need to be partially encoded by first-order mechanisms tuned to the carrier's spatial frequency, but not its orientation (see Experiment 6.1). This is unlikely, given the close correlation between the tuning properties of these two characteristics. However this would require explicit testing of carrier spatial frequency selectivity to exclude this possibility.

In conclusion, the results described in this chapter suggest that the neurons which encode event duration do not show any apparent selectivity to low-level (orientation) or high-level (facial identity) features. However, they do suggest that a more global form of processing such as that performed by second-order, contrast mechanisms could be involved in encoding event duration. To investigate this further, future work should focus on visual stimuli that contain no consistent first-order cues to stimulus duration, such as dynamic luminance noise convolved with a contrast envelope. These stimuli would isolate second-order pathways, reducing potential contamination from first-order signals.

Chapter 7 – The role of discrepancy and cue reliability in audio-visual duration perception

Introduction

The world around us is inherently multisensory, and the brain is tasked with deciding whether the barrage of afferent signals or 'cues' arise from common or separate external sources. The integration of cues arising from a common source has important behavioural advantages such as improved speech interpretation (Sumby and Pollack 1954; Ross et al. 2007), voice recognition (Von Kriegstein and Giraud 2006) and object motion detection (Kim et al. 2012). Even when auditory and visual duration signals are generated simultaneously by a common source, small physical duration discrepancies will arise via differences in the generative motor durations of (e.g.) vocal cord engagement and facial feature movements. For the signal's recipient, these discrepancies are then amplified or reduced by variation in their neural representation within visual and auditory sensory pathways.

An unresolved question is how the brain determines whether these discrepancies are small enough to have arisen despite generation by a common source or do in fact signal multiple, potentially unrelated external sources.

In the spatial domain, dominant theoretical frameworks for understanding multisensory integration contend that discrepant stimulus pairs are resolved in favour of the more reliable signal. In other words, the least variable

information source dominates the multisensory percept (Ernst and Banks 2002; Battaglia et al. 2003; Alais and Burr 2004; Helbig and Ernst 2007; Moro et al. 2014). In the spatial domain, visual positional thresholds can be an order of magnitude lower than their auditory counterparts, leading to the ventriloquist effect: attraction of perceived auditory location towards nearby visual signals (Pick et al. 1969; Welch and Warren 1980; Alais and Burr 2004).

Conversely, auditory temporal sensitivity exceeds visual sensitivity (Grondin and Rousseau 1991; Grondin 1993; Wearden et al. 1998; Grondin et al. 2001; Recanzone 2003; Lapid et al. 2009; Stauffer et al. 2012). Correspondingly, there are multiple examples of visual temporal perception being biased in the direction of concurrently presented audition information such as in the perception of rate (Gebhard and Mowbray 1959; Shipley 1964; Repp and Penel 2002; Recanzone 2003), duration (Klink et al. 2011; Grassi and Pavan 2012; Shi et al. 2013), temporal order/position (Fendrich and Corballis 2001; Morein-Zamir et al. 2003) and numerosity (Shams et al. 2000). However, audition can continue to bias visual temporal perception even when unimodal sensitivities are matched (Ortega et al. 2014), suggesting that unimodal temporal sensitivity alone cannot provide a complete explanation for how sound and vision interact in the temporal domain.

When measuring the relative perceptual dominance of one modality over another, the standard approach has been to introduce small, (typically sub-threshold) physical cross-modal cue discrepancies (Burr et al. 2009a; Chen and Yeh 2009; Bausenhart et al. 2013; Hartcher-O'Brien et al. 2014). In order

to map out the role of discrepancy *per se* it becomes necessary to measure bias across a much greater range of cue discrepancy magnitudes. Where these discrepancies become supra-threshold, the locus of attention must be directed to the 'task relevant' and away from the 'task irrelevant' modalities.

For the perception of auditory and visual temporal rate, robust integration gives way to cue segregation as rate discrepancy exceeds ~5Hz (Roach et al., 2006). For audio-visual duration perception, the role of cue discrepancy is less clear. Intuitively, gross duration discrepancies might be expected to promote cue segregation. Whilst this has been reported for stimulus combinations where auditory duration is ~5x visual duration (Klink et al 2011), integration can persist despite 3x stimulus differences (Romei et al 2011). Comparison across studies is complicated by the way that duration discrepancies are modulated across different baseline duration ranges. These duration discrepancies can be defined in at least two ways: (1) proportional differences (i.e. the ratio of one duration to another), or (2) arithmetic differences (one duration subtracted from the other). The latter could allow cue segregation based on onset and/or offset asynchrony alone whereas the former requires the explicit extraction of, and subsequent comparison between, unimodal durations. An unresolved question is whether integration mechanisms make use of one or both of these cues when formulating a multisensory representation of event duration.

Here this question is addressed by examining audio-visual duration perception under conditions of cue conflict. Observers made unimodal auditory or visual duration discrimination judgements in the presence of task-irrelevant 'distracter' durations presented in the opposite modality. A range of

durations are utilised that simultaneously test integration mechanisms under conditions of differential cue reliability, *proportional* duration discrepancies and *arithmetic* duration discrepancies.

Experiment 7.1

7.1.1 Methods

7.1.1.1 Observers

Seven observers (four naïve) took part in the first experiment. All observers gave their informed, written consent to participate, and had normal or corrected to normal vision and hearing at the time of the experiment.

7.1.1.2 Stimuli and apparatus

All visual stimuli were presented on a gamma-corrected Compaq P1220 CRT monitor with a refresh rate of 100Hz and a resolution of 1280x1024. This was connected to a 2x2.26GHz Quad-core Apple Mac Pro desktop computer running Mac OS 10.6.8. All stimuli were generated using Matlab 7.9.0 (Mathworks, MA) running the Psychtoolbox Extension version 3.0 (Brainard and Pelli, 1997, www.psychtoolbox.org). The physical durations and temporal alignment of all auditory and visual stimuli were verified using a dual-channel oscilloscope. Auditory stimuli were intervals filled with white noise (72 dBs). Visual stimuli were isotropic, luminance defined Gaussian 'blobs' (mean luminance 81cd/m^2) presented to the centre of the screen against a uniform grey background of 37cd/m^2.

7.1.1.3 Procedure

Observers made two alternative, forced choice (2AFC) duration discrimination judgements between a visual reference stimulus and a visual test stimulus (see Figure 7.1a). They were asked to decide ‘which flash had the longer duration?’ and respond via a key press, which subsequently triggered the presentation of the next stimulus pair. The reference stimulus duration was fixed at 320ms duration and presented concurrently with an auditory ‘distracter’ stimulus, aligned by the temporal midpoint (see Figure 7.1a). A symmetric range of distracters was deployed comprising 11 auditory durations spanning 0.25 log units of proportional duration discrepancy either side of the 320ms reference (180, 200, 220, 250, 290, 320, 360, 400, 450, 510, or 570ms). Logarithmic spacing of the distracter values ensured that auditory durations greater than 320ms were proportionally discrepant from the 320ms visual reference stimulus to the same degree as those less than 320ms. Visual test stimuli varied in seven steps around 320ms, and were randomly interleaved within a method of constant stimuli. The presentation order of reference and test stimuli was also randomised.

The interstimulus interval (ISI) between reference and test stimuli was randomly jittered between 750ms and 1250ms. Observers were instructed to ignore auditory stimuli and base their judgments on visual duration information alone. Each experimental block consisted of three repetitions of each auditory distracter duration and lasted approximately thirty minutes. Seven blocks were added together to give a total of twenty one repetitions of each distracter-test stimulus combination.

The resulting psychometric functions were fitted with a logistic function of the form:

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

Where μ represents the Point of Subjective Equality (PSE) corresponding to 50% 'test longer than reference' responses (the physical test duration that produces a perceptual match with the 320ms visual reference stimulus) and θ is an estimate of the observer's duration discrimination threshold (half the difference between the values corresponding of 27% and 73% 'test longer' responses).

7.1.2 Results and discussion

Psychometric functions were constructed for each auditory distracter duration, plotting the proportion of "test longer" responses against the physical visual test durations. From these the PSE was then extracted. A sample psychometric function from a single representative observer is shown in Figure 7.1b, where the 320ms visual reference stimulus was paired with a 180ms auditory distracter. Here, a test duration of 309ms has perceptual equivalence with the 320ms visual reference stimulus.

Figure 7.1c shows PSE values (averaged across observers) for each distracter condition plotted as a function of the physical auditory distracter duration. Short auditory distracters induce attraction-type compression of perceived visual reference duration: perceptual distortion in the same direction as the audio-visual duration discrepancy (see Figure 7.1c light grey

region). The extent of this distortion increases with progressive decreases in auditory distracter duration. This trend persists at the extreme left hand side of the function where distracter duration was 140ms (0.25 log steps) shorter than the reference. When distracter duration was longer than the reference

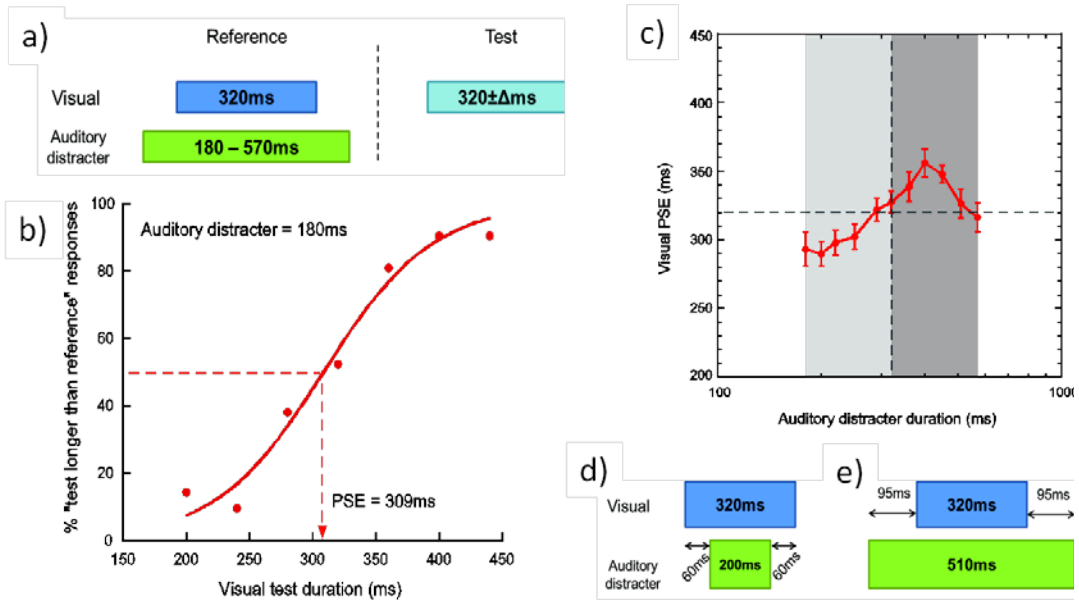

Figure 7.1: a) Schematic showing the experimental task. Observers made 2AFC duration discrimination judgements between a 320ms visual reference stimulus (dark blue rectangle) paired (via its temporal midpoint) with a variable duration auditory distracter (green rectangle – shown to be longer in duration than the reference in the example above), and a variable visual test stimulus (light blue rectangle).**b)** A sample psychometric function for naïve representative observer JS, showing the percentage of "test longer than reference" responses as a function of physical visual test duration. In this example, the 320ms visual reference was perceptually matched to a 308ms visual test stimulus when the former was paired with a 180ms auditory distracter. **c)** Mean tuning function (n=7) from Experiment 7.1 showing perceived visual reference duration across all eleven auditory distracter durations. The vertical dashed line represents the distracter duration which is physically identical to the visual reference stimulus (320ms). Grey regions highlight auditory distracters which are both shorter (light grey) and longer (dark grey) than the visual reference. Error bars represent the SEM. **d & e)** Schematic highlighting how duration discrepancies can remain constant in proportional terms (both the 200ms and 510ms auditory distracters provide 0.2 log steps of duration discrepancy with the 320ms visual reference stimulus) but different in arithmetic terms (200ms is 120ms shorter than 320ms whereas 510ms is 190ms longer than 320ms).

duration (see data points to the right of the dashed vertical line) attractive interaction between the two persists (see Figure 7.1c dark grey region), but now induces expansion of perceived visual duration. Beyond distracter durations of 400ms, attraction declines as PSE returns towards veridical (horizontal dashed line).

Attraction-type interaction (see Figure 7.1c, distracters from 180 – 400ms) is consistent with classical models of sensory cue integration, which evoke differential reliability as a key factor influencing the extent to which each cue contributes towards the integrated multisensory percept. Such attraction-type interaction has been well documented in both the spatial (Ernst and Banks 2002; Battaglia et al. 2003; Alais and Burr 2004; Heron et al. 2004) and temporal (Walker and Scott 1981; Donovan et al. 2004; Roach et al. 2006; Chen and Yeh 2009; Klink et al. 2011; Romei et al. 2011; Sarmiento et al. 2012; Hartcher-O'Brien et al. 2014; Ortega et al. 2014) domains, suggesting that it represents a universal feature of multisensory integration. However it is not clear why this attraction should operate over such an asymmetrical range of duration discrepancies (from log steps of 0.25 below to 0.1 above the reference duration).

In determining the perceptual integration/segregation balance it is possible that the proportional audio-visual duration discrepancy (i.e. reference:distracter duration ratio) is less important than arithmetic differences between the two. This could explain the asymmetry seen in the Figure 7.1c's tuning function: longer distracters will have disproportionately larger onset/offset asynchronies than the corresponding shorter distracters with matched proportional differences (see Figure 7.1d&e).

To answer this question, Experiment 7.2 was designed to examine the role of arithmetic versus proportional differences in governing duration integration. Two new reference durations (160ms and 640ms) were employed with new visual test and distracter ranges centred on these values. The proportionality between reference and distracter durations was identical to Experiment 7.1, but the associated arithmetic duration differences were varied (Figure 7.1d&e).

If the tuning profiles for the 160ms and 640ms ranges form scaled, self-similar versions of the 320ms range's tuning profile (Figure 7.1c) it could be concluded that durational proportionality controls the magnitude of audio-visual duration interaction. If, however arithmetic differences control interaction, a progressive flattening of the tuning profiles might be expected as non-linear increases in arithmetic discrepancy are induced across 160, 320 and 640ms ranges.

Experiment 7.2

7.2.1 Methods

Six observers (three naïve) took part in Experiment 7.2. The procedure and stimuli were identical to the previous experiment, with the exception that the duration of the visual reference stimuli was either 160ms or 640ms. Each of these reference stimuli were coupled with their own range of 11 distracter stimuli and again each of these distracter values varied in 0.05 log steps. This provided duration discrepancies which, in proportional terms, were

symmetrical around the reference duration. The key difference was that that these proportional differences now induced very different arithmetic duration discrepancies.

7.2.2 Results and discussion

For the 160ms visual reference data (Figure 7.2a green data), perceived visual duration is attracted towards the duration signed by the auditory distracter durations. This pattern of attraction takes the form of perceptual

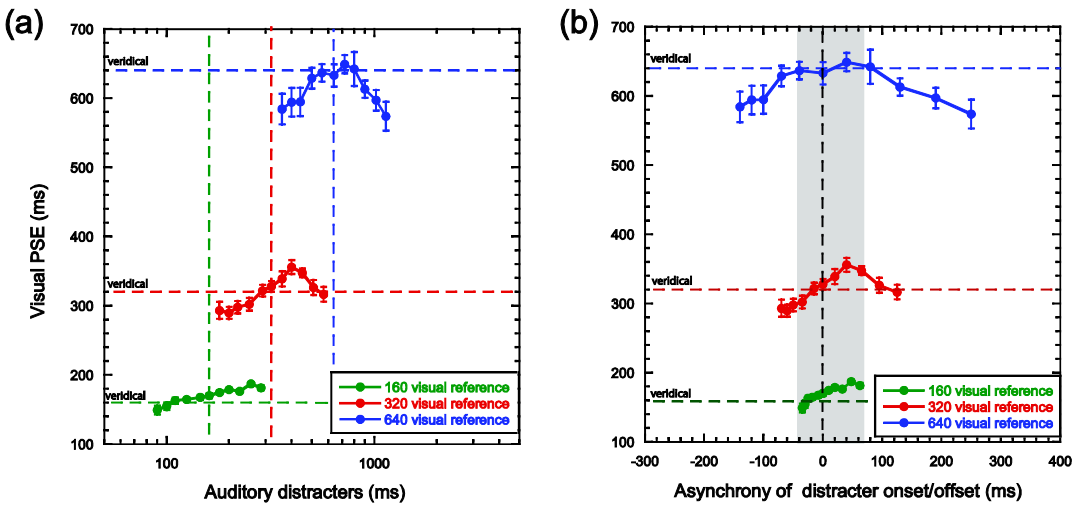

Figure 7.2: Mean tuning functions for the 160ms (green data, n=6), 320ms (red data, n=7) and 640ms (blue data, n=6) visual reference stimuli, showing: **a)** Visual PSE against the physical distracter durations. Vertical dashed lines represent the auditory distracter durations that were physically identical to the visual reference in each range. **b)** Visual PSE plotted in terms of the onset/offset asynchrony between reference and distracter (0.5x the total arithmetic duration difference). Negative/positive values refer to distracters that are shorter/longer than the reference. The black vertical dashed line represents the physical auditory distracter durations that were identical to the visual reference in each range (thus having no asynchrony with the reference). The light grey region highlights the range of arithmetic duration discrepancies which result in attraction-type interaction across all 3 tuning functions. For both plots, horizontal dashed lines represent veridical perception of the visual reference duration and error bars represent the between-observer SEM.

compression/expansion for auditory distracters shorter/longer than 160ms. This shows some parallels with the pattern of interaction observed at the 320ms range (Figure 7.1c and 7.2a (red symbols)) and suggests that duration integration could be dictated by a combination of differential cue reliability and arithmetic duration discrepancies. This scenario is illustrated in Figure 7.2b where the attraction-type interaction persists over the arithmetic difference range (onset or offset asynchronies of -35ms to +60ms) that is common to all three reference durations (Figure 7.2b – grey region), albeit of a smaller magnitude at 640ms (Figure 7.2b, blue tuning function).

One factor that may limit the apparent magnitude of the PSE shifts observed in the plots presented thus far is that of individual differences in perceived visual duration when visual reference and distracter durations were physically equal. For example, when auditory and visual durations were both 640ms, PSEs varied from 584ms - 704ms across observers. The reasons for inter-observer variations in these conditions could reflect the extent to which task irrelevant durations divert attention from the task relevant reference duration. Typically, directing attention away from duration judgments induces compressions of perceived duration (Zakay and Block 1996; Brown 1997). These biases change the overall height of each observer's entire tuning function. Averaging across observers therefore has the effect of compressing group mean functions vertically around their midpoints. In order to isolate the effects of duration discrepancy *per se,* each individual observer's PSE values were divided by their PSE value when distracter and reference durations were physically matched.

This normalisation allows the tuning functions to be collapsed vertically and therefore facilitates comparison of duration discrepancy effect across

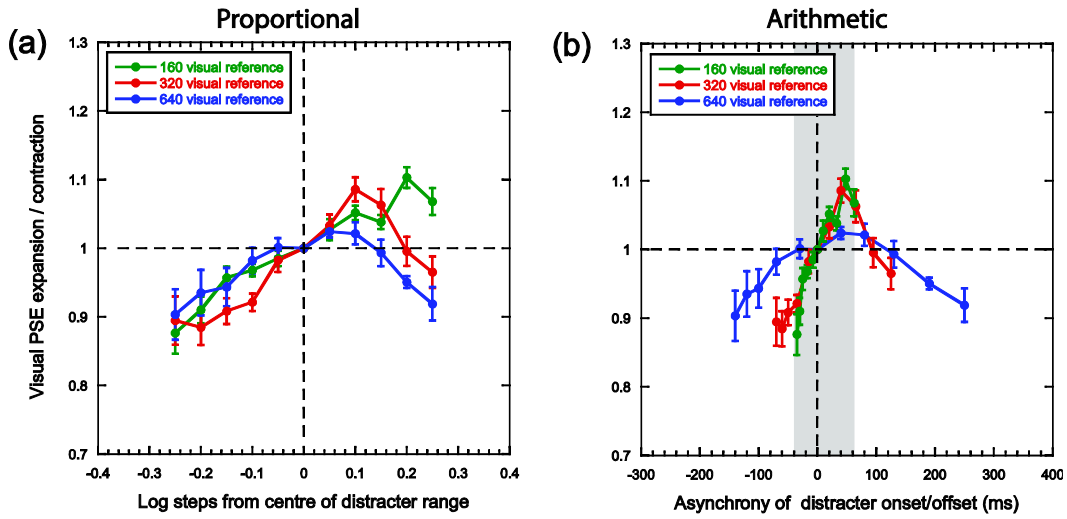

Figure 7.3: Mean tuning functions for the 160ms (green data, n=6), 320ms (red data, n=7) and 640ms (blue data, n=6) ranges, where perceived visual duration is expressed as a ratio of each individual observer's PSE when distracter duration =reference duration **a)** Mean tuning functions plotted as log step differences from the centre of the distracter range. Vertical dashed line represents a physical match between auditory distracter duration and visual reference duration. **b)** Mean ratio tuning functions plotted in terms of distracter onset/offset asynchrony (0.5x the total asynchrony). Negative values are distracters which are shorter than the reference stimulus, and positive values are distracters which are longer than the reference. The onset/offset asynchrony (SOA) demonstrates a much wider arithmetic spread for the 640ms range (blue data) compared to the 320ms range (red data). This in turn has a greater arithmetic spread than the 160ms range (green data). The black vertical dashed line represents the physical auditory distracter duration that was identical to the visual reference in each range (and thus had no onset/offset asynchrony). The light grey region highlights the range of arithmetic duration discrepancies which result in attraction-type interaction across all 3 tuning functions. On both plots, error bars show the SEM.

duration ranges. In addition to this vertical normalisation, all three distracter ranges have now been collapsed horizontally around the centre of each distracter range. This range is now expressed as either proportional (Figure 7.3a) or arithmetic (Figure 7.3b) differences from the visual reference

duration and therefore allows visualisation of the degree to which the tuning functions form scaled, self-similar versions of one another.

If proportional duration differences were critical in determining the shape of these normalised tuning functions they should superimpose on top of one another in Figure 7.3a. Clearly, these functions diverge from one another when distracters are longer than the reference. This lack of symmetry either side of these functions' mid-points makes it unlikely that duration differences *per se* are driving perceptual interaction between reference and distracter durations.

However, when plotted in arithmetic terms, attraction-type interaction is more symmetrical either side of zero and across an asynchrony range common to all three functions (onset or offset asynchronies of -35ms to +60ms, Figure 7.3b - grey region) This commonality suggests that - across this range - the observed (attraction-type) interaction could be a product of relative cue reliability (auditory dominance over vision) and arithmetic duration differences (attraction increases with increasing discrepancy). Despite normalisation, the magnitude of auditory interaction with 640ms visual reference durations is notably weaker than its 160ms and 320ms reference duration counterparts. Attraction persists for distracters outside this range but only for distracters increasingly shorter than the reference (Figure 7.3b - leftwards extremes of the red and blue tuning functions). Conversely, attraction gives way to repulsion for longer distracters (Figure 7.3b - rightwards extremes of the red and blue tuning functions): progressive *increases* in distracter duration cause a progressive *decrease* in perceived visual reference duration.

Reliability-based cue integration frameworks are compatible with attraction-type interaction but do not offer any obvious explanation as to why grossly longer auditory distracters should cause repulsion of perceived visual duration away from the distracter duration. This raises the possibility that interaction around the 640ms reference duration may not be driven by differential cue reliability. In addition, either attraction or repulsion-type interaction persists across the entire distracter range (Figure 7.3a&b). We wondered if a transition to cue segregation could be induced by extending the distracter range to provide larger onset/offset asynchronies and whether cue interaction patterns are similar when the relative reliability of reference and distracter durations is reversed.

Experiment 7.3

7.3.1 Methods

Four observers (two naïve) took part in Experiment 7.3. The stimuli and procedures were identical to Experiment 7.2 with the following exceptions: the reference duration was 640ms throughout and the range of distracter durations was extended from 320ms – 2040ms (sampled more coarsely in 0.1 log steps), providing 9 distracter durations. Observers performed two separate versions of the task: a visual duration discrimination task with auditory distracters and an auditory duration discrimination task with visual distracters (the order of tasks was randomly split between observers, and modality of the reference stimulus was maintained in separate blocks). In all cases observers were instructed to ignore the distracter durations.

7.3.2 Results and discussion

For visual discrimination judgements made alongside auditory distracters (Figure 7.4, blue data), the pattern of attraction and repulsion seen around 640ms in Experiment 7.2 is replicated. When auditory judgements are made alongside visual distracters (Figure 7.4a red data) the tuning profile is markedly similar: both relatively short and long visual distracters induce perceptual compression of auditory duration. This could reflect similar duration discrimination sensitivities for visual and auditory judgments around the 640ms reference stimulus. However, examination of these threshold values (Figure 7.4b) shows that they conform to the widely-reported pattern of relatively higher auditory duration discrimination sensitivity (Grondin and

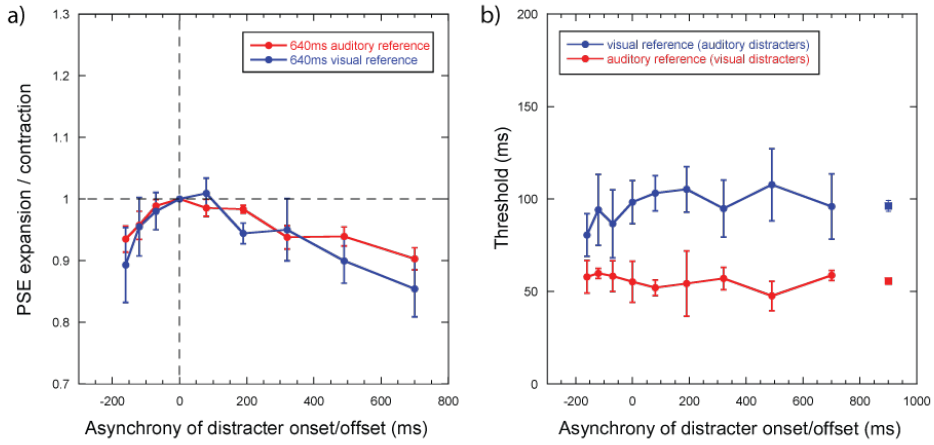

Figure 7.4: a) Normalised mean tuning functions (n=4) for the visual (blue data points) and auditory (red data points) tasks, plotted as a function of distracter onset/offset asynchrony (0.5x the total asynchrony). Negative values refer to distracter durations which are physically shorter than the reference stimulus, and positive values refer to distracter durations which are physically longer than the reference. The vertical dashed line represents a distracter onset/offset asynchrony of 0ms. **b)** Thresholds (averaged across observers) from the visual and auditory task's data sets shown in **(a)**. Square data points represent thresholds averaged across distracter durations. In both plots, error bars represent the SEM.

Rousseau 1991; Mattes and Ulrich 1998; Ortega et al. 2009; Stauffer et al. 2012).

Within sensory reliability frameworks, this sensitivity imbalance predicts grossly reduced, or non-existent (Chen and Yeh 2009; Klink et al. 2011; Bausenhart et al. 2013) biases in perceived auditory duration. Figure 7.4a also shows that extending the range of reference-distracter duration discrepancies failed to restore veridical perception of either visual or auditory reference durations. It therefore seems unlikely that gross cue discrepancy and/or differential cue reliability can explain duration interactions across auditory and visual domains at the 640ms range.

An alternative explanation is provided by the possibility of interactions between attention and the asynchrony levels induced by large arithmetic duration discrepancies. In the current study, observers were instructed to direct their attention towards the reference and test durations alone. However, the asynchrony between the relative temporal locations of the reference and distracter stimuli's onsets and offsets provide highly salient temporal discontinuities of the type known to be effective in capturing exogenous attention within and between sensory modalities (Spence et al. 2001; Talsma et al. 2010). In attentional terms, these transients are most effective when they contain the abrupt (i.e. square wave) temporal profiles provided by the distracter onsets/offsets in the present series of experiments (Van der Burg et al. 2008; Van der Burg et al. 2010; Kösem and Van Wassenhove 2012).

As mentioned earlier, directing attention towards/away from a stimulus typically induces perceptual overestimation/underestimation of its duration. Early demonstrations of these effects were provided by studies employing cognitively demanding tasks alongside the discrimination of supra-second intervals (Zakay and Block 1996; Brown 1997). More recently, the pattern of results has been widened to include attentional-modulation of sub-second intervals, both within dual-task (Chen and O'Nell 2001; Rammsayer and Ulrich 2005; Casini et al. 2009; Cicchini and Morrone 2009) and exogenous cueing situations (Mattes and Ulrich 1998; Yeshurun and Marom 2008; Osugi et al. 2016). Regarding the latter, if attention is diverted away from the spatial location of an upcoming stimulus by 'invalid' cues its perceived duration is compressed. If a temporal analogue is responsible for the perceptual compression seen in Figure 7.4a, onset/offset asynchrony would need be of sufficiently supra-threshold magnitude to efficiently divert attention from the reference stimulus towards asynchrony itself. The resulting attentional deficit would affect the perceived duration of the reference but not the test stimulus. This effect would reduce PSE values, as per Figure 7.4a. By this logic, when asynchrony is sub-threshold *and* distracter stimuli are more temporally reliable than reference stimuli, perceived duration is attracted towards the distracter. This proposition is broadly compatible with the pattern of interaction depicted within the shaded region of Figure 7.3b. Conversely, when asynchrony is sub-threshold *and* distracters are *less* temporally reliable than reference durations both relative cue-reliability and attentional mechanisms predict reduced or absent distracter-reference interaction.

This hypothesis was tested in a further experiment where a range of visual distracters were coupled with 320ms auditory stimuli. The visual distracter stimuli's durations were chosen to create a range of relatively small onset/offset asynchrony values (from -70ms to 125ms).

Experiment 7.4

7.4.1 Methods

Seven observers (3 naïve) took part in Experiment 7.4. The stimuli and procedure were the same as Experiment 7.1, except that observers made auditory duration discrimination judgements in the presence of visual distracter stimuli (centred on a 320ms reference duration).

7.4.2 Results and discussion

Results are shown in Figure 7.5 (red data) alongside visual judgements with auditory distracters (replotted from Figure 7.3b, blue data). Clearly, there is no systematic pattern of interaction between auditory reference and visual distracter when onset/offset asynchrony magnitude is limited to -70ms to +125ms (see Figure 7.5). This is in sharp contrast to the pattern observed in Experiment 7.1 where the same range of asynchronies induced attraction-type interaction towards auditory distracters (see Figure 7.5 blue data, replotted from Figure 7.3b). Clearly, without sufficient asynchrony to activate alternative mechanisms, the relative temporal reliabilities of reference and distracter dominate duration perception.

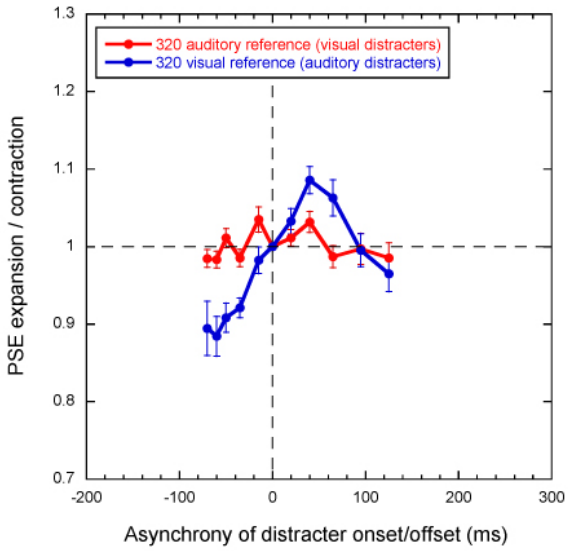

Figure 7.5: Mean normalised tuning functions for auditory duration judgements with visual distracters (n=7, red data) alongside visual judgments with auditory distracters (n = 7, blue data, replotted from Figure 7.3b). For both data sets, the reference judgment was centred on 320ms. The amount of PSE expansion and contraction is plotted in terms of distracter onset/offset asynchrony (0.5*total asynchrony). Negative/positive values refer to distracters that are shorter/longer than the reference. Error bars represent the SEM.

General discussion

The current study has shown that perceived visual duration can be attracted towards or repelled from the duration signalled by concurrent auditory durations. Specifically, when the arithmetic discrepancy between these durations is small *and* relative temporal sensitivity favours the distracter duration, visual duration was found to be distorted in the direction of audition. Over the range where attraction was consistent, the pattern was approximately linear and, when plotted in terms of arithmetic differences, showed notable overlap across duration ranges (Figure 7.3b).

When duration discrepancy was large, a counterintuitive pattern of interaction was found: both grossly shorter *and* longer distracters induced

robust perceptual compression of visual duration. Critically, this pattern persists (Experiment 7.3) even when the temporal reliability imbalance between reference and distracter durations makes the opposite prediction. This strongly suggests that the mechanism underpinning duration interaction at the extremes of the present asynchrony ranges has a supramodal processing locus. Finally, the results of Experiment 7.4 underscore the interaction between onset/offset asynchrony magnitude and temporal reliability. Duration perception was found to be veridical when asynchrony was limited *and* distracter stimuli were less temporally reliable that reference stimuli.

Taken together, the findings of the current study suggest two - potentially distinct - mechanisms operating in duration integration. One is suggested to be driven by low-level sensory characteristics, the other by the global, supra-modal dynamics of attention. The former dominates when small duration discrepancies induce onset/offset asynchronies likely to be sub-threshold on a majority of trials. Under these conditions, relative temporal reliability is key: distracter durations only induce attraction-type interaction when imbued with superior temporal reliability (e.g. Figure 7.5). This finding is in broad agreement with literature examining the relative perceptual weighting of auditory and visual duration cues when discrepant duration combinations generate small arithmetic audio-visual discrepancies of approximately <100ms (Donovan et al. 2004; Burr et al. 2009a; Romei et al. 2011; Bausenhart et al. 2013; Hartcher-O'Brien et al. 2014). In the current study evidence is presented of an attractive-type bias across an onset/offset asynchrony range of ~ ±100ms (Figure 7.3b). The mechanism driving

attraction within this range could take at least two forms. The perceived onsets and/or offsets of the reference stimulus could be distorted towards their distracter counterparts, a process of 'temporal ventriloquism' (Scheier et al. 1999; Morein-Zamir et al. 2003; Vroomen and de Gelder 2004; Burr et al. 2009a). Alternatively, integration could be from a weighted averaging process that transpires subsequent to the extraction of unimodal duration estimates (Hartcher-O'Brien et al. 2014). Although both mechanisms are compatible with the reference-distracter interaction observed across small asynchronies, the latter should induce interaction tuning profiles that are symmetrical when plotted in terms of duration discrepancy. Comparison of Figures 7.3a & b suggests that interaction between onsets and offsets better predicts the relationship between small magnitude cue discrepancy and audio-visual duration integration.

The literature documenting onset/offset asynchrony detection thresholds for filled durations quotes values ranging from ~ 50-200ms, depending on stimulus characteristics (Vatakis and Spence 2006; Stekelenburg and Vroomen 2007; Van Wassenhove et al. 2007; Boenke et al. 2009; Vroomen and Stekelenburg 2011; Kuling et al. 2012). Although asynchrony detection thresholds were not measured directly, it is likely that the instruction to ignore distracter stimuli (as opposed to making explicit judgments about the temporal order of reference-distracter onset/offset) elevated each observer's effective asynchrony detection thresholds away from the lower end of this range. Arguably, the range over which attraction-type interaction persists is narrower for positive asynchronies where the distracter's presentation starts before *and* finishes after the reference duration's presentation (e.g. Figure

7.5a – blue data). This could reflect heightened sensitivity to offset asynchrony. However, that trend is not a trend supported by the relevant literature (e.g. Boenke et al. 2009; Kuling et al. 2012). An alternative interpretation is that the perceptual expansion of perceived duration (~ 0 to +100ms) represents the interaction between two opposing forces: one driving sensory, attraction-type expansion and the other driving attention-induced perceptual compression. In this scenario, the former dominates up to the expansionary peak at ~ +40ms but is curtailed by the latter as asynchrony further increases. When distracter presentation starts after *and* finishes before the reference duration's presentation compression is found for all levels of (-ve) asynchrony. This could be driven by a single attraction-type mechanism. However, this is thought to be unlikely for the following reasons. First, low-level interaction between other forms of auditory and visual temporal signals dissipates as the magnitude of the discrepancy between them becomes grossly supra-threshold (Roach et al. 2006; Kuling et al. 2013). Second, it is unclear why this mechanism would continue to operate when temporal reliability makes the opposite prediction, but only when asynchrony is large (Figure 7.4a – red data).

Perhaps the pattern of attraction-type interaction could be explained by a form of response bias where observers adopt a strategy of making a discrimination judgment between distracter and test stimulus durations? If they behaved in this way on a small number of trials (e.g. when reference-test duration differences were close to their discrimination threshold) PSE would be biased in the direction of the distracter. However, in this scenario the measured psychometric function would in fact form an amalgam between

two functions: one centred on the reference-test PSE, the other centred on the distracter-test PSE. Merging these functions leads to flattening of the combined function's slope. If this strategy is to explain PSE over the range where attraction-type interaction might operate (-140 to +100ms in Figure 7.3b) threshold elevation should be noted which tracks the extent of any attraction-type PSE shifts over the same asynchrony range. Threshold data presented in Figure 7.6 does not support this conclusion: introducing distracter-reference discrepancy does not lead to any systematic pattern of threshold elevation. In addition, it is unclear why observers would adopt this strategy only when asynchrony was < ~+100ms.

However, it is conceivable this strategy (a form of response bias) may explain other examples of duration interaction with large audio-visual duration discrepancies. For example, Klink et al. (2011) found expansive/compressive attraction-type bias of visual duration by concurrently presented auditory durations. This bias increased approximately linearly up to ±300ms (the maximum used in their study). In that study, test and reference visual durations were of identical duration, thus incentivising observers to make (much easier) comparisons between auditory distracter and visual test. Their design did not allow independent measures of visual bias or duration discrimination sensitivity, making it difficult to rule out this strategy as an explanation for attraction-type bias with large duration discrepancies. The same issue limits comparison with other related studies of audio-visual duration integration where the perceptual metrics reported are non-specific measures of visual performance such as *d'* or '% correct' (Donovan et al. 2004; Romei et al. 2011).

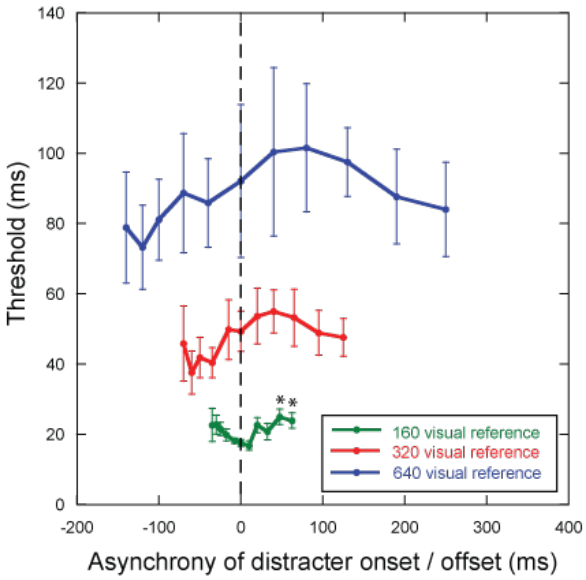

Figure 7.6: Mean visual duration discrimination threshold data for the 160ms range (green data, n=6), 320ms range (red data, n=7) and 640ms range (blue data, n=6) from Experiment 7.1 and 7.2 plotted as a function of distracter onset/offset asynchrony. The vertical dashed line represents a distracter onset/offset asynchrony of 0ms. Negative values refer to distracter durations which are physically shorter than the reference stimulus, and positive values refer to distracter durations which are physically longer than the reference. Black asterisks denote threshold values which were significantly different ($p<0.05$) from the threshold when distracter = reference (for each range) Error bars represent the SEM.

Relatedly, Ortega and colleagues (2014) found that single interval discrimination of visual durations showed compressive/expansive perceptual bias towards ±300ms shorter/longer concurrently presented auditory durations. In this scenario, observers reporting auditory - rather than visual - duration on some trials would shift PSE in the direction of the auditory stimulus – a pattern consistent with their finding that *"all bimodal stimuli were perceived to be about the same duration as the auditory-alone stimulus. JNDs for the bimodal stimuli are also more similar to the JND for the auditory-alone stimulus than to the JND for the visual-alone stimulus"*.

The results of Experiments 7.3 and 7.4 (Figure 7.4a and 7.5) suggest that a common, supramodal mechanism drives the perceptual compression observed when large duration discrepancies induce correspondingly large magnitude asynchrony. The literature documenting interaction between attention and duration suggests that the relative temporal distance between the transients associated with the distracter's onset and the reference duration's onset represents an effective vehicle for capturing exogenous attention, thereby diverting endogenous attention from the reference stimulus (Spence et al. 2001; Talsma et al. 2010). This attentional deficit could be responsible for compression of perceived reference stimulus duration. If so, it should be possible to amplify this effect via experimental manipulations known to modulate attentional control. For example, expectation about the temporal structure of upcoming events can strongly influence the allocation of attention across these events (Barnes and Jones 2000; Jones et al. 2002; Miller et al. 2013). At long duration ranges (e.g. 640ms – Figure 7.4a) the interleaving of distracter durations also had the effect of interleaving grossly supra-threshold asynchronies of random magnitude and polarity. Anecdotally, observers described this temporal structure as a highly distracting series of transients with temporally unpredictable onsets.

The contribution of temporal irregularity to compression of the reference duration was tested by increasing the temporal jitter of the distracter-reference onsets and offsets whilst keeping the total asynchrony (i.e. arithmetic duration difference) constant. To this end, part of Experiment 7.2 was repeated with five observers (three naïve) using only the 640ms visual reference, and a more coarsely sampled range of auditory distracter stimuli.

The key difference with Experiment 7.2 was that trials were now equally divided between centre aligned distracters (as per Experiment 7.2) and two 'catch trial' conditions where distracters were aligned with the reference by either their onset or offset. This created additional temporal uncertainty about the trial-to-trial variation in reference-distracter temporal structure: two thirds of trials contained either no onset or offset asynchrony (but double the opposite polarity) compared to their centre-aligned counterparts of identical distracter duration and total asynchrony. If this new temporal structure

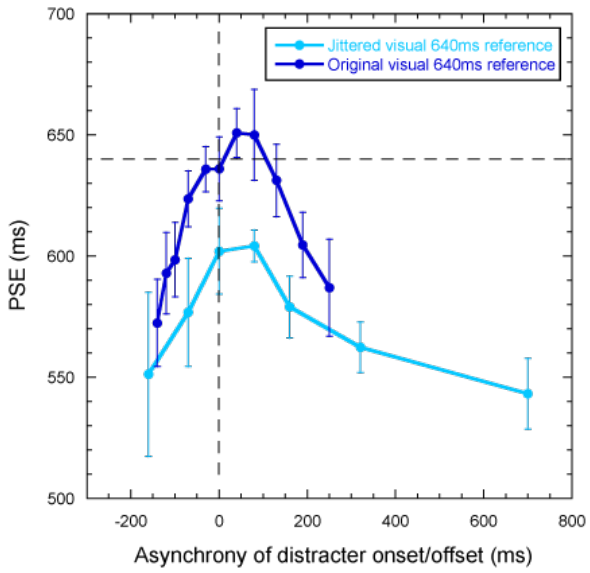

Figure 7.7: Mean tuning functions for the new 'jittered', centre-aligned condition (light blue data, n=5) against the 640ms visual reference data replotted from Figure 7.2b (data blue data, n=6). Perceived duration of the visual reference is shown in terms of the onset/offset asynchrony between reference and distracter (0.5x the total arithmetic duration difference). Error bars represent the SEM.

increases the trial-to-trial difficulty of attending to the reference duration, an attention-based mechanism would further compress its perceived duration.

The new (jittered) centre-aligned results are shown in Figure 7.7 (light blue data) alongside those for the original (blocked, centre-aligned) 640ms range

data replotted from Figure 7.2b (dark blue data). Increasing the salience of the distracter's onsets/offsets by increasing their trial-to-trial temporal jitter induces a reduction in PSE for all distracter durations. This PSE reduction represents further perceptual compression of the reference stimulus duration. Importantly, the total asynchrony and temporal structure of the onset/offset is identical for both of Figure 7.7's data sets. The only difference is the temporal context in which they were presented. It is perhaps notable that the degree of additional compression appears approximately uniform across asynchrony levels, suggesting the effect of the temporal jitter was distributed globally across trials.

This result further suggests a key role for interaction between attention and large magnitude asynchrony in driving temporal compression. A dependence on attentional allocation would have parallels with other reports of an 'attentional bottleneck' in the processing of event duration (Morgan et al. 2008; Brown 2010; Ruthruff and Pashler 2010) and audio-visual asynchrony (Fujisaki et al. 2005).

How might attention mediate temporal compression? It has been argued that within a pacemaker-accumulator setting (Creelman 1962; Treisman 1963), reducing selective attention 'narrows the attentional gate' between pacemaker and accumulator (Zakay and Block 1995). If this restricts the number of 'pulses' passed from the former to the latter, estimates of temporal extent decrease accordingly. Although the neurophysiological basis of a pacemaker, accumulator or attentional gate remains elusive, this framework could provide a conceptual explanation for the compressive effects reported here.

Alternatively, an attentional deficit could lower the amplitude of the stimulus' neural representation via a process of response gain reduction (Reynolds and Heeger 2009). If perceived duration scales with the magnitude of neural activity associated with a stimulus (Eagleman and Pariyadath 2009), attention-dependent reductions in this magnitude would therefore reduce perceived duration. A similar neural modulation could distort duration perception within 'duration channels' (Walker et al. 1981; Ivry 1996; Becker and Rasmussen 2007; Heron et al. 2012) by selectively manipulating response gain within channels or by shifting channel tuning preferences (David et al. 2008). Either of these manipulations would provide a neural pathway for attention to skew the population response (across channels) towards shorter durations.

In summary, the interaction between conflicting auditory and visual duration signals has been shown to form a complex pattern that varies with relative temporal reliability and duration discrepancy. When duration discrepancy is small and temporal reliability favours task-irrelevant distracter stimuli, perceived duration follows the pattern of attraction predicted by current sensory reliability-based models of cue integration. However, this pattern breaks down when duration discrepancy increases in a way that introduces large onset/offset asynchronies. Under these conditions, reference stimuli were compressed by distracter stimuli, irrespective of whether the distracters signalled longer/shorter durations or were imbued with higher/lower temporal reliability. Whilst the reasons underlying this compression remain unclear, follow up experiments suggest that the transient nature of abrupt onsets and offsets is an effective manipulator of attentional focus. When this focus is

diverted from the reference to the transients themselves, the resultant attentional deficit leads to temporal compression. The inhomogeneity of interaction patterns across duration discrepancies contrasts markedly with the spatial tuning of the ‘ventriloquist effect’ (Howard and Templeton 1966; Bertelson and Radeau 1981) which suggests that there may be fundamental differences between the way the brain combines sensory information across time and space.

Chapter 8: Conclusions

Temporal estimates are formed continually, and often without conscious awareness. The ability to accurately process durations in the sub-second range is particularly important for interpreting speech, co-ordinating movement and decision making. Our perception of time is further mediated by cognitive factors such as attention, and may be biased by recent sensory history.

The work presented in this thesis adds to the growing body of research into the mechanisms subserving human time perception. Whilst early models posit that duration is processed by an amodal, centralised clock (Creelman 1962; Treisman 1963; Gibbon and Church 1984; Gibbon et al. 1984), evidence presented here challenges this view in several important ways and favours localised, sensory specific temporal encoding.

Chapters 4 – 6 focused on an interaction between recent sensory history, perceived duration and visual stimulus characteristics. Specifically, adaptation was used to generate duration aftereffects, which were used to probe the mechanism(s) underlying duration encoding. By generating duration aftereffects and then manipulating the similarity of adapting and test stimulus characteristics, it was possible to ascertain the duration encoding mechanism's selectivity for these characteristics. Using the visual system's known hierarchical structure (Hubel and Wiesel 1968; Gross et al. 1972; Zeki 1973; Desimone et al. 1984; Felleman and Van Essen 1991; Grill-Spector and Malach 2004) a range of characteristics were chosen which attempted to selectively target different processing stages within this hierarchy.

Chapter 4 used the most well documented variation in selectivity throughout the visual system: tuning for spatial location (Hubel and Wiesel 1968; Gattass et al. 1981; Gattass et al. 1988; Blatt et al. 1990; Xu et al. 2001; Dumoulin and Wandell 2008). Duration aftereffects were found to spread across a relatively large region of space which was approximately 5x larger than the adapting stimulus. This spatial spread could be consistent with duration encoding by relatively late stage visual filtering underpinned by receptive fields with a large, fixed diameter. This possibility was ruled out by follow up experiments that showed a proportional relationship between stimulus size and aftereffect spread: as stimulus size increases, the region into which the aftereffect spreads increases by the same (~5x) multiple.

The proportionality is evocative of visual spatial mechanisms which pool input across multiple, proportionally smaller filters. These spatially abutting inputs tile a region of space which form the receptive field of the second (pooled) stage mechanism. Within the visual system this has strong parallels with the processing relationship between first-order, luminance selective mechanisms and their second-order, contrast/texture selective counterparts (Bergen 1991; Sutter et al. 1995; Zhou and Baker 1996; Westrick and Landy 2013).

Chapter 5 then utilised another feature that is closely tied to a mechanism's position within the visual processing hierarchy: the degree to which visual mechanisms are binocular and selective for retinal disparity. To this end two complementary approaches were used. The first of these was the degree to which duration aftereffects transferred between adapted and non-adapted eyes. Aftereffects showed robust (79%) transfer, indicating that the

mechanism responsible for generating the aftereffect must receive substantial input from both eyes. The second approach employed a novel method for delivering temporal information, where stimulus duration was defined by retinal disparity in the absence of any monocular cues. Again, robust duration aftereffects were observed. This finding confirms that the duration adaptation mechanism must be binocular, and that it resides at or beyond the level of neurons tuned for retinal disparity. This necessitates a cortical locus which, given the scale dependent tuning observed in the preceding chapter, is potentially of an extrastriate nature.

These findings suggest that visual duration is unlikely to be encoded at the very earliest stages of visual processing. This hypothesis was tested in Chapter 6 where luminance defined visual orientation was varied across adapt and test phases. Duration aftereffects showed no selectivity for this stimulus feature. Although this finding could reflect a sub-cortical (pre-orientation selectivity) basis (Johnston et al. 2006), the broad spatial tuning described in Chapter 4 and the binocularity described in Chapter 5 make this unlikely. One visual mechanism displaying coarse selectivity for spatial location is that underlying the processing of facial identity (Leopold et al. 2001; Afraz and Cavanagh 2008). This coarse spatial selectivity is thought to arise via pooling across multiple low-level stimulus characteristics. Might duration be a similarly late stage stimulus characteristic? Although a trend towards selectivity for adapt-test changes in facial identity was noted, this failed to reach statistical significance. This could reflect duration encoding at a later stage than facial identity processing but is perhaps more likely to reflect earlier encoding which is tied to stimulus feature(s) which are then

subsequently pooled by facial identity processing mechanisms. To test this latter possibility we introduced adapt-test changes in stimulus size where size itself was defined by spatial variations in Gabor contrast envelope. Detecting these changes requires a mechanism that can pool the output of multiple luminance detecting filters tuned to the characteristics of the Gabor's carrier. By comparing across these filters, contrast envelope can be extracted. Although aftereffects did transfer across a 1º increase/decrease change in contrast envelope size, a significant reduction in aftereffect magnitude was associated with this size change. Taken together the results of Chapter 6 reveal selectivity for contrast-defined but not luminance-defined stimulus characteristics. This finding is consistent with the broad, scale-dependent tuning for spatial location documented in Chapter 4.

Chapter 7 then went on to examine the strategies governing multisensory integration of duration signals. A series of experiments focused on the perceptual distortions of duration that arise when audio-visual cues to duration are placed in conflict. When small duration discrepancies are introduced between visual reference stimuli and concurrently presented (but task irrelevant) auditory stimuli, perceived visual duration is attracted towards that signalled by audition. This pattern of interaction is broadly consistent with literature examining spatial and temporal interactions under conditions of small (sub-threshold) cue conflict: multisensory judgments are biased towards the modality with the higher precision (Ernst and Banks 2002; Alais and Burr 2004; Burr et al. 2009a; Hartcher-O'Brien et al. 2014). However, this pattern of interaction was not symmetrical: increases in distracter duration curtailed the attraction-type interaction more rapidly than decreases.

Yet, when plotted in terms of onset/offset asynchrony, a greater degree of commonality was evident for interaction patterns across the 160, 320 and 640ms ranges.

When reference-distracter discrepancy is relatively large, perceived duration of the reference is compressed, irrespective of whether the distracter durations were longer or shorter than the reference and irrespective of relative sensory reliability. This pattern of interaction cannot be explained by relative duration differences (which remained proportionally constant across longer/shorter distracter ranges) or relative sensory reliability (which predicts only attraction type integration or complete cue segregation). These characteristics suggest that when onset/offset asynchrony is sufficiently large a global, supramodal influence drives perceived duration via a central mechanism. When the predictability of large-magnitude onsets/offsets was reduced by increasing their trial-to-trial temporal jitter, duration underestimation increased across the range of distracters. It is argued that this mechanism is underpinned by the asynchrony between the distracter-reference onsets/offsets capturing exogenous attention. Increasing the magnitude of their asynchrony increases their salience thereby increasing their ability to divert attention from the reference stimulus duration. If onset/offset asynchrony is responsible for an attentional deficit, this could be consistent with the wide body of literature reporting duration underestimation under conditions of divided attention (e.g. Macar et al. 1994; Brown 1997; Casini and Macar 1997; Cicchini and Morrone 2009).

If cross-modal attention can explain the interaction between grossly discrepant auditory and visual durations, would more extreme effects be seen when the two durations share within-modality attentional resources? Unimodal dual-task studies dividing attention across multiple visual spatial locations show severe performance deficits in temporal discrimination tasks (Morgan et al. 2008; Cheng et al. 2014). However it remains unclear whether perceived temporal extent is underestimated in a similar manner to that reported here for audio-visual interaction.

In summary, the findings of this thesis' adaption experiments are consistent with local, sensory specific duration encoding mechanisms. The encoding mechanism shows selectivity for stimulus features which have a cortical locus. This could form a parallel mechanism, distinct from other local duration encoding mechanisms argued to have a basis in neural energy readout (Pariyadath and Eagleman 2007; Eagleman and Pariyadath 2009) or the temporal properties of the magnocellular (Johnston et al. 2008; Ayhan et al. 2009; Bruno et al. 2011) (or extrastriate motion, (Burr et al. 2007; Fornaciai et al. 2016)) pathways. Alternatively, duration could have parallels with visual motion: a visual characteristic encoded at multiple visual scales which form a cascade of interacting processing stages.

Whilst encoding by multiple, interacting local mechanisms remain a possibility, a unitary centralised mechanism is not consistent with the sensory specificity documented elsewhere (Walker et al. 1981; Becker and Rasmussen 2007; Heron et al. 2012; Li et al. 2015b) or any of the stimulus specificity reported here. In this context, the robust cross-modal interaction between conflicting audio-visual durations is unlikely to be supported by

attention-dependent modulation of a centralised mechanism such as a pacemaker-accumulator. A more likely explanation is that the putative influence of attention is driven by a mechanism operating at a stage subsequent to the encoding of auditory and visual durations by local, sensory-specific mechanisms within unimodal brain areas.

References

Aaen-Stockdale, C., Hotchkiss, J., Heron, J. and Whitaker, D. (2011) Perceived time is spatial frequency dependent. *Vision Research*, 51 (11), 1232-1238.

Adams, D. L., Sincich, L. C. and Horton, J. C. (2007) Complete pattern of ocular dominance columns in human primary visual cortex. *The Journal of Neuroscience,* 27 (39), 10391-10403.

Afraz, S. R. and Cavanagh, P. (2008) Retinotopy of the face aftereffect. *Vision Research,* 48 (1), 42-54.

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

Aldrich, M. S., Alessi, A. G., Beck, R. W. and Gilman, S. (1987) Cortical blindness: etiology, diagnosis, and prognosis. *Annals of Neurology,* 21 (2), 149-158.

Allan, L. G. (1984) Contingent aftereffects in duration judgments. *Annals of the New York Academy of Sciences* 423 (1), 116-130.

Allan, L. G. (2002) The location and interpretation of the bisection point. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology,* 55 (1), 43-60.

Allan, L. G. and Gibbon, J. (1991) Human bisection at the geometric mean. *Learning and Motivation,* 22 (1-2), 39-58.

Allan, L. G., Siegel, S., Toppan, P. and Lockhead, G. R. (1991) Assessment of the McCollough effect by a shift in psychometric function. *Bulletin of the Psychonomic Society,* 29 (1), 21-24.

Allman, J. M. and Kaas, J. H. (1974) The organization of the second visual area (V II) in the owl monkey: a second order transformation of the visual hemifield. *Brain Research,* 76 (2), 247-265.

Allman, M. J. and Meck, W. H. (2012) Pathophysiological distortions in time perception and timed performance. *Brain,* 135 (3), 656-677.

Allman, M. J., Teki, S., Griffiths, T. D. and Meck, W. H. (2014) Properties of the internal clock: first-and second-order principles of subjective time. *Annual Review of Psychology,* 65, 743-771.

Angrilli, A., Cherubini, P., Pavese, A. and Manfredini, S. (1997) The influence of affective factors on time perception. *Perception & Psychophysics,* 59 (6), 972-982.

Artieda, J., Pastor, M. A., Lacruz, F. and Obeso, J. A. (1992) Temporal discrimination is abnormal in Parkinson's disease. *Brain,* 115 (1), 199-210.

Aubie, B., Sayegh, R. and Faure, P. A. (2012) Duration tuning across vertebrates. *Journal of Neuroscience,* 32 (18), 6373 - 6390.

Ayhan, I., Bruno, A., Nishida, S. and Johnston, A. (2009) The spatial tuning of adaptation-based time compression. *Journal of Vision,* 9 (11), doi: 10.1167/9.11.2.

Ayhan, I., Bruno, A., Nishida, S. Y. and Johnston, A. (2011) Effect of the luminance signal on adaptation-based time compression. *Journal of Vision,* 11 (7), doi: 10.1167/11.7.22.

Baccus, S. A. and Meister, M. (2002) Fast and slow contrast adaptation in retinal circuitry. *Neuron,* 36 (5), 909-919.
Baker, C. L. and Mareschal, I. (2001) Processing of second-order stimuli in the visual cortex. *Progress in Brain Research,* 134, 171-191.
Ball, D. M., Arnold, D. H. and Yarrow, K. (2017) Weighted integration suggests that visual and tactile signals provide independent estimates about duration. *Journal of Experimental Psychology: Human Perception and Performance,* 43 (5), 868-880.
Bangert, A. S., Reuter-Lorenz, P. A. and Seidler, R. D. (2011) Dissecting the clock: Understanding the mechanisms of timing across tasks and temporal intervals. *Acta Psychologica,* 136 (1), 20-34.
Barlow, H. B., Blakemore, C. and Pettigrew, J. D. (1967) The neural mechanism of binocular depth perception. *The Journal of Physiology,* 193 (2), 327-342.
Barnes, R. and Jones, M. R. (2000) Expectancy, attention, and time. *Cognitive Psychology,* 41 (3), 254-311.
Bartolo, R. and Merchant, H. (2009) Learning and generalization of time production in humans: rules of transfer across modalities and interval durations. *Experimental Brain Research,* 197 (1), 91-100.
Battaglia, P. W., Jacobs, R. A. and 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 (7), 1391-1397.
Bausenhart, K. M., de la Rosa, M. D. and Ulrich, R. (2013) Multimodal integration of time: visual and auditory contributions to perceived duration and sensitivity. *Experimental Psychology,* 61, doi:10.1027/1618-3169/a000249.
Becker, M. W. and Rasmussen, I. P. (2007) The rhythm aftereffect: support for time sensitive neurons with broad overlapping tuning curves. *Brain and Cognition,* 64 (3), 274-281.
Ben Hamed, S., Duhamel, J.R., Bremmer, F. and Graf, W. (2001) Representation of the visual field in the lateral intraparietal area of macaque monkeys: a quantitative receptive field analysis. *Experimental Brain Research,* 140 (2), 127-144.
Bergen, J. R. (1991) Theories of visual texture perception. In Regan, D. (editor) *Vision and Visual Dysfunction.* Vol. 10B. New York: Macmillan. 114-134.
Berson, D. M., Dunn, F. A. and Takao, M. (2002) Phototransduction by retinal ganglion cells that set the circadian clock. *Science,* 295 (5557), 1070-1073.
Bertelson, P. and Radeau, M. (1981) Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Attention, Perception and Psychophysics,* 29 (6), 578-584.
Bjarnason, G. A., Jordan, R. C. K. and Sothern, R. B. (1999) Circadian variation in the expression of cell-cycle proteins in human oral epithelium. *American Journal of Pathology,* 154 (2), 613-622.
Blakemore, C. and Campbell, F. W. (1969) On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology,* 203 (1), 237-260.

Blakemore, C. and Sutton, P. (1969) Size adaptation a new aftereffect. *Science,* 166 (3902), 245-247.
Blasdel, G. G. and Fitzpatrick, D. (1984) Physiological organization of layer 4 in macaque striate cortex. *Journal of Neuroscience,* 4 (3), 880-895.
Blatt, G. J., Andersen, R. A. and Stoner, G. R. (1990) Visual receptive field organization and cortico-cortical connections of the lateral intraparietal area (area LIP) in the macaque. *Journal of Comparative Neurology,* 299 (4), 421-445.
Block, R. A. and Zakay, D. (1996) Models of psychological time revisited. In H., H. (editor) *Time and Mind.* Kirkland, WA: Hogrefe and Huber. 171-195.
Block, R. A. and Zakay, D. (1997) Prospective and retrospective duration judgments: a meta-analytic review. *Psychonomic Bulletin & Review,* 4 (2), 184-197.
Bobko, D. J., Schiffman, H. R., Castino, R. J. and Chiappetta, W. (1977) Contextual effects in duration experience. *The American Journal of Psychology*, 90 (4), 577-586.
Boenke, L. T., Deliano, M. and Ohl, F. W. (2009) Stimulus duration influences perceived simultaneity in audiovisual temporal-order judgment. *Experimental Brain Research,* 198 (2-3), 233-244.
Brainard, D. H. (1997) The psychophysics toolbox. *Spatial Vision,* 10 (4), 433-436.
Brand, A., Urban, R. and Grothe, B. (2000) Duration tuning in the mouse auditory midbrain. *Journal of Neurophysiology,* 84 (4), 1790-1799.
Bratzke, D., Seifried, T. and Ulrich, R. (2012) Perceptual learning in temporal discrimination: asymmetric cross-modal transfer from audition to vision. *Experimental Brain Research,* 221 (2), 205-210.
Brigner, W. L. (1986) Effect of perceived brightness on perceived time. *Perceptual and Motor Skills,* 63 (2), 427-430.
Brooks, V. B. and Thach, W. T. (1981) Cerebellar control of posture and movement. *Handbook of Physiology, The Nervous System, Motor Control*, 877-946.
Brown, S. W. (1985) Time perception and attention - the effects of prospective versus retrospective paradigms and task demands on perceived duration. *Perception & Psychophysics,* 38 (2), 115-124.
Brown, S. W. (1997) Attentional resources in timing: interference effects in concurrent temporal and nontemporal working memory tasks. *Perception & Psychophysics,* 59 (7), 1118-1140.
Brown, S. W. (2006) Timing and executive function: Bidirectional interference between concurrent temporal production and randomization tasks. *Memory & Cognition,* 34 (7), 1464-1471.
Brown, S. W. (2010) Timing, resources and interference: attentional modulation of time perception. In Nobre, A. C. and Coull, J. T. (editors) *Attention and Time.* Oxford: Oxford University Press. 107-122.
Brown, S. W. and Merchant, S. M. (2007) Processing resources in timing and sequencing tasks. *Perception & Psychophysics,* 69 (3), 439-449.
Brown, S. W. and West, A. N. (1990) Multiple timing and the allocation of attention. *Acta Psychologica,* 75 (2), 103-121.

Bruce, C., Desimone, R. and Gross, C. G. (1981) Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology,* 46 (2), 369-384.
Bruno, A., Ayhan, I. and Johnston, A. (2010) Retinotopic adaptation-based visual duration compression. *Journal of Vision,* 10 (10), doi:10.1167/10.10.30.
Bruno, A., Ayhan, I. and Johnston, A. (2011) Duration expansion at low luminance levels. *Journal of Vision,* 11 (14), doi:10.1167/11.14.13.
Bryce, D. and Bratzke, D. (2015) Multiple timing of nested intervals: further evidence for a weighted sum of segments account. *Psychonomic Bulletin and Review,* 23 (1), 317-323.
Bryce, D., Seifried-Dübon, T. and Bratzke, D. (2015) How are overlapping time intervals perceived? Evidence for a weighted sum of segments model. *Acta Psychologica,* 156, 83-95.
Bueti, D., Walsh, V., Frith, C. and Rees, G. (2008) Different brain circuits underlie motor and perceptual representations of temporal intervals. *Journal of Cognitive Neuroscience,* 20 (2), 204-214.
Buhusi, C. V. and Meck, W. H. (2005) What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews: Neuroscience,* 6 (10), 755-765.
Buhusi, C. V. and Oprisan, S. A. (2013) Time-scale invariance as an emergent property in a perceptron with realistic, noisy neurons. *Behavioural Processes,* 95, 60-70.
Buonomano, D. V. (2000) Decoding temporal information: a model based on short-term synaptic plasticity. *The Journal of Neuroscience,* 20 (3), 1129-1141.
Buonomano, D. V. and Karmarkar, U. R. (2002) How do we tell time? *Neuroscientist,* 8 (1), 42-51.
Buonomano, D. V. and Maass, W. (2009) State-dependent computations: spatiotemporal processing in cortical networks. *Nature Reviews Neuroscience,* 10 (2), 113-125.
Buonomano, D. V. and Merzenich, M. M. (1995) Temporal information transformed into a spatial code by a neural network with realistic properties. *Science,* 267 (5200), 1028-1030.
Burkhalter, A. and Van Essen, D. C. (1986) Processing of color, form and disparity information in visual areas VP and V2 of ventral extrastriate cortex in the macaque monkey. *The Journal of Neuroscience,* 6 (8), 2327-2351.
Burr, D., Banks, M. S. and Morrone, M. C. (2009a) Auditory dominance over vision in the perception of interval duration. *Experimental Brain Research,* 198 (1), 49-57.
Burr, D., Silva, O., Cicchini, G. M., Banks, M. S. and Morrone, M. C. (2009b) Temporal mechanisms of multimodal binding. *Proceedings of the Royal Society B,* 276 (1663), 1761-1769.
Burr, D., Tozzi, A. and Morrone, M. C. (2007) Neural mechanisms for timing visual events are spatially selective in real-world coordinates. *Nature Neuroscience,* 10 (4), 423-425.
Burr, D. C., Cicchini, G. M., Arrighi, R. and Morrone, M. C. (2011) Spatiotopic selectivity of adaptation-based compression of event duration. *Journal of Vision,* 11 (2), doi:10.1167/11.2.21.

Cahoon, D. and Edmonds, E. M. (1980) The watched pot still won't boil - expectancy as a variable in estimating the passage of time. *Bulletin of the Psychonomic Society,* 16 (2), 115-116.
Cai, Z. G. and Wang, R. (2014) Numerical magnitude affects temporal memories but not time encoding. *PLoS ONE,* 9 (1), e83159.
Calvert, G. A. (2001) Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cerebral Cortex,* 11 (12), 1110-1123.
Campbell, F. W. and Maffei, L. (1971) The tilt after-effect: a fresh look. *Vision Research,* 11 (8), 833-840.
Casini, L., Burle, B. and Nguyen, N. (2009) Speech perception engages a general timer: Evidence from a divided attention word identification task. *Cognition,* 112 (2), 318-322.
Casini, L. and Ivry, R. B. (1999) Effects of divided attention on temporal processing in patients with lesions of the cerebellum or frontal lobe. *Neuropsychology,* 13 (1), 10-21.
Casini, L. and Macar, F. (1997) Effects of attention manipulation on judgments of duration and of intensity in the visual modality. *Memory and Cognition,* 25 (6), 812-818.
Casseday, J. H. and Covey, E. (1992) Frequency tuning properties of neurons in the inferior colliculus of an FM bat. *Journal of Comparative Neurology,* 319 (1), 34-50.
Casseday, J. H., Ehrlich, D. and Covey, E. (1994) Neural tuning for sound duration: role of inhibitory mechanisms in the inferior colliculus. *Science,* 264, 847-850.
Cavanagh, P. and Mather, G. (1989) Motion: the long and short of it. *Spatial Vision,* 4 (2), 103-129.
Chen, G.-D. (1998) Effects of stimulus duration on responses of neurons in the chinchilla inferior colliculus. *Hearing Research,* 122, 142-150.
Chen, K.-M. and Yeh, S.-L. (2009) Asymmetric cross-modal effects in time perception. *Acta Psychologica,* 130 (3), 225-234.
Chen, Z. and O'Nell, P. (2001) Processing demand modulates the effects of spatial attention on the judged duration of a brief stimulus. *Attention, Perception and Psychophysics,* 63 (7), 1229-1238.
Cheng, X., Yang, Q., Han, Y., Ding, X. and Fan, Z. (2014) Capacity limit of simultaneous temporal processing: how many concurrent 'clocks' in vision? *PloS ONE,* 9 (3), e91797.
Cheong, S. K., Tailby, C., Solomon, S. G. and Martin, P. R. (2013) Cortical-like receptive fields in the lateral geniculate nucleus of marmoset monkeys. *Journal of Neuroscience,* 33 (16), 6864-6876.
Chiba, A., Oshio, K.-i. and Inase, M. (2008) Striatal neurons encode temporal information in duration discrimination task. *Experimental Brain Research,* 186 (4), 671-676.
Chubb, C. and Sperling, G. (1988) Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A,* 5 (11), 1986-2007.
Church, R. M. and Deluty, M. Z. (1977) Bisection of temporal intervals. *Journal of Experimental Psychology-Animal Behavior Processes,* 3 (3), 216-228.

Church, R. M. and Gibbon, J. (1982) Temporal Generalization. *Journal of Experimental Psychology-Animal Behavior Processes,* 8 (2), 165-186.
Cicchini, G. M. and Morrone, M. C. (2009) Shifts in spatial attention affect the perceived duration of events. *Journal of Vision,* 9 (1), doi:10.1167/9.1.9
Cohen, J., Hansel, C. E. M. and Sylvester, J. D. (1953) A new phenomenon in time judgment. *Nature,* 172, 901.
Corrow, S. L., Dalrymple, K. A. and Barton, J. J. S. (2016) Prosopagnosia: current perspectives. *Eye and Brain,* 8, 165 - 175.
Coslett, H. B., Wiener, M. and Chatterjee, A. (2010) Dissociable neural systems for timing: evidence from subjects with basal ganglia lesions. *PLoS ONE,* 5 (4), e10324.
Coull, J. and Nobre, A. (2008) Dissociating explicit timing from temporal expectation with fMRI. *Current Opinion in Neurobiology,* 18 (2), 137-144.
Coull, J. T., Cheng, R. K. and Meck, W. H. (2011) Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacology,* 36 (1), 3-25.
Coull, J. T., Nazarian, B. and Vidal, F. (2008) Timing, storage, and comparison of stimulus duration engage discrete anatomical components of a perceptual timing network. *Journal of Cognitive Neuroscience,* 20 (12), 2185-2197.
Coull, J. T., Vidal, F., Nazarian, B. and Macar, F. (2004) Functional anatomy of the attentional modulation of time estimation. *Science,* 303 (5663), 1506-1508.
Cowan, R. L. and Wilson, C. J. (1994) Spontaneous firing patterns and axonal projections of single corticostriatal neurons in the rat medial agranular cortex. *Journal of Neurophysiology,* 71 (1), 17-32.
Cox, D. D. (2014) Do we understand high-level vision? *Current Opinion in Neurobiology,* 25, 187-193.
Creelman, C. D. (1962) Human discrimination of auditory duration. *The Journal of the Acoustical Society of America,* 34 (5), 582-593.
Curran, W. and Benton, C. P. (2012) The many directions of time. *Cognition,* 122 (2), 252-257.
Curran, W., Benton, C. P., Harris, J. M., Hibbard, P. B. and Beattie, L. (2016) Adapting to time: duration channels do not mediate human time perception. *Journal of Vision,* 16 (5), doi:10.1167/16.5.4
David, S. V., Hayden, B.Y., Mazer, J. A. and Gallant, J. L. (2008) Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. *Neuron,* 59 (3), 509-521.
De Valois, R. L., Albrecht, D. G. and Thorell, L. G. (1982a) Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research,* 22 (5), 545-559.
De Valois, R. L., Yund, E. W. and Hepler, N. (1982b) The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research,* 22 (5), 531-544.
DeAngelis, G. C. and Newsome, W. T. (1999) Organization of disparity selective neurons in macaque area MT. *The Journal of Neuroscience,* 19 (4), 1398-1415.

Desimone, R., Albright, T.D., Gross, C. G. and Bruce, C. (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience,* 4 (8), 2051-2062.
Desimone, R. and Gross, C. G. (1979) Visual areas in the temporal cortex of the macaque. *Brain Research,* 178 (2), 363-380.
Desimone, R. and Schein, S. J. (1987) Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *Journal of Neurophysiology,* 57 (3), 835-868.
Dhruv, N. T. and Carandini, M. (2014) Cascaded effects of spatial adaptation in the early visual system. *Neuron,* 81 (3), 529-535.
Dhruv, N. T., Tailby, C., Sokol, S. H. and Lennie, P. (2011) Multiple adaptable mechanisms early in the primate visual pathway. *Journal of Neuroscience,* 31 (42), 15016-15025.
Di Chiara, G., Morelli, M. and Consolo, S. (1994) Modulatory functions of neurotransmitters in the striatum: ACh/dopamine/NMDA interactions. *Trends in Neurosciences,* 17 (6), 228-233.
Dickinson, J. E., Almeida, R.A., Bell, J. and Badcock, D. R. (2010) Global shape aftereffects have a local substrate: A tilt aftereffect field. *Journal of Vision,* 10 (13), doi:10.1167/10.13.5.
Dickinson, J. E. and Badcock, D. R. (2013) On the hierarchical inheritance of aftereffects in the visual system. *Frontiers in Psychology,* 4, doi:10.3389/fpsyg.2013.00472.
Dickinson, J. E., Mighall, H.K., Almeida, R. A., Bell, J. and Badcock, D. R. (2012) Rapidly acquired shape and face aftereffects are retinotopic and local in origin. *Vision Research,* 65, 1 - 11.
Dolores de la Rosa, M. and Bausenhart, K. (2013) Multimodal integration of interval duration: Temporal ventriloquism or changes in pacemaker rate? *Timing & Time Perception,* 1 (2), 189-215.
Dong, P. and Wyer, R. S. (2014) How time flies: the effects of conversation characteristics and partner attractiveness on duration judgments in a social interaction. *Journal of Experimental Social Psychology,* 50, 1-14.
Donovan, C. L., Lindsay, D. S. and Kingstone, A. (2004) Flexible and abstract resolutions to crossmodal conflicts. *Brain and Cognition,* 56 (1), 1-4.
Dormal, V., Seron, X. and Pesenti, M. (2006) Numerosity-duration interference: A Stroop experiment. *Acta Psychologica,* 121 (2), 109-124.
Driver, J. and Noesselt, T. (2008) Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron,* 57 (1), 11-23.
Droit-Volet, S., Brunot, S. and Niedenthal, P. (2004) Perception of the duration of emotional events. *Cognition and Emotion,* 18 (6), 849-858.
Dumoulin, S. O. and Wandell, B. A. (2008) Population receptive field estimates in human visual cortex. *Neuroimage,* 39 (2), 647-660.
Durstewitz, D. (2003) Self organizing neural integrator predicts interval times through climbing activity. *The Journal of Neuroscience,* 23 (12), 5342-5353.
Duysens, J., Schaafsma, S. J. and Orban, G. A. (1996) Cortical off response tuning for stimulus duration. *Vision Research,* 36 (20), 3243-3251.

Eagleman, D. M. and Pariyadath, V. (2009) Is subjective duration a signature of coding efficiency? *Philosophical Transactions of the Royal Society B: Biological Sciences,* 364 (1525), 1841-1851.
Ellemberg, D., Allen, H.A. and Hess, R. F. (2004) Investigating local network interactions underlying first-and second-order processing. *Vision Research,* 44 (15), 1787-1797.
Ernst, M. O. and Banks, M. S. (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature,* 415 (6870), 429-433.
Espinosa-Fernandez, L., Vacas, L. D., Garcia-Viedma, M. D., Garcia-Gutierrez, A. and Colmenero, C. J. T. (2004) Temporal performance in 4-8 year old children. The effect of chronometric information in task execution. *Acta Psychologica,* 117 (3), 295-312.
Fahy, F. L., Riches, I. P. and Brown, M. W. (1993) Neuronal activity related to visual recognition memory: long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Experimental Brain Research,* 96 (3), 457-472.
Faisal, A. A. and Wolpert, D. M. (2009) Near optimal combination of sensory and motor uncertainty in time during a naturalistic perception-action task. *Journal of Neurophysiology,* 101 (4), 1901-1912.
Falchier, A., Clavagnier, S., Barone, P. and Kennedy, H. (2002) Anatomical evidence of multimodal integration in primate striate cortex. *Journal of Neuroscience,* 22 (13), 5749-5759.
Faure, P. A., Fremouw, T., Casseday, J. H. and Covey, E. (2003) Temporal masking reveals properties of sound-evoked inhibition in duration-tuned neurons of the inferior colliculus. *Journal of Neuroscience,* 23 (7), 3052-3065.
Favreau, O. E. (1976) Motion aftereffects - evidence for parallel processing in motion perception. *Vision Research,* 16 (2), 181-186.
Fayolle, S., Gil, S. and Droit-Volet, S. (2015) Fear and time: fear speeds up the internal clock. *Behavioural Processes,* 120, 135-140.
Fechner, G. (1860) *Elemente dur psychophysik.* Leipzig, Germany: Breitkopf and Härtel.
Felleman, D. J. and Van Essen, D., C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex,* 1 (1), 1-47.
Felleman, D. J. and Vanessen, D. C. (1987) Receptive-field properties of neurons in area V3 of macaque monkey extrastriate cortex. *Journal of Neurophysiology,* 57 (4), 889-920.
Fendrich, R. and Corballis, P. M. (2001) The temporal cross-capture of audition and vision. *Perception & Psychophysics,* 63 (4), 719-725.
Ferrandez, A.-M., Hugueville, L., Lehéricy, S., Poline, J.-B., Marsault, C. and Pouthas, V. (2003) Basal ganglia and supplementary motor area subtend duration perception: an fMRI study. *Neuroimage,* 19 (4), 1532-1544.
Ferrara, A., Lejeune, H. and Wearden, J. H. (1997) Changing sensitivity to duration in human scalar timing: An experiment, a review, and some possible explanations. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology,* 50 (3), 217-237.

Fierro, B., Palermo, A., Puma, A., Francolini, M., Panetta, M. L., Daniele, O., Brighina and F. (2007) Role of the cerebellum in time perception: a TMS study in normal subjects. *Journal of the Neurological Sciences,* 263 (1), 107-112.

Fornaciai, M., Arrighi, R. and Burr, D. C. (2016) Adaptation-induced compression of event time occurs only for translational motion. *Scientific Reports,* 6:23341, doi:10.1038/srep23341.

Foster, K. H., Gaska, J. P., Nagler, M. and Pollen, D. A. (1985) Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *The Journal of Physiology,* 365 (1), 331-363.

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

Freeman, J. and Simoncelli, E. P. (2011) Metamers of the ventral stream. *Nature Neuroscience,* 14 (9), 1195-1201.

Frisby, J. P. (1980) *Seeing: Illusion, Brain and Mind.* Oxford: Oxford University Press.

Fujisaki, W., Nishida and S.Y. (2005) Temporal frequency characteristics of synchrony–asynchrony discrimination of audio-visual signals. *Experimental Brain Research,* 166 (3-4), 455-464.

Galazyuk, A. V. and 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,* 180 (4), 301-311.

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

Garcia-Perez, M. A. (2014) Does time ever fly or slow down? The difficult interpretation of psychophysical data on time perception. *Frontiers in Human Neuroscience,* 8, doi:10.3389/fnhum.2014.00415.

Gattass, R., Gross, C. G. and Sandell, J. H. (1981) Visual topography of V2 in the macaque. *The Journal of Comparative Neurology,* 201 (4), 519-539.

Gattass, R., Sousa, A. P. B. and Gross, C. G. (1988) Visuotopic organization and extent of V3 and V4 of the macaque. *The Journal of Neuroscience,* 8 (6), 1831-1845.

Gebhard, J. and Mowbray, G. (1959) On discriminating the rate of visual flicker and auditory flutter. *The American Journal of Psychology,* 72 (4), 521-529.

Gescheider, G. A. (1976) *Psychophysics: method and theory*. Hillside, N. J.: L. Erlbaum Associates

Getty, D. J. (1975) Discrimination of short temporal intervals - comparison of 2 models. *Perception & Psychophysics,* 18 (1), 1-8.

Gheorghiu, E., Kingdom, F. A. A., Thai, M. T. and Sampasivam, L. (2009) Binocular properties of curvature-encoding mechanisms revealed through two shape after-effects. *Vision Research,* 49 (14), 1765-1774.

Gibbon, J. (1971) Scalar timing and semi-Markov chains in free-operant avoidance. *Journal of Mathematical Psychology,* 8 (1), 109-138.

Gibbon, J. (1972) Timing and discrimination of shock density in avoidance. *Psychological Review,* 79 (1), 68-92

Gibbon, J. (1977) Scalar expectancy theory and Weber's law in animal timing. *Psychological Review,* 84 (3), 279-325
Gibbon, J. (1991) Origins of scalar timing. *Learning and Motivation,* 22 (1), 3-38.
Gibbon, J. (1992) Ubiquity of scalar timing with a Poisson clock. *Journal of Mathematical Psychology,* 36 (2), 283-293.
Gibbon, J. and Church, R. M. (1984) Sources of variance in an information processing theory of timing. In Roitblat, H. L., Bever, T. G., and Terrace, H. S. (editors) *Animal cognition.* Vol. 1. Hillsdale, NJ: Erlbaum. 465 - 487.
Gibbon, J., Church, R. M. and Meck, W. H. (1984) Scalar timing in memory. *Annals of the New York Academy of Sciences,* 423 (1), 52-77.
Gibbon, J., Malapani, C., Dale, C. L. and Gallistel, C. (1997) Toward a neurobiology of temporal cognition: advances and challenges. *Current Opinion in Neurobiology,* 7 (2), 170-184.
Gibson, J. J. (1933) Adaptation, after-effect and contrast in the perception of curved lines. *Journal of Experimental Psychology,* 16 (1), 1-31.
Gibson, J. J. and Radner, M. (1937) Adaptation, after-effect and contrast in the perception of tilted lines. I. Quantitative studies. *Journal of Experimental Psychology,* 20 (5), 453-467
Gil, S. and Droit-Volet, S. (2011) “Time flies in the presence of angry faces”… depending on the temporal task used! *Acta Psychologica,* 136 (3), 354-362.
Goldstone, S., Boardman, W. K. and Lhamon, W. T. (1959) Intersensory comparisons of temporal judgments. *Journal of Experimental Psychology,* 57 (4), 243-248.
Goldstone, S. and Goldfarb, J. L. (1963) Judgment of filled and unfilled durations: intersensory factors. *Perceptual and Motor Skills,* 17 (3), 763-774.
Goldstone, S. and Lhamon, W. T. (1974) Studies of auditory - visual differences in human time judgement. 1. Sounds are judged longer than lights. *Perceptual and Motor Skills,* 39 (1), 63-82.
Goldstone, S., Lhamon, W. T. and Sechzer, J. (1978) Light intensity and judged duration. *Bulletin of the Psychonomic Society,* 12 (1), 83-84.
Gorea, A. (2011) Ticks per thought or thoughts per tick? A selective review of time perception with hints on future research. *Journal of Physiology,* 105 (4), 153-163.
Graham, N. V. (2011) Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): useful additions of the last 25 years. *Vision Research,* 51 (13), 1397-1430.
Grahn, J. A. and Brett, M. (2009) Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex,* 45 (1), 54-61.
Grantham, D. W. and Wightman, F. L. (1979) Auditory motion aftereffects. *Perception & Psychophysics,* 26 (5), 403-408.
Grassi, M. and Pavan, A. (2012) The subjective duration of audiovisual looming and receding stimuli. *Attention, Perception & Psychophysics,* 74 (6), 1321-1333.
Green, D. M. and Swets, J. A. (1966) *Signal detection theory and psychophysics.* Vol. 1., New York: Wiley

Grill-Spector, K., Henson, R. and Martin, A. (2006) Repetition and the brain: neural models of stimulus specific effects. *Trends in Cognitive Sciences,* 10 (1), 14-23.
Grill-Spector, K. and Malach, R. (2004) The human visual cortex. *Annual Review of Neuroscience,* 27, 649-677.
Grondin, S. (1993) Duration discrimination of empty and filled intervals marked by auditory and visual signals. *Perception & Psychophysics,* 54 (3), 383-394.
Grondin, S. (2010) Unequal Weber fractions for the categorization of brief temporal intervals. *Attention, Perception, & Psychophysics,* 72 (5), 1422-1430.
Grondin, S. (2012) Violation of the scalar property for time perception between 1 and 2 seconds: evidence from interval discrimination, reproduction, and categorization. *Journal of Experimental Psychology: Human Perception and Performance,* 38 (4), 880-890
Grondin, S., Meilleur-Wells, G. and Lachance, R. (1999) When to start explicit counting in a time-intervals discrimination task: a critical point in the timing process of humans. *Journal of Experimental Psychology: Human Perception and Performance,* 25 (4), 993-1004.
Grondin, S., Meilleur-Wells, G., Ouellette, C. and Macar, F. (1998) Sensory effects on judgments of short time intervals. *Psychological Research,* 61 (4), 261-268.
Grondin, S., Ouellet, B. and Roussel, M. (2004) Benefits and limits of explicit counting for discriminating temporal intervals. *Canadian Journal of Experimental Psychology,* 58 (1), 1-12.
Grondin, S., Ouellet, B. and Roussel, M.-È. (2001) About optimal timing and stability of Weber fraction for duration discrimination. *Acoustical Science and Technology,* 22 (5), 370-372.
Grondin, S. and Rousseau, R. (1991) Judging the relative duration of multimodal short empty time intervals. *Perception & Psychophysics,* 49 (3), 245-256.
Gross, C. G., Rochamir.Ce and Bender, D. B. (1972) Visual properties of neurons in inferotemporal cortex of macaque. *Journal of Neurophysiology,* 35 (1), 96-111.
Halsband, U., Ito, N., Tanji, J. and Freund, H. J. (1993) The role of premotor cortex and the supplementary motor area in the temporal control of movement in man. *Brain,* 116 (1), 243-266.
Hardy, N. and Buonomano, D. (2016) Neurocomputational models of interval and pattern timing. *Current Biology,* 8, 250-257.
Hari, R. and Salmelin, R. (1997) Human cortical oscillations: a neuromagnetic view through the skull. *Trends in Neurosciences,* 20 (1), 44-49.
Harrington, D. L., Boyd, L.A., Mayer, A. R., Sheltraw, D. M., Lee, R. R., Huang, M. and Rao, S. M. (2004) Neural representation of interval encoding and decision making. *Cognitive Brain Research,* 21 (2), 193-205.
Harrington, D. L., Haaland, K. Y. and Hermanowitz, N. (1998a) Temporal processing in the basal ganglia. *Neuropsychology,* 12 (1), 3-12

Harrington, D. L., Haaland, K. Y. and Knight, R. T. (1998b) Cortical networks underlying mechanisms of time perception. *Journal of Neuroscience,* 18 (3), 1085-1095.
Harrington, D. L., Zimbelman, J. L., Hinton, S. C. and Rao, S. M. (2010) Neural modulation of temporal encoding, maintenance, and decision processes. *Cerebral Cortex,* 20, 1274-1285.
Hartcher-O'Brien, J., Di Luca, M. and Ernst, M. O. (2014) The duration of uncertain times: audiovisual information about intervals is integrated in a statistically optimal fashion. *PloS ONE,* 9 (3), e89339.
Haxby, J. V. H., E.A.Gobbini, M.I. (2000) The distributed human neural system for face perception. *Trends in Cognitive Sciences,* 4 (6), 223-233.
Hayashi, M. J., Ditye, T., Harada, T., Hashiguchi, M., Sadato, N., Carlson, S., Walsh, V. and Kanai, R. (2015) Time adaptation shows duration selectivity in the human parietal cortex. *PLOS Biology,* 13 (9), e1002262.
He, J. F., Hashikawa, T., Ojima, H. and Kinouchi, Y. (1997) Temporal integration and duration tuning in the dorsal zone of cat auditory cortex. *Journal of Neuroscience,* 17 (7), 2615-2625.
Helbig, H. B. and Ernst, M. O. (2007) Optimal integration of shape information from vision and touch. *Experimental Brain Research,* 179 (4), 595-606.
Hellström, Å. (1985) The time-order error and its relatives: mirrors of cognitive processes in comparing. *Psychological Bulletin,* 97 (1), 35-61
Hellström, Å. and Rammsayer, T. H. (2004) Effects of time-order, interstimulus interval, and feedback in duration discrimination of noise bursts in the 50-and 1000-ms ranges. *Acta Psychologica,* 116 (1), 1-20.
Heron, J., Aaen-Stockdale, C., Hotchkiss, J., Roach, N. W., McGraw, P. V. and Whitaker, D. (2012) Duration channels mediate human time perception. *Proceedings of the Royal Society B-Biological Sciences,* 279 (1729), 690-698.
Heron, J., Hotchkiss, J., Aaen-Stockdale, C., Roach, N. W. and Whitaker, D. (2013) A neural hierarchy for illusions of time: duration adaptation precedes multisensory integration. *Journal of Vision,* 13 (14), doi:10.1167/13.14.4
Heron, J., Whitaker, D. and McGraw, P. V. (2004) Sensory uncertainty governs the extent of audio-visual interaction. *Vision Research,* 44 (25), 2875-2884.
Hicks, R. E., Miller, G. W. and Kinsbourne, M. (1976) Prospective and retrospective judgments of time as a function of amount of information processed. *The American Journal of Psychology,* 89 (4), 719-730.
Hinton, S. C., Harrington, D. L., Binder, J. R., Durgerian, S. and Rao, S. M. (2004) Neural systems supporting timing and chronometric counting: an FMRI study. *Cognitive Brain Research,* 21 (2), 183-192.
Hore, J., Watts, S., Tweed, D. and Miller, B. (1996) Overarm throws with the nondominant arm: kinematics of accuracy. *Journal of Neurophysiology,* 76 (6), 3693-3704.

Howard, I. P. and Templeton, W. B. (1966) *Human Spatial Orientation.* Oxford, England: John Wiley and Sons.
Howell, E. R. and Hess, R. F. (1978) The functional area for summation to threshold for sinusoidal gratings. *Vision Research,* 18 (4), 369-374.
Hubel, D. H. (1960) Single unit activity in lateral geniculate body and optic tract of unrestrained cats. *The Journal of Physiology,* 150 (1), 91-104.
Hubel, D. H. and Wiesel, T. N. (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology,* 160 (1), 106-154.
Hubel, D. H. and Wiesel, T. N. (1968) Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology,* 195 (1), 215-243.
Hubel, D. H. and Wiesel, T. N. (1970) Stereoscopic vision in macaque monkey: cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature,* 225 (5227), 41-42.
Hubel, D. H. and Wiesel, T. N. (1972) Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. *Journal of Comparative Neurology,* 146 (4), 421-450.
Hutchinson, C. V. and Ledgeway, T. (2010) Spatial summation of first-order and second-order motion in human vision. *Vision Research,* 50 (17), 1766-1774.
Ishai, A., Schmidt, C. F. and Boesiger, P. (2005) Face perception is mediated by a distributed cortical network. *Brain Research Bulletin,* 67 (1), 87-93.
Ivry, R. B. (1996) The representation of temporal information in perception and motor control. *Current Opinion in Neurobiology,* 6 (6), 851-857.
Ivry, R. B. and 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 (1), 3-18.
Ivry, R. B. and Keele, S. W. (1989) Timing functions of the cerebellum. *Journal of Cognitive Neuroscience,* 1 (2), 136-152.
Ivry, R. B. and Richardson, T. C. (2002) Temporal control and coordination: the multiple timer model. *Brain and Cognition,* 48 (1), 117-132.
Ivry, R. B. and Schlerf, J. E. (2008) Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences,* 12 (7), 273-280.
Ivry, R. B. and Spencer, R. (2004a) The neural representation of time. *Current Opinion in Neurobiology,* 14 (2), 225-232.
Ivry, R. B., Spencer, R. M., Zelaznik, H. N. and Diedrichsen, J. (2002) The cerebellum and event timing. *Annals of the New York Academy of Sciences,* 978 (1), 302-317.
Ivry, R. B. and Spencer, R. M. C. (2004b) Evaluating the role of the cerebellum in temporal processing: beware of the null hypothesis. *Brain,* 127 (8), doi:10.1093/brain/awh226
Jagota, A., de la Iglesia, H. O. and Schwartz, W. J. (2000) Morning and evening circadian oscillations in the suprachiasmatic nucleus in vitro. *Nature Neuroscience,* 3 (4), 372-376.
Jahanshahi, M., Jones, C. R., Dirnberger, G. and Frith, C. D. (2006) The substantia nigra pars compacta and temporal processing. *The Journal of Neuroscience,* 26 (47), 12266-12273.

Jain, A., Bansal, R., Kumar, A. and Singh, K. D. (2015) A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students. *International Journal of Applied and Basic Medical Research,* 5 (2), 124-127.

Jamieson, D. G. and Petrusic, W. M. (1975) The dependence of time-order error direction on stimulus range. *Canadian Journal of Psychology,* 29 (3), 175-183

Janssen, P. and Shadlen, M. N. (2005) A representation of the hazard rate of elapsed time in macaque area LIP. *Nature Neuroscience,* 8 (2), 234-241.

Janssen, P., Vogels, R., Liu, Y. and Orban, G. A. (2003) At least at the level of inferior temporal cortex, the stereo correspondence problem is solved. *Neuron,* 37 (4), 693-701.

Janssen, P., Vogels, R. and Orban, C. A. (2000) Selectivity for 3D shape that reveals distinct areas within macaque inferior temporal cortex. *Science,* 288 (5473), 2054-2056.

Janssen, P., Vogels, R. and Orban, G. A. (1999) Macaque inferior temporal neurons are selective for disparity-defined three-dimensional shapes. *Proceedings of the National Academy of Sciences,* 96 (14), 8217-8222.

Jantzen, K. J. S., F. L. and Kelso, J. A. S. (2005) Functional MRI reveals the existence of modality and coordination-dependent timing networks. *Neuroimage,* 25 (4), 1031-1042.

Jazayeri, M. and Shadlen, M. N. (2015) A neural mechanism for sensing and reproducing a time interval. *Current Biology,* 25 (20), 2599-2609.

Jeffress, L. A. (1948) A place theory of sound localization. *Journal of Comparative and Physiological Psychology,* 41 (1), 35-39.

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

Johnston, A. (1987) Spatial scaling of central and peripheral contrast-sensitivity functions. *Journal of the Optical Society of America A,* 4 (8), 1583-1593.

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

Johnston, A., Bruno, A., Watanabe, J., Quansah, B., Patel, N., Dakin, S. and Nishida, S. (2008) Visually-based temporal distortion in dyslexia. *Vision Research,* 48 (17), 1852-1858.

Jones, C. R., Malone, T. J., Dirnberger, G., Edwards, M. and Jahanshahi, M. (2008) Basal ganglia, dopamine and temporal processing: performance on three timing tasks on and off medication in Parkinson's disease. *Brain and Cognition,* 68 (1), 30-41.

Jones, C. R., Rosenkranz, K., Rothwell, J. C. and Jahanshahi, M. (2004) The right dorsolateral prefrontal cortex is essential in time reproduction: an investigation with repetitive transcranial magnetic stimulation. *Experimental Brain Research,* 158 (3), 366-372.

Jones, M. R., Moynihan, H., MacKenzie, N. and Puente, J. (2002) Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science,* 13 (4), 313-319.

Julesz, B. (1971) *Foundations of Cyclopean Perception.* Chicago: University of Chicago Press.
Jurado, M. B. and Rosselli, M. (2007) The elusive nature of executive functions: a review of our current understanding. *Neuropsychology Review,* 17 (3), 213-233.
Kaas, J. H. and Hackett, T. A. (2000) Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences of the United States of America,* 97 (22), 11793-11799.
Kalenscher, T., Ohmann, T., Windmann, S., Freund, N. and Güntürkün, O. (2006) Single forebrain neurons represent interval timing and reward amount during response scheduling. *European Journal of Neuroscience,* 24 (10), 2923-2931.
Kanai, R., Lloyd, H., Bueti, D. and Walsh, V. (2011) Modality-independent role of the primary auditory cortex in time estimation. *Experimental Brain Research,* 209 (3), 465-471.
Kanwisher, N., McDermott, J. and Chun, M. M. (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience,* 17 (11), 4302-4311.
Karmarkar, U. R. and Buonomano, D. V. (2007) Timing in the absence of clocks: encoding time in neural network states. *Neuron,* 53 (3), 427-438.
Kashino, M. and Nishida, S. y. (1998) Adaptation in the processing of interaural time differences revealed by the auditory localization aftereffect. *The Journal of the Acoustical Society of America,* 103 (6), 3597-3604.
Kaskan, P. M., Lu, H. D. D., Dillenburger, B. C., Kaas, J. H. and Roe, A. W. (2009) The organization of orientation-selective, luminance-change and binocular-preference domains in the second (V2) and third (V3) visual areas of new world owl monkeys as revealed by intrinsic signal optical imaging. *Cerebral Cortex,* 19 (6), 1394-1407.
Kelly, J. B. G., S. Lynn and Beaver, C. J. (1991) Sound frequency and binaural properties of single neurons in rat inferior colliculus. *Hearing Research,* 56 (1-2), 273-280.
Kim, R., Peters, M. A. K. and Shams, L. (2012) 0+1 > 1: how adding noninformative sound improves performance on a visual task. *Psychological Science,* 23 (1), 6-12.
Kingdom, F. A. and Keeble, D. R. (1996) A linear systems approach to the detection of both abrupt and smooth spatial variations in orientation-defined textures. *Vision Research,* 36 (3), 409-420.
Kingdom, F. A. and Keeble, D. R. (1999) On the mechanism for scale invariance in orientation-defined textures. *Vision Research,* 39 (8), 1477-1489.
Klapproth, F. (2011) Temporal decision making in simultaneous timing. *Frontiers in Integrative Neuroscience,* 5 (71), doi:10.3389/fnint.2011.00071
Klink, P. C., Montijn, J. S. and van Wezel, R. J. A. (2011) Crossmodal duration perception involves perceptual grouping, temporal ventriloquism, and variable internal clock rates. *Attention Perception & Psychophysics,* 73 (1), 219-236.

Klumpp, R. G. and Eady, H. R. (1956) Some measurements of interaural time difference thresholds. *The Journal of the Acoustical Society of America,* 28 (5), 859-860.
Kobatake, E., Tanaka and K. (1994) Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology,* 71 (3), 856-867.
Koch, G., Oliveri, M., Torriero, S., Salerno, S., Gerfo, E. L. and Caltagirone, C. (2007) Repetitive TMS of cerebellum interferes with millisecond time processing. *Experimental Brain Research,* 179 (2), 291-299.
Kohn, A. (2007) Visual adaptation: physiology, mechanisms, and functional benefits. *Journal of Neurophysiology,* 97 (5), 3155-3164.
Kohn, A. and Movshon, J. A. (2003) Neuronal adaptation to visual motion in area MT of the macaque. *Neuron,* 39 (4), 681-691.
Krauchi, K. (2002) How is the circadian rhythm of core body temperature regulated? *Clinical Autonomic Research,* 12 (3), 147-149.
Kuffler, S. W. (1953) Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology,* 16 (1), 37-68.
Kuling, I. A., Kohlrausch, A. and Juola, J. F. (2013) Quantifying temporal ventriloquism in audiovisual synchrony perception. *Attention, Perception and Psychophysics,* 75 (7), 1583-1599.
Kuling, I. A., Van Eijk, R.L.J., Juola, J. F. and Kohlrausch, A. (2012) Effects of stimulus duration on audio-visual synchrony perception. *Experimental Brain Research,* 221 (4), 403-412.
Kösem, A. and Van Wassenhove, V. (2012) Temporal structure in audiovisual sensory selection. *PloS ONE,* 7 (7), e40936.
Laje, R. and Buonomano, D. V. (2013) Robust timing and motor patterns by taming chaos in recurrent neural networks *Nature Neuroscience,* 16 (7), 925-933.
Lapid, E., Ulrich, R. and Rammsayer, T. (2009) Perceptual learning in auditory temporal discrimination: No evidence for a cross-modal transfer to the visual modality. *Psychonomic Bulletin & Review,* 16 (2), 382-389.
Larsson, J. and Harrison, S. J. (2015) Spatial specificity and inheritance of adaptation in human visual cortex. *Journal of Neurophysiology,* 114 (2), 1211-1226.
Larsson, J., Landy, M. S. and Heeger, D. J. (2006) Orientation-selective adaptation to first-and second-order patterns in human visual cortex. *Journal of Neurophysiology,* 95 (2), 862-881.
Latimer, K., Curran, W. and Benton, C. P. (2014) Direction-contingent duration compression is primarily retinotopic. *Vision Research,* 105, 47-52.
Leary, C. J., Edwards, C. J. and 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 (21), 5481-5493.
Lebedev, M. A., O'Doherty, J. E. and Nicolelis, M. A. (2008) Decoding of temporal intervals from cortical ensemble activity. *Journal of Neurophysiology,* 99 (1), 166-186.

Lee, B. B., Martin, P. R. and Grünert, U. (2010) Retinal connectivity and primate vision. *Progress in Retinal and Eye Research,* 29 (6), 622-639.
Lee, H. A. and Lee, S. H. (2012) Hierarchy of direction-tuned motion adaptation in human visual cortex. *Journal of Neurophysiology,* 107 (8), 2163-2184.
Lee, K.-H., Egleston, P. N., Brown, W. H., Gregory, A. N., Barker, A. T. and Woodruff, P. W. R. (2007) The role of the cerebellum in subsecond time perception: evidence from repetitive transcranial magnetic stimulation. *Journal of Cognitive Neuroscience,* 19 (1), 147-157.
Lehmkuhle, S. W. and Fox, R. (1976) Measuring interocular transfer. *Vision Research,* 16 (4), 428-430.
Lejeune, H. (1998) Switching or gating? The attentional challenge in cognitive models of psychological time. *Behavioural Processes,* 44 (2), 127-145.
Lejeune, H. and Wearden, J. H. (2009) Vierordt's the experimental study of the time sense (1868) and its legacy. *European Journal of Cognitive Psychology,* 21 (6), 941-960.
Leon, M. I. and Shadlen, M. N. (2003) Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron,* 38 (2), 317-327.
Leopold, D. A., Bondar, I. V. and Giese, M. A. (2006) Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature,* 442 (7102), 572-575.
Leopold, D. A., O'Toole, A. J., Vetter, T. and Blanz, V. (2001) Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience,* 4 (1), 89-94.
Levinson, E. and Sekuler, R. (1976) Adaptation alters perceived direction of motion. *Vision Research,* 16 (7), 779-781.
Levitan, C. A., Ban, Y.-H. A., Stiles, N. R. and Shimojo, S. (2015) Rate perception adapts across the senses: evidence for a unified timing mechanism. *Scientific Reports,* 5:8857, doi:10.1038/srep08857
Levitan, I. B. and Kaczmarek, L. K. (2002) *The neuron: cell and molecular biology.* Oxford: Oxford University Press.
Levitt, J. B., Kiper, D. C. and Movshon, J. A. (1994) Receptive fields and functional architecture of macaque V2. *Journal of Neurophysiology,* 71 (6), 2517-2542.
Lewis, P. A. and Miall, R. C. (2003a) Brain activation patterns during measurement of sub-and supra-second intervals. *Neuropsychologica,* 41 (12), 1583-1592.
Lewis, P. A. and Miall, R. C. (2003b) Distinct systems for automatic and cognitively controlled time measurement: evidence from neuroimaging. *Current Opinion in Neurobiology,* 13 (2), 250-255.
Lewis, P. A. and Miall, R. C. (2009) The precision of temporal judgement: milliseconds, many minutes, and beyond. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 364 (1525), 1897-1905.
Ley, I., Haggard, P. and Yarrow, K. (2009) Optimal integration of auditory and vibrotactile information for judgments of temporal order. *Journal of Experimental Psychology: Human Perception and Performance,* 35 (4), 1005-1019.

Li, B., Yuan, X., Chen, Y., Liu, P. and Huang, X. (2015a) Visual duration aftereffect is position invariant. *Frontiers in Psychology,* 6:1536, doi:10.3389/fpsyg.2015.01536

Li, B. L., Yuan, X. Y. and Huang, X. T. (2015b) The aftereffect of perceived duration is contingent on auditory frequency but not visual orientation. *Scientific Reports,* 5:10124, doi:10.1038/srep10124.

Li, G. and Baker, C. L. (2012) Functional organization of envelope-responsive neurons in early visual cortex: Organization of carrier tuning properties. *Journal of Neuroscience,* 32 (22), 7538-7549.

Li, G., Yao, Z., Wang, Z., Yuan, N., Talebi, V., Tan, J., Wang, Y., Zhou, Y. and Baker, C. L. (2014) Form-cue invariant second-order neuronal responses to contrast modulation in primate area V2. *Journal of Neuroscience,* 34 (36), 12081-12092.

Llinás, R. (1988) The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. *Science,* 242, 1654-1664.

Logan, A. J., Wilkinson, F., Wilson, H. R., Gordon, G. E. and Loffler, G. (2016) The Caledonian face test: A new test of face discrimination. *Vision Research,* 119, 29-41.

Logothetis, N. K., Pauls, J. and Poggio, T. (1995) Shape representation in the inferior temporal cortex of monkeys. *Current Biology,* 5 (5), 552-563.

Ma, X. and Suga, N. (2001) Corticofugal modulation of duration-tuned neurons in the midbrain auditory nucleus in bats. *Proceedings of the National Academy of Sciences,* 98 (24), 14060-14065.

Macar, F., Grondin, S. and Casini, L. (1994) Controlled attention sharing influences time estimation. *Memory and Cognition,* 22 (6), 673-686.

Macar, F., Lejeune, H., Bonnet, M., Ferrara, A., Pouthas, V., Vidal, F. and Maquet, P. (2002) Activation of the supplementary motor area and of attentional networks during temporal processing. *Experimental Brain Research,* 142 (4), 475-485.

Macías, S., Hechavarría, J. C., Kössl, M. and Mora, E. C. (2013) Neurons in the inferior colliculus of the mustached bat are tuned both to echo-delay and sound duration. *Neuroreport,* 24 (8), 404-409.

Malik, J. and Perona, P. (1990) Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America,* 7 (5), 923-932.

Mangels, J. A., Ivry, R. B. and Shimizu, N. (1998) Dissociable contributions of the prefrontal and neocerebellar cortex to time perception. *Cognitive Brain Research,* 7 (1), 15-39.

Maquet, P., Lejeune, H., Pouthas, V., Bonnet, M., Casini, L., Macar, F., Timsit-Berthier, M., Vidal, F., Ferrara, A., Degueldre, C. and Quaglia, L. (1996) Brain activation induced by estimation of duration: a PET study. *Neuroimage,* 3 (2), 119-126.

Mareschal, I. and Baker, C. L. (1998) Temporal and spatial response to second-order stimuli in cat area 18. *Journal of Neurophysiology,* 80 (6), 2811-2823.

Mareschal, I. and Baker, C. L. (1999) Cortical processing of second-order motion. *Visual Neuroscience,* 16 (3), 527-540.

Marr, D. (1982) *Vision - A computational investigation into the human representation and processing of visual information.* San Francisco: WH Freeman.
Matell, M. S. and Meck, W. H. (2000) Neuropsychological mechanisms of interval timing behavior. *Bioessays,* 22 (1), 94-103.
Matell, M. S. and Meck, W. H. (2004) Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Cognitive Brain Research,* 21 (2), 139-170.
Matell, M. S., Meck, W. H. and Nicolelis, M. A. (2003) Interval timing and the encoding of signal duration by ensembles of cortical and striatal neurons. *Behavioral Neuroscience,* 117 (4), 760-773
Mattes, S. and Ulrich, R. (1998) Directed attention prolongs the perceived duration of a brief stimulus. *Perception & Psychophysics,* 60 (8), 1305-1317.
Matthews, W. J., Stewart, N. and Wearden, J. H. (2011) Stimulus Intensity and the Perception of Duration. *Journal of Experimental Psychology-Human Perception and Performance,* 37 (1), 303-313.
Mauk, M. D. and Buonomano, D. V. (2004) The neural basis of temporal processing. *Annual Review of Neuroscience,* 27, 307-340.
Maunsell, J. H. and Newsome, W. T. (1987) Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience,* 10 (1), 363-401.
Maunsell, J. H. R. and Vanessen, D. C. (1983) Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology,* 49 (5), 1148-1167.
Mayhew, J. E. W. and Anstis, S. M. (1972) Movement aftereffects contingent on color, intensity, and pattern. *Perception & Psychophysics,* 12 (1B), 77-85.
McCollough, C. (1965) Color adaptation of edge-detectors in the human visual system. *Science,* 149 (3688), 1115-1116.
Meck, W. H., Penney, T. B. and Pouthas, V. (2008) Cortico-striatal representation of time in animals and humans. *Current Opinion in Neurobiology,* 18 (2), 145-152.
Menon, R. S., Ogawa, S., Strupp, J. P. and Uğurbil, K. (1997) Ocular dominance in human V1 demonstrated by functional magnetic resonance imaging. *Journal of Neurophysiology,* 77 (5), 2780-2787.
Merchant, H., Harrington, D. L. and Meck, W. H. (2013a) Neural basis of the perception and estimation of time. *Annual Review of Neuroscience,* 36, 313-336.
Merchant, H., Pérez, O., Zarco, W. and Gámez, J. (2013b) Interval tuning in the primate medial premotor cortex as a general timing mechanism. *The Journal of Neuroscienc,e* 33 (21), 9082-9096.
Merchant, H., Zarco, W. and 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 (2), 939-949.
Merchant, H., Zarco, W., Pérez, O., Prado, L. and Bartolo, R. (2011) Measuring time with different neural chronometers during a synchronization-continuation task. *Proceedings of the National Academy of Sciences,* 108 (49), 19784-19789.

Merzenich, M. M. and Brugge, J. F. (1973) Representation of the cochlear partition on the superior temporal plane of the macaque monkey. *Brain Research,* 50 (2), 275-296.
Miall, C. (1989) The storage of time intervals using oscillating neurons. *Neural Computation,* 1 (3), 359-371.
Miles, L. E. M., Raynal, D. M. and Wilson, M. A. (1977) Blind man living in normal society has circadian-rhythms of 24.9 hours. *Science,* 198 (4315), 421-423.
Miller, J. E., Carlson, L.A. and McAuley, J. D. (2013) When what you hear influences when you see: listening to an auditory rhythm influences the temporal allocation of visual attention. *Psychological Science,* 24 (1), 11-18.
Mills, A. W. (1958) On the minimum audible angle. *The Journal of the Acoustical Society of America,* 30 (4), 237-246.
Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y. and Tanji, J. (2009) Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nature Neuroscience,* 12 (4), 502-507.
Mitchell, D. E. and Muir, D. W. (1976) Does the tilt after-effect occur in the oblique meridian? *Vision Research,* 16 (6), 609-613.
Mitchell, D. E. and Ware, C. (1974) Interocular transfer of a visual after-effect in normal and stereoblind humans. *Journal of Physiology,* 236 (3), 707-721.
Mo, S. S. (1975) Temporal reproduction of duration as a function of numerosity. *Bulletin of the Psychonomic Society,* 5 (2), 165-167.
Mo, S. S. and Michalski, V. A. (1972) Judgment of temporal duration of area as a function of stimulus configuration. *Psychonomic Science,* 27 (2), 97-98.
Montaser-Kouhsari, L., Landy, M. S., Heeger, D. J. and Larsson, J. (2007) Orientation-selective adaptation to illusory contours in human visual cortex. *Journal of Neuroscience,* 27 (9), 2186-2195.
Morein-Zamir, S., Soto-Faraco, S. and Kingstone, A. (2003) Auditory capture of vision: examining temporal ventriloquism. *Cognitive Brain Research,* 17 (1), 154-163.
Morgan, M. J., Giora, E. and Solomon, J. A. (2008) A single “stopwatch” for duration estimation, a single “ruler” for size. *Journal of Vision,* 8 (2), doi:10.1167/8.2.14
Morillon, B., Kell, C. A. and Giraud, A. L. (2009) Three stages and four neural systems in time estimation. *Journal of Neuroscience,* 29 (47), 14803-14811.
Moro, S. S., Harris, L. R. and Steeves, J. K. E. (2014) Optimal audiovisual integration in people with one eye. *Multisensory Research,* 27 (3-4), 173-188.
Morrone, M. C., Ross, J. and Burr, D. (2005) Saccadic eye movements cause compression of time as well as space. *Nature Neuroscience,* 8 (7), 950-954.
Motter, B. C. and Mountcastle, V. B. (1981) The functional properties of the light-sensitive neurons of the posterior parietal cortex studied in waking monkeys: foveal sparing and opponent vector organization. *Journal of Neuroscience,* 1 (1), 3-26.

Movshon, J. A., Thompson, I. D. and Tolhurst, D. J. (1978) Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *The Journal of Physiology,* 283, 53-77.
Murch, G. M. (1972) Binocular relationships in a size and color orientation specific aftereffect. *Journal of Experimental Psychology,* 93 (1), 30-34.
Nenadic, I., Gaser, C., Volz, H. P., Rammsayer, T., Häger, F. and Sauer, H. (2003) Processing of temporal information and the basal ganglia: new evidence from fMRI. *Experimental Brain Research,* 148 (2), 238-246.
Nichelli, P., Alway, D. and Grafman, J. (1996) Perceptual timing in cerebellar degeneration. *Neuropsychologia,* 34 (9), 863-871.
Nichelli, P., Clark, K., Hollnagel, C. and Grafman, J. (1995) Duration processing after frontal lobe lesions. *Annals of the New York Academy of Sciences,* 769 (1), 183-190.
Nishida, S., Ashida, H. and Sato, T. (1994) Complete interocular transfer of motion aftereffect with flickering test. *Vision Research,* 34 (20), 2707-2716.
Noulhiane, M., Mella, N., Samson, S., Ragot, R. and Pouthas, V. (2007) How emotional auditory stimuli modulate time perception. *Emotion,* 7 (4), 697-704
Oprisan, S. A. and Buhusi, C. V. (2013) How noise contributes to time-scale invariance of interval timing. *Physical Review E,* 87 (5), doi:10.1103/PhysRevE.87.052717
Ortega, L., Guzman-Martinez, E., Grabowecky, M. and Suzuki, S. (2012) Flicker adaptation of low-level cortical visual neurons contributes to temporal dilation. *Journal of Experimental Psychology: Human Perception and Performance,* 38 (6), 1380-1389
Ortega, L., Guzman-Martinez, E., Grabowecky, M. and Suzuki, S. (2014) Audition dominates vision in duration perception irrespective of salience, attention, and temporal discriminability. *Attention, Perception & Psychophysics,* 76 (5), 1485-1502.
Ortega, L., Lopez, F. and Church, R. M. (2009) Modality and intermittency effects on time estimation. *Behavioural Processes,* 81 (2), 270-273.
Oshio, K. I., Chiba, A. and Inase, M. (2008) Temporal filtering by prefrontal neurons in duration discrimination. *European Journal of Neuroscience,* 28 (11), 2333-2343.
Osugi, T., Takeda, Y. and Murakami, I. (2016) Inhibition of return shortens perceived duration of a brief visual event. *Vision Research,* 128, 39-44.
Pariyadath, V. and Eagleman, D. (2007) The effect of predictability on subjective duration. *PloS ONE,* 2 (11), e1264.
Pariyadath, V. and Eagleman, D. M. (2008) Brief subjective durations contract with repetition. *Journal of Vision,* 8 (16) doi:10.1167/8.16.11
Parker, K. L., Chen, K.-h., Kingyon, J. R., Cavanagh, J. F. and Narayanan, N. S. (2014) D 1 -dependent 4 Hz oscillations and ramping activity in rodent medial frontal cortex during interval timing. *Journal of Neuroscience,* 34 (50), 16774–16783.
Pastor, M., Artieda, J., Jahanshahi, M. and Obeso, J. (1992) Time estimation and reproduction is abnormal in Parkinson's disease. *Brain,* 115 (1), 211-225.

Pasupathy, A. and Connor, C. E. (1999) Responses to contour features in macaque area V4. *Journal of Neurophysiology,* 82 (5), 2490-2502.
Patterson, C. A., Wissig, S. C. and Kohn, A. (2014) Adaptation disrupts motion integration in the primate dorsal stream. *Neuron,* 81 (3), 674-686.
Penney, T. B., Gibbon, J. and Meck, W. H. (2000) Differential effects of auditory and visual signals on clock speed and temporal memory. *Journal of Experimental Psychology: Human Perception and Performance,* 26 (6), 1770-1787
Penton-Voak, I. S., Edwards, H., Percival, A. and Wearden, J. H. (1996) Speeding up an internal clock in humans? Effects of click trains on subjective duration. *Journal of Experimental Psychology: Animal Behavior Processes,* 22 (3), 307-320.
Perez-Gonzalez, D., Malmierca, M. S., Moore, J. M., Hernandez, O. and 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 (2), 823-836.
Perrault, N. and Picton, T. W. (1984) Event-related potentials recorded from the scalp and nasopharynx. I. N1 and P2. *Electroencephalography and Clinical Neurophysiology,* 59 (3), 177-194.
Perrett, D. I., Rolls, E. T. and Caan, W. (1982) Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research,* 47 (3), 329-342.
Pick, H. L., Warren, D. H. and Hay, J. C. (1969) Sensory conflict in judgments of spatial direction. *Perception & Psychophysics,* 6 (4), 203-205.
Piras, F. and Coull, J. T. (2011) Implicit, predictive timing draws upon the same scalar representation of time as explicit timing. *PloS ONE,* 6 (3), e18203.
Poggio, G., F., Gonzalez, F. and Krause, F. (1988) Stereoscopic mechanisms in monkey visual cortex, binocular correlation and disparity selectivity. *The Journal of Neuroscience,* 8 (12), 4531-4550.
Poggio, G. and Fischer, B. (1977) Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology,* 40 (6), 1392-1405.
Port, N. L., Kruse, W., Lee, D. and Georgopoulos, A. P. (2001) Motor cortical activity during interception of moving targets. *Journal of Cognitive Neuroscience,* 13 (3), 306-318.
Pouthas, V., George, N., Poline, J. B., Pfeuty, M., VandeMoorteele, P. F., Hugueville, L., Ferrandez, A. M., Lehericy, S., LeBihan, D. and Renault, B. (2005) Neural network involved in time perception: an fMRI study comparing long and short interval estimation. *Human Brain Mapping,* 25 (4), 433-441.
Priebe, N. J., Churchland, M. M. and Lisberger, S. G. (2002) Constraints on the source of short-term motion adaptation in macaque area MT. I. The role of input and intrinsic mechanisms. *Journal of Neurophysiology,* 88 (1), 354-369.

Priebe, N. J. and Lisberger, S. G. (2002) Constraints on the source of short-term motion adaptation in macaque area MT. II. Tuning of neural circuit mechanisms. *Journal of Neurophysiology,* 88 (1), 370-382.
Prince, S. J. D., Pointon, A. D., Cumming, B. G. and Parker, A. J. (2002) Quantitative analysis of the responses of V1 neurons to horizontal disparity in dynamic random-dot stereograms. *Journal of Neurophysiology,* 87 (1), 191-208.
Quraishi, S., Heider, B. and Siegel, R. M. (2007) Attentional modulation of receptive field structure in area 7a of the behaving monkey. *Cerebral Cortex,* 17 (8), 1841-1857.
Rammsayer, T. and Classen, W. (1997) Impaired temporal discrimination in Parkinson's disease: temporal processing of brief durations as an indicator of degeneration of dopaminergic neurons in the basal ganglia. *International Journal of Neuroscience,* 91 (1-2), 45-55.
Rammsayer, T. and Ulrich, R. (2005) No evidence for qualitative differences in the processing of short and long temporal intervals. *Acta Psychologica,* 120 (2), 141-171.
Rammsayer, T. and Verner, M. (2015) Larger visual stimuli are perceived to last longer from time to time : The internal clock is not affected by nontemporal visual stimulus size. *Journal of Vision,* 15 (3), doi:10.1167/15.3.5
Rammsayer, T. H. (1999) Neuropharmacological evidence for different timing mechanisms in humans. *The Quarterly Journal of Experimental Psychology: Section B,* 52 (3), 273-286.
Rammsayer, T. H. (2010) Differences in duration discrimination of filled and empty auditory intervals as a function of base duration. *Attention, Perception, & Psychophysics,* 72 (6), 1591-1600.
Rammsayer, T. H. and Lima, S. D. (1991) Duration discrimination of filled and empty auditory intervals: Cognitive and perceptual factors. *Perception & Psychophysics,* 50 (6), 565-574.
Rauschecker, J. P., Tian, B., Pons, T. and Mishkin, M. (1997) Serial and parallel processing in rhesus monkey auditory cortex. *Journal of Comparative Neurology,* 382 (1), 89-103.
Recanzone, G. H. (2003) Auditory influences on visual temporal rate perception. *Journal of Neurophysiology,* 89 (2), 1078-1093.
Remington, R. W., Johnston, J. C. and Yantis, S. (1992) Involuntary attentional capture by abrupt onsets. *Perception & Psychophysics,* 51 (3), 279-290.
Renoult, L., Roux, S. and Riehle, A. (2006) Time is a rubberband: neuronal activity in monkey motor cortex in relation to time estimation. *European Journal of Neuroscience,* 23 (11), 3098-3108.
Repp, B. H. and Penel, A. (2002) Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance,* 28 (5), 1085-1099
Reutimann, J., Yakovlev, V., Fusi, S. and Senn, W. (2004) Climbing neuronal activity as an event-based cortical representation of time. *The Journal of Neuroscience,* 24 (13), 3295-3303.
Reynolds, J. H. and Heeger, D. J. (2009) The normalization model of attention. *Neuron,* 61 (2), 168-185.

Riesenhuber, M. and Poggio, T. (1999) Hierarchical models of object recognition in cortex. *Nature Neuroscience,* 2 (11), 1019-1025.
Roach, N. W., Heron, J. and McGraw, P. V. (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 (1598), 2159-2168.
Rochefort, N. L., Buzás, P., Quenech'Du, N., Koza, A., Eysel, U. T., Milleret, C. and Kisvárday, Z. F. (2009) Functional selectivity of interhemispheric connections in cat visual cortex. *Cerebral Cortex,* 19 (10), 2451-2465.
Rockland, K. S. and Ojima, H. (2003) Multisensory convergence in calcarine visual areas in macaque monkey. *International Journal of Psychphysiology,* 50 (1), 19-26.
Rolls, E. and Baylis, G. (1986) Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Experimental Brain Research,* 65 (1), 38-48.
Rolls, E. T. (1984) Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Human Neurobiology,* 3 (4), 209-222.
Rolls, E. T., Milward and T. (2000) A model of invariant object recognition in the visual system: learning rules, activation functions, lateral inhibition, and information-based performance measures. *Neural Computation,* 12 (11), 2547-2572.
Romei, V., De Haas, B., Mok, R. M. and Driver, J. (2011) Auditory stimulus timing influences perceived duration of co-occurring visual stimuli. *Frontiers in Psychology,* 2:215, doi:10.3389/fpsyg.2011.00215
Rose, D. and Summers, J. (1995) Duration illusions in a train of visual stimuli. *Perception,* 24 (10), 1177-1187.
Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C. and Foxe, J. J. (2007) Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex,* 17 (5), 1147-1153.
Rousseau, L. and Rousseau, R. (1996) Stop—reaction time and the internal clock. *Perception & Psychophysics,* 58 (3), 434-448.
Rousseau, R., Poirier, J. and Lemyre, L. (1983) Duration discrimination of empty time intervals marked by intermodal pulses. *Perception & Psychophysics,* 34 (6), 541-548.
Roy, J.-P., Komatsu, H. and Wurtz, R. W. (1992) Disparity sensitivity of neurons in monkey extrastriate area MST. *The Journal of Neuroscience,* 12 (7), 2478-2492.
Ruthruff, E. and Pashler, H. (2010) Mental timing and the central attentional bottleneck. In Nobre, A. C. and Coull, J. T. (editors) *Attention and Time.* Oxford: Oxford University Press. 123 - 136.
Sack, R. L., Lewy, A. J., Blood, M. L., Keith, L. D. and Nakagawa, H. (1992) Circadian rhythm abnormalities in totally blind people - incidence and clinical significance. *Journal of Clinical Endocrinology and Metabolism,* 75 (1), 127-134.

Sagi, D. (1990) Detection of an orientation singularity in Gabor textures: effect of signal density and spatial-frequency. *Vision Research,* 30 (9), 1377-1388.
Salvioni, P., Murray, M. M., Kalmbach, L. and Bueti, D. (2013) How the visual brain encodes and keeps track of time. *The Journal of Neuroscience,* 33 (30), 12423-12429.
Sarmiento, B. R., Shore, D. I., Milliken, B. and Sanabria, D. (2012) Audiovisual interactions depend on context of congruency. *Attention, Perception, & Psychophysics,* 74 (3), 563-574.
Scheier, C. R., Nijhawan, R. and Shimojo, S. (1999) Sound alters visual temporal resolution. *Investigative Ophthalmology and Visual Science,* 40 (4), S792.
Schiffman, H. R. and Bobko, D. J. (1974) Effects of stimulus complexity on the perception of brief temporal intervals. *Journal of Experimental Psychology,* 103 (1), 156-159.
Schubotz, R. I., Friederici, A. D. and Yves von Cramon, D. (2000) Time perception and motor timing: a common cortical and subcortical basis revealed by fMRI. *Neuroimage,* 11 (1), 1-12.
Schwartz, E. L., Desimone, R., Albright, T. D. and Gross, C. G. (1983) Shape recognition and inferior temporal neurons. *Proceedings of the National Academy of Sciences,* 80 (18), 5776-5778.
Seifried, T. and Ulrich, R. (2010) Does the asymmetry effect inflate the temporal expansion of odd stimuli? *Psychological Research,* 74 (1), 90-98.
Seilheimer, R. L., Rosenberg, A. and Angelaki, D. E. (2014) Models and processes of multisensory cue combination. *Current Opinion in Neurobiology,* 25, 38-46.
Seitz, A. R., Kim, R. and Shams, L. (2006) Sound facilitates visual learning. *Current Biology,* 16 (14), 1422-1427.
Seriès, P., Stocker, A. A. and Simoncelli, E. P. (2009) Is the homunculus “aware” of sensory adaptation? *Neural Computation,* 21 (12), 3271-3304.
Shams, L., Kamitani, Y. and Shimojo, S. (2000) Illusions: what you see is what you hear. *Nature,* 408 (6814), 788.
Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J. and Ekelid, M. (1995) Speech recognition with primarily temporal cues. *Science,* 270 (5234), 303-304.
Shelton, J. and Kumar, G. P. (2010) Comparison between auditory and visual simple reaction times. *Neuroscience and Medicine,* 1, 30-32.
Shi, Z. H., Ganzenmuller, S. and Muller, H. J. (2013) Reducing bias in auditory duration reproduction by integrating the reproduced signal. *Plos ONE,* 8 (4), e62065
Shih, L. Y. L., Kuo, W.-J., Yeh, T.-C., Tzeng, O. J. L. and Hsieh, J.-C. (2009) Common neural mechanisms for explicit timing in the sub-second range. *Neuroreport,* 20 (10), 897-901.
Shipley, T. (1964) Auditory flutter-driving of visual flicker. *Science (New York, N.Y.)* 145 (3638), 1328-30.
Shuler, M. G. and Bear, M. F. (2006) Reward timing in the primary visual cortex. *Science,* 311 (5767), 1606-1609.

Smith, J. G., Harper, D. N., Gittings, D. and Abernethy, D. (2007) The effect of Parkinson's disease on time estimation as a function of stimulus duration range and modality. *Brain and Cognition,* 64 (2), 130-143.
Smith, M. A., Kohn, A. and Movshon, J.A., (2007) Glass pattern responses in macaque V2 neurons. *Journal of Vision,* 7 (3), doi:10.1167/7.3.5.
Snowden, R., Snowden, R. J., Thompson, P. and Troscianko, T. (2012) *Basic vision: an introduction to visual perception.* Oxford University Press.
Solomon, S. G., Peirce, J. W., Dhruv, N. T. and Lennie, P. (2004) Profound contrast adaptation early in the visual pathway. *Neuron,* 42 (1), 155-162.
Spence, C., Shore, D. I. and Klein, R. M. (2001) Multisensory prior entry. *Journal of Experimental Psychology: General,* 130 (4), 799-832.
Spencer, R. M., Zelaznik, H. N., Diedrichsen, J. and Ivry, R. B. (2003) Disrupted timing of discontinuous but not continuous movements by cerebellar lesions. *Science,* 300 (5624), 1437-1439.
Staddon, J. E. R. and Higa, J. J. (1999) Time and memory: Towards a pacemaker-free theory of interval timing. *Journal of the Experimental Analysis of Behavior,* 71 (2), 215-251.
Stauffer, C. C., Haldemann, J., Troche, S. J. and Rammsayer, T. H. (2012) Auditory and visual temporal sensitivity: evidence for a hierarchical structure of modality-specific and modality-independent levels of temporal information processing. *Psychological Research,* 76 (1), 20-31.
Stein, B. E. and Stanford, T. R. (2008) Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience,* 9 (4), 255-266.
Stekelenburg, J. J. and Vroomen, J. (2007) Neural correlates of multisensory integration of ecologically valid audiovisual events. *Journal of Cognitive Neuroscience,* 19 (12), 1964-1973.
Stocker, A. A. and Simoncelli, E. P. (2009) Visual motion aftereffects arise from a cascade of two isomorphic adaptation mechanisms. *Journal of Vision,* 9 (9), doi:10.1167/9.9.9.
Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., Port, M. and Porter, N. R. (2001) When is now? Perception of simultaneity. *Proceedings of the Royal Society B: Biological Sciences,* 268 (1462), 31-38.
Stuss, D. T. and Benson, D. F. (1986) *The Frontal Lobes.* New York: Raven Press.
Sukumar, S. and Waugh, S. J. (2007) Separate first-and second-order processing is supported by spatial summation estimates at the fovea and eccentrically. *Vision Research,* 47 (5), 581-596.
Sumby, W. H. and Pollack, I. (1954) Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America,* 26 (2), 212-215.
Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M. and Egner, T. (2008) Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience,* 11 (9), 1004-1006.

Sutter, A., Sperling, G. and Chubb, C. (1995) Measuring the spatial frequency selectivity of second-order texture mechanisms. *Vision Research,* 35 (7), 915-924.
Sáry, G., Vogels, R. and Orban, G. (1993) Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science,* 260 (5110), 995-997.
Talsma, D., Senkowski, D., Soto-Faraco, S. and Woldorff, M. G. (2010) The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences,* 14 (9), 400-410.
Tanaka, H., Uka, T., Yoshiyama, K., Kato, M. and Fujita, I. (2001) Processing of shape defined by disparity in monkey inferior temporal cortex. *Journal of Neurophysiology,* 85 (2), 735-744.
Tanaka, M. (2007) Cognitive signals in the primate motor thalamus predict saccade timing. *The Journal of Neuroscience,* 27 (44), 12109-12118.
Tanner, T. A., Patton, R. M. and Atkinson, R. C. (1965) Intermodality judgments of signal duration. *Psychonomic Science,* 2 (1-12), 271-272.
Teki, S., Grube, M., Kumar, S. and Griffiths, T. D. (2011) Distinct neural substrates of duration-based and beat-based auditory timing. *The Journal of Neuroscience,* 31 (10), 3805-3812.
Thomas, E. A. and Weaver, W. B. (1975) Cognitive processing and time perception. *Attention, Perception and Psychophysics,* 17 (4), 363-367.
Thompson, P. and Burr, D. (2009) Visual aftereffects. *Current Biology,* 19 (1), R11-14.
Tillmann, B., Stevens, C. and Keller, P. E. (2011) Learning of timing patterns and the development of temporal expectations. *Psychological Research,* 75 (3), 243-258.
Tomassini, A., Gori, M., Burr, D., Sandini, G. and Morrone, C. (2011) Perceived duration of visual and tactile stimuli depends on perceived speed. *Frontiers in Integrative Neuroscience,* 5, doi:10.3389/fnint.2011.00051
Tootell, R. B., Silverman, M. S. and De Valois, R. L. (1981) Spatial frequency columns in primary visual cortex. *Science,* 214 (4522), 813-815.
Tregellas, J. R., Davalos, D. B. and Rojas, D. C. (2006) Effect of task difficulty on the functional anatomy of temporal processing. *Neuroimage,* 32 (1), 307-315.
Treisman, M. (1963) Temporal discrimination and the indifference interval. Implications for a model of the "internal clock". *Psychological Monographs,* 77 (13), 1-31.
Treisman, M., Faulkner, A., Naish, P. L. and Brogan, D. (1990) The internal clock: Evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception,* 19 (6), 705-743.
Tse, P. U., Intriligator, J., Rivest, J. and Cavanagh, P. (2004) Attention and the subjective expansion of time. *Perception & Psychophysics,* 66 (7), 1171-1189.
Uka, T., Tanaka, H., Yoshiyama, K., Kato, M. and Fujita, I. (2000) Disparity selectivity of neurons in monkey inferior temporal cortex. *Journal of Neurophysiology,* 84 (1), 120-132.

Ulrich, R., Nitschke, J. and Rammsayer, T. (2006a) Crossmodal temporal discrimination: Assessing the predictions of a general pacemaker-counter model. *Perception & Psychophysics,* 68 (7), 1140-1152.
Ulrich, R., Nitschke, J. and Rammsayer, T. (2006b) Perceived duration of expected and unexpected stimuli. *Psychological Research,* 70 (2), 77-87.
Vakrou, C., Whitaker, D. and McGraw, P. V. (2007) Extrafoveal viewing reveals the nature of second-order human vision. *Journal of Vision,* 7 (14), doi:10.1167/7.14.13.
Van de Burg, E., Alais, D. and Cass, J. (2015) Audiovisual temporal recalibration occurs independently at two different time scales. *Scientific Reports,* 5 (14526), doi:10.1038/srep14526.
Van der Burg, E., Cass, J., Olivers, C. N., Theeuwes, J. and Alais, D. (2010) Efficient visual search from synchronized auditory signals requires transient audiovisual events. *PloS ONE,* 5 (5), e10664.
Van der Burg, E., Olivers, C. N., Bronkhorst, A. W. and Theeuwes, J. (2008) Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance,* 34 (5), 1053-1065.
Van Essen, D. C., Newsome, W. T. and Maunsell, J. H. (1984) The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability. *Vision Research,* 24 (5), 429-448.
Van Rijn, H., Gu, B.-M. and Meck, W. H. (2014) Dedicated clock/timing-circuit theories of time perception and timed performance. In Merchant, H. A. D. L. V. (editor) *Neurobiology of Interval Timing.* New York: NY: Springer-Verlag. 75-99.
Van Rijn, H. and Taatgen, N. A. (2008) Timing of multiple overlapping intervals: How many clocks do we have? *Acta Psychologica,* 129 (3), 365-375.
Van Wassenhove, V., Grant, K. W. and Poeppel, D. (2007) Temporal window of integration in auditory-visual speech perception. *Neuropsychologica,* 45 (3), 598-607.
Vatakis, A. and Spence, C. (2006) Audiovisual synchrony perception for music, speech, and object actions. *Brain Research,* 1111 (1), 134-142.
Vidal, J., Giard, M. H., Roux, S., Barthelemy, C. and Bruneau, N. (2008) Cross-modal processing of auditory–visual stimuli in a no-task paradigm: A topographic event-related potential study. *Clinical Neurophysiology,* 119 (4), 763-771.
Vierordt, K. (1868) *Der Zeitsinn nach Versuchen.* Tubingen, Germany: H. Laupp.
Von Der Heyclt, R., Peterhans, E. and Baurngartner, G. (1984) Illusory contours and cortical neuron responses. *Science,* 224, 1260-1262.
Von der Heydt, R. and Peterhans, E. (1989) Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *Journal of Neuroscience,* 9 (5), 1731-1748.
Von Der Heydt, R., Peterhans, E. and Baumgartner, G. (1984) Illusory contours and cortical neuron responses. *Science,* 224, 1260-1262.

Von der Heydt, R., Zhou, H. and Friedman, H. S. (2000) Representation of stereoscopic edges in monkey visual cortex. *Vision Research,* 40 (15), 1955-1967.
Von Kriegstein, K. and Giraud, A.-L. (2006) Implicit multisensory associations influence voice recognition. *PLoS Biology,* 4 (10), e326.
Vroomen, J. and de Gelder, B. (2004) Temporal ventriloquism: sound modulates the flash-lag effect. *Journal of Experimental Psychology: Human Perception and Performance,* 30 (3), 513-518.
Vroomen, J. and Keetels, M. (2010) Perception of intersensory synchrony: a tutorial review. *Attention, Perception & Psychophysics,* 72 (4), 871-884.
Vroomen, J., Keetels, M., de Gelder, B. and Bertelson, P. (2004) Recalibration of temporal order perception by exposure of audio-visual asynchrony. *Cognitive Brain Research,* 22 (1), 32-35.
Vroomen, J. and Stekelenburg, J. J. (2011) Perception of intersensory synchrony in audiovisual speech: not that special. *Cognition,* 118 (1), 75-83.
Walker, J. T. and Irion, A. L. (1979) Two new contingent aftereffects: perceived auditory duration contingent on pitch and on temporal order. *Perception & Psychophysics,* 26 (3), 241-244.
Walker, J. T., Irion, A. L. and Gordon, D. G. (1981) Simple and contingent aftereffects of perceived duration in vision and audition. *Perception & Psychophysics,* 29 (5), 475-486.
Walker, J. T. and Scott, K. J. (1981) Auditory-visual conflicts in the perceived duration of lights, tones, and gaps. *Journal of Experimental Psychology-Human Perception and Performance,* 7 (6), 1327-1339.
Walsh, V. (2003) A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences,* 7 (11), 483-488.
Wang, J., Van Wijhe, R., Chen, Z. and Yin, S. (2006) Is duration tuning a transient process in the inferior colliculus of guinea pigs? *Brain Research,* 1114 (1), 63-74.
Ware, C. and Mitchell, D. E. (1974) Interocular transfer of various visual aftereffects in normal and stereoblind observers. *Vision Research,* 14 (8), 731-734.
Warm, J. S., Stutz, R. M. and Vassolo, P. A. (1975) Intermodal transfer in temporal discrimination. *Perception & Psychophysics,* 18 (4), 281-286.
Warren, D. H., Welch, R. B. and McCarthy, T. J. (1981) The role of visual-auditory “compellingness” in the ventriloquism effect: implications for transitivity among the spatial senses. *Perception & Psychophysics,* 30 (6), 557-564.
Wearden, J. (1999) “Beyond the fields we know...”: exploring and developing scalar timing theory. *Behavioural Processes,* 45 (1), 3-21.
Wearden, J. and Lejeune, H. (2008a) Scalar properties in human timing: conformity and violations. *The Quarterly Journal of Experimental Psychology,* 61 (4), 569-587.
Wearden, J. and McShane, B. (1988) Interval production as an analogue of the peak procedure: evidence for similarity of human and animal timing processes. *The Quarterly Journal of Experimental Psychology,* 40 (4), 363-375.

Wearden, J., Smith-Spark, J., Cousins, R., Edelstyn, N., Cody, F. and O'Boyle, D. (2008) Stimulus timing by people with Parkinson's disease. *Brain and Cognition,* 67 (3), 264-279.
Wearden, J. H. (1991a) Do humans possess an internal clock with scalar timing properties? *Learning and Motivation,* 22 (1-2), 59-83.
Wearden, J. H. (1991b) Human performance on an analogue of an interval bisection task. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology,* 43 (1), 59-81.
Wearden, J. H. (1992) Temporal generalization in humans. *Journal of Experimental Psychology-Animal Behavior Processes,* 18 (2), 134-144.
Wearden, J. H., Denovan, L., Fakhri, M. and Haworth, R. (1997) Scalar timing in temporal generalization in humans with longer stimulus durations. *Journal of Experimental Psychology-Animal Behavior Processes,* 23 (4), 502-511.
Wearden, J. H., Edwards, H., Fakhri, M. and 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 (2), 97-120.
Wearden, J. H. and Ferrara, A. (1995) Stimulus spacing effects in temporal bisection by humans. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology,* 48 (4), 289-310.
Wearden, J. H. and Lejeune, H. (2008b) Scalar properties in human timing: conformity and violations. *Quarterly Journal of Experimental Psychology,* 61 (4), 569-587.
Wearden, J. H., Norton, R., Martin, S. and Montford-Bebb, O. (2007) Internal clock processes and the filled-duration illusion. *Journal of Experimental Psychology-Human Perception and Performance,* 33 (3), 716-729.
Wearden, J. H., Todd, N. P. M. and Jones, L. A. (2006) When do auditory/visual differences in duration judgements occur? *Quarterly Journal of Experimental Psychology,* 59 (10), 1709-1724.
Webster, M. A. (2011) Adaptation and visual coding. *Journal of Vision,* 11 (5), doi:10.1167/11.5.3.
Webster, M. A., Kaping, D., Mizokami, Y. and Duhamel, P. (2004) Adaptation to natural facial categories. *Nature,* 428 (6982), 557-561.
Webster, M. A. and MacLin, O. H. (1999) Figural aftereffects in the perception of faces. *Psychonomic Bulletin and Review*, 6 (4), 647-653.
Welch, R. B. and Warren, D. H. (1980) Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin,* 88 (3), 638-667.
Westheimer, G. (1999) Discrimination of short time intervals by the human observer. *Experimental Brain Research,* 129 (1), 121-126.
Westheimer, G. and McKee, S. P. (1977) Spatial configurations for visual hyperacuity. *Vision Research,* 17 (8), 941-947.
Westrick, Z. M. and Landy, M. S. (2013) Pooling of first-order inputs in second-order vision. *Vision Research,* 91, 108-117.
Wiener, M., Kliot, D., Turkeltaub, P. E., Hamilton, R. H., Wolk, D. A. and Coslett, H. B. (2012) Parietal influence on temporal encoding indexed

by simultaneous transcranial magnetic stimulation and electroencephalography. *Journal of Neuroscience,* 32 (35), 12258-12267.
Wiener, M., Matell, M. S. and Coslett, H. B. (2011) Multiple mechanisms for temporal processing. *Frontiers in Integrative Neuroscience,* 5, doi:10.3389/fnint.2011.00031
Wiener, M., Turkeltaub, P. and Coslett, H. B. (2010) The image of time: a voxel-wise meta-analysis. *Neuroimage,* 49 (2), 1728-1740.
Wilson, C. J. (1995) The contribution of cortical neurons to the firing pattern of striatal spiny neurons. In: Houk, J. C., Davis, J. L. and Beiser, D. G. (eds). *Models of Information Processing in the Basal Ganglia*, Cambridge: MIT Press, 29-50.
Wilson, H. R., Ferrera, V. P. and Yo, C. (1992) A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience,* 9 (1), 79-97.
Wilson, H. R., Loffler, G. and Wilkinson, F. (2002) Synthetic faces, face cubes, and the geometry of face space. *Vision Research,* 42 (27), 2909-2923.
Wittmann, M. (2013) The inner sense of time: how the brain creates a representation of duration. *Nature Reviews Neuroscience,* 14 (3), 217-223.
Wohlgemuth, A. (1911) *On the after-effect of seen movement*. Cambridge: Cambridge University Press, 1-117
Xu, H., Dayan, P., Lipkin, R. M. and Qian, N. (2008) Adaptation across the cortical hierarchy: low-level curve adaptation affects high-level facial-expression judgments. *Journal of Neuroscience,* 28 (13), 3374-3383.
Xu, X., Ichida, J. M., Allison, J. D., Boyd, J. D., Bonds, A. B. and Casagrande, V. A. (2001) A comparison of koniocellular, magnocellular and parvocellular receptive field properties in the lateral geniculate nucleus of the owl monkey (*Aotus trivirgatus*). *The Journal of Physiology,* 531 (1), 203-218.
Xuan, B., Zhang, D., He, S. and Chen, X. (2007) Larger stimuli are judged to last longer. *Journal of Vision,* 7 (10), doi:10.1167/7.10.2
Xue, J. T., Ramoa, A. S., Carney, T. and Freeman, R. D. (1987) Binocular interaction in the dorsal lateral geniculate nucleus of the cat. *Experimental Brain Research,* 68 (2), 305-310.
Yamazaki, T. and Tanaka, S. (2005) Neural modeling of an internal clock. *Neural Computation,* 17 (5), 1032-1058.
Yates, M. J., Loetscher, T. and Nicholls, M. E. (2012) A generalized magnitude system for space, time, and quantity? A cautionary note. *Journal of Vision,* 12 (7), doi:10.1167/12.7.9
Yeshurun, Y. and Marom, G. (2008) Transient spatial attention and the perceived duration of brief visual events. *Visual Cognition,* 16 (6), 826-848.
Yue, X., Cassidy, B. S., Devaney, K. J., Holt, D. J. and Tootell, R. B. (2011) Lower-level stimulus features strongly influence responses in the fusiform face area. *Cerebral Cortex,* 21 (1), 35-47.
Zakay, D. (1993) Relative and absolute duration judgments under prospective and retrospective paradigms. *Attention, Perception and Psychophysics,* 54 (5), 656-664.

Zakay, D. and Block, R. A. (1995) An attentional gate model of prospective time estimation. In: Richelle, M., DeKeyser, V., d'Ydewalle, G. and Vandierendonck, A. (Eds), *Time and the dynamic control of behavior*, Liege, Belgium: Universite de Liege, 167-178.

Zakay, D. and Block, R. A. (1996) The role of attention in time estimation processes. In: Pastor, M. A. and Artieda, J. (eds) *Time, internal clocks and movement.* Vol. 115. North Holland: Elsevier Science. 143-164.

Zeiler, M. D. and Hoyert, M. S. (1989) Temporal reproduction. *Journal of the Experimental Analysis of Behavior,* 52 (2), 81-95.

Zeiler, M. D., Scott, G. K. and Hoyert, M. S. (1987) Optimal temporal differentiation. *Journal of the Experimental Analysis of Behavior,* 47 (2), 191-200.

Zeki, S., Watson, J., Lueck, C., Friston, K. J., Kennard, C. and Frackowiak, R. (1991) A direct demonstration of functional specialization in human visual cortex. *The Journal of Neuroscience,* 11 (3), 641-649.

Zeki, S. M. (1973) Colour coding in rhesus monkey prestriate cortex. *Brain Research,* 53 (2), 422-427.

Zeki, S. M. (1978) Uniformity and diversity of structure and function in rhesus monkey prestriate visual cortex. *Journal of Physiology,* 277, 273-290.

Zelkind, I. (1972) Factors in time estimation and a case for the internal clock. *The Journal of General Psychology,* 88 (2), 295-301.

Zhou, B., Yang, S., Mao, L. and Han, S. (2014) Visual feature processing in the early visual cortex affects duration perception. *Journal of Experimental Psychology: General,* 143 (5), 1893-1902

Zhou, Y. X. and Baker, C. L. (1993) A processing stream in mammalian visual cortex neurons for non-Fourier responses. *Science,* 261, 98-101.

Zhou, Y. X. and Baker, C. L. (1996) Spatial properties of envelope-responsive cells in area 17 and 18 neurons of the cat. *Journal of Neurophysiology,* 75 (3), 1038-1050.

Zihl, J., Von Cramon, D. and Mai, N. (1983) Selective disturbance of movement vision after bilateral brain damage. *Brain,* 106 (2), 313-340.