Bayesian Time Perception

Darren Rhodes

A thesis submitted to the University of Birmingham

for the degree of Doctor of Philosophy

September 2015

School of Psychology

College of Life and Environmental Sciences

University of Birmingham

UNIVERSITYOF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

Acknowledgments

ways. Steve and Phil get particular praise for football-related chat. I have loved working with you all and look forward to keeping in touch. Thank you.

I am from the Black Country, and as such, my friends help with bringing back down to Earth – thanks mates: Percie, bere Kingy, Pete, Tony, Batesy, Danny B, Shirg, TJ, Lank, The Pigeon Loft; Warner, Hicky, Zak, Joe, Tay, Nath, Paul, and also Graeme & Stu Brookes. But also those outside the Black Country and beyond: Sam, Dan, Pip, Rob, Jimm and Matt & Sara (and family).

I also have to thank Walsall F.C for keeping me positive throughout the whole PhD. Like a PhD, there are ups and downs when supporting the Saddlers – but we can always smile at the end. I would also like to thank Romaine Sawyers' brilliance for making weekends less painful when sat in a never-ending pile of programming or marking.

Thank you for the endless support Alan Rhodes, Jayne Rhodes, Chris Rhodes, Jamie Rhodes and all of my extended family. I wouldn't be on this journey if it were not for your love – so thanks a million! Finally, I would like to say thank you to Danielle Insley (also to Kevin, Bernie and Katie Insley too) – her love and support have kept me going through this experience and I look forward to the future with you.

Dedication

I would like to dedicate this thesis to my Grandparents: John & Kathleen Rhodes and

Beryl & Reginald Paskin.

Table of Contents

Bayesian Time Perception

List of Figures

List of Tables

Summary

Time is an elemental dimension of human perception, cognition and action. Perceiving time is essential for everyday activities and survival. Given such importance, innumerable studies have investigated the perception of time over the last 100 years, but the computational basis for the processing of temporal information remains unknown. This thesis aims to understand the mechanisms underlying the perceived timing of stimuli. We propose a novel *Bayesian* model of when stimuli are perceived, that is consistent with the predictive coding framework – such a perspective to how the brain deals with temporal information forms the core of this thesis. We theorize that that the brain takes prior expectations about when a stimulus might occur in the future (prior distribution) and combines it with current sensory evidence (likelihood function) in order to generate a percept of perceived timing (posterior distribution).

In *Chapter 1*, I will present a brief history of research and the methods used in time perception. I will then discuss the psychophysical approach to time, before discussing extant models of time perception and advancing inconsistencies between each account that this thesis aims to bridge the gap between. Further, I will introduce how time in the brain has been modelled using Bayesian Decision Theory. In *Chapters 2-4*, we use human psychophysics (experimental methods to quantify behaviour in the perceptual system) to show that the brain may bias perception to make slightly irregularly timed stimuli be reported as more regular.

To explain our findings, in *Chapter 2* we advance a Bayesian model of perceived timing of individual stimuli. We suggest that the brain uses temporal expectations to bias perception in a way that slightly earlier or later than expected stimuli are perceived closer to expectation. Critically, our model shows that regularly timed stimuli are perceptually

accelerated and thus moved away from expectation. The key to being able to explain this effect is by relaxing the assumption of normality that is employed in models of time perception – as the processing of time is necessarily bound by its anisotropy. In *Chapter 3*, we show how an environment of irregularity can cause regularly timed sequences to be perceived as irregular whilst *Chapter 4* shows how a change in the reliability of a signal can cause an increased attraction towards expectation. Then, in *Chapter 5*, I will summarize our work and then present an outlook for how our model can advance future research in temporal perception.

Chapter 1 General Introduction

Time is a fundamental variable that pervades all sensory, motor and cognitive processes. Organisms, such as human beings, must quantify time in order to survive and interact with the environment efficiently and successfully. Time is central to our everyday lives, from playing sports, speaking, dancing, singing, or playing music – to our sleep-wake cycle. Though an important dimension of perception, a slight unease may fill the reader when researchers refer to 'time perception'. The fields of colour, object, taste, olfactory, distance, speech and depth perception all investigate tangible physical properties, whereas the dimension of time is invisible and transient – in fact, one could ask whether time even exists at all – as the theories of relativity suggest that all moments in the past, present and future are equally real – rendering the specious present something of an illusion (Callender, 2010; Davies, 2002; Einstein, 1916; James, 1890).

1.1.1 Scales of Time

In contrast to senses such light that we can only perceive a limited spectrum of, time is perceived over a broad scale from microseconds to circadian rhythms. Circadian rhythms are based on 24-hour light/dark cycle due to the movement of the Earth in relation to the Sun, which helps control waking times, sleep times and metabolic fitness (Buhusi & Meck, 2005; Czeisler et al., 1999). At the millisecond range, time is critical for speech generation (Schirmer, 2004), recognition (Mauk & Buonomano, 2004) and motor control (Edwards, Alder, & Rose, 2002); whilst at the interval range, time is crucial for foraging behaviour (Henderson, Hurly, Bateson, & Healy, 2006; Meck, 2003), decision making (Brody, Hernández, Zainos, & Romo, 2003), sequential actions (Bortoletto, Cook, & Cunnington, 2011) and associative learning (Gallistel & Gibbon, 2000), and has been demonstrated in

many species of non-human animals, such as birds (Bateson & Kacelnik, 1997; Buhusi, Sasaki, & Meck, 2002; Henderson et al., 2006; Ohyama, Gibbon, Deich, & Balsam, 1999), rodents (Buhusi et al., 2002; Gallistel, King, & McDonald, 2004), fish (Drew, Zupan, Cooke, Couvillon, & Balsam, 2005), primates (Gribova, Donchin, Bergman, Vaadia, & de Oliveira, 2002; Janssen & Shadlen, 2005) as well as in human infants (Brannon, Roussel, Meck, & Woldorff, 2004) and adults (Church & Deluty, 1977; Gibbon, Church, & Meck, 1984).

Millisecond, interval and circadian scales are believed to have different (or competing) computational or neural mechanisms underpinning them (Buhusi & Meck, 2005; Ivry & Schlerf, 2008; Merchant & de Lafuente, 2014). This thesis focuses on human behaviour and perception in the hundreds of milliseconds scale and as such I will later describe historical accounts of how the brain may deal with interval timing on such a scale. 'Timing' can mean both *how long* an event lasted (the duration of an interval delimited by two stimuli), or conversely, *when* an event transpired (Merchant & de Lafuente, 2014). All of the extant models of temporal perception are concerned with *how long* the perceived duration between two events is. The central aim of this thesis, however, is to elucidate how the brain may estimate *when* an event occurred in the world. I will firstly discuss the methods employed in time perception research before describing how current models can explain temporal processing. I will then introduce recent research that suggests the brain uses a Bayesian inferential processing approach to estimate interval timing.

1.2   General Methods in Time Perception

If a mechanism for time perception exists in the brain – what should it do? One might argue that an optimal mechanism would try to perceive time as close to veridical (physical) time as possible. Thus, the two main dependent variables in time perception research historically concern the *mean accuracy* and *variability* of temporal estimates. Estimates of

a temporal characteristic, such as the duration of an event, are prone to temporal distortions by stimulus properties (Horr & Di Luca, 2015a; 2015b; Thomas & Brown, 1974; Wearden, Norton, Martin, & Montford-Bebb, 2007), complexity (Schiffman & Bobko, 1977), sensory modality (Goldstone & Lhamon, 1974; Wearden, Edwards, Fakhri, & Percival, 1998; Wearden, Todd, & Jones, 2006), and context (Dyjas & Ulrich, 2014); and as such, the mean accuracy of an estimate deviates from real time. Whilst the mean accuracy may approximate real time, the system may be poor and as such the variability in the system may sometimes lead to experiencing an event as shorter or longer than the physical duration (Grondin, 2010). Thus, some researchers are concerned with measuring and reducing the variability in temporal estimates which has been classically researched using finger tapping (see, Repp, 2005; Repp & Su, 2013, for a review).

1.2.1 Prospective or Retrospective Tasks?

An explicit distinction exists between *prospective* and *retrospective* timing tasks (Block & Zakay, 1997; Brown & Stubbs, 1988; 1992; Tobin, Bisson, & Grondin, 2010). Historically, prospective timing tasks involve the subject knowing that the experiment or trial is explicitly about timing – an example of which is asking a participant to determine the temporal order between two stimuli. Conversely, in retrospective timing tasks, the subject is not aware that the experiment is about time. For example, a subject may watch a video or listen to a series of events, but only after the task will a temporal judgment be made, such as 'how much time passed watching the video?'.

Though it may be argued that retrospective timing tasks are more ecologically valid – as we rarely monitor the timing or duration of events at each moment in time – retrospective timing tasks are considered to involve memory-based aspects of cognition whereas prospective timing tasks are grounded more in the perceptual processes concerning the millisecond to second range of time, and as such are less affected by

higher-order processing (Brown & Stubbs, 1988; Grondin, 2010). The experimental work in this thesis deals with prospective timing tasks as it allows a purer method of investigation into the perception of *when* an event occurred, as well as being less corrupted by higher-order processing.

1.2.2 Traditional Methods of Investigation

Traditionally, psychologists have mainly used four methods to investigate the perception of time – but explicitly to study duration perception (Bindra & Waksberg, 1956; Block & Gruber, 2014; Grondin, 2010; Hancock & Block, 2012; Wallace & Rabin, 1960; Zakay, 1993). The first two tasks involve producing or reproducing a target interval (Goldstone, 1968). In a temporal *reproduction* task, a participant is asked to reproduce the duration of a continuous or delimited interval, whereas in a *production* task, an investigator tells the participant to push a button when the target interval has been reproduced or tap twice to mark the start and end of the interval. The third method is *verbal estimation*, where a target interval is presented to a participant and they are subsequently required to report in seconds or minutes the length of this duration (Vierordt, 1868).

The fourth method, the *method of comparison*, requires participants to respond whether a standard or comparison interval is shorter (or longer) than the other. In general, the standard interval remains constant over the experiment but the comparison interval changes in duration from trial-to-trial though the intervals are either delimited by a stimulus (empty intervals), but could also be continuous (filled interval) sounds or flashes (Bald, Berrien, Price, & Sprague, 1942; Dinnerstein & Zlotogura, 1968; Hamlin, 1895; Höring, 1864; Spence, Shore, & Klein, 2001; Zampini, Shore, & Spence, 2003). Using this method, some studies have highlighted how temporal characteristics of the experiment can reliably change perception. For example, if the first interval of two empty intervals is very brief (< 250ms) and the temporal distance between them is less than 100ms, then the

second interval is perceived as being much shorter than the first (Arao, Suetomi, & Nakajima, 2000). This effect, the *time-shrinking illusion*, is an illusion of time like another reported effect – the *filled duration illusion* – where filled intervals are perceived longer than empty ones (Adams, 1977; Buffardi, 1971; Horr & Di Luca, 2015a; Thomas & Brown, 1974; Wearden et al., 2007). Participants may also be asked to report whether the final (comparison) interval in a sequence of more than one standard interval is perceived to be shorter or longer. In this methodology, it has been observed that the ability to discriminate the durations increases as a function of the amount of stimuli in a sequence (Drake & Botte, 1993; ten Hoopen, Van Den Berg, Memelink, Bocanegra, & Boon, 2011; Miller & McAuley, 2005).

In the forced-choice version of the method of comparison, the standard interval is presented first and followed by the comparison interval (or vice versa), but the participant must always respond whether the comparison stimulus was either shorter or longer than the standard (a *Two-Alternative-Forced-Choice*, 2AFC) – even if the participant perceives the two intervals as being the same duration. Interestingly, it has been advocated that duration discrimination performance – the ability to successfully differentiate whether the standard or comparison is longer – is better when the standard is presented first (Dyjas & Ulrich, 2014; Grondin & McAuley, 2009; Lapid, Ulrich, & Rammsayer, 2008).

1.2.3 From Perceived Duration to Perceived Timing

To this point, I have described several methods of psychological inquiry that seek to understand how the brain may perceive time. Temporal reproduction and production, verbal estimation and the method of comparison have been used classically to assess the perceived duration of events. The method of comparison involves discriminating whether a standard interval is shorter or longer in duration than a comparison interval. Of central interest to this thesis, is the perception of *when* an event occurs rather than how long an

event lasts. In order to understand how we could measure the perceived timing of a stimulus, I will briefly introduce the reader to *psychophysics* and how the method of comparison can be used to estimate when a stimulus is perceived.

1.3   Perceived Timing & Psychophysics

1.3.1 Psychophysical Methods

Psychophysics is the scientific investigation of the functional interrelations between the physical and phenomenal world (W. H. Ehrenstein & Ehrenstein, 1999; Fechner, 1860). The aim of psychophysics is to quantify and measure subjective experience by determining the relationship between perception and physical stimuli. Such as in the method of comparison outlined above, a central tenet of modern psychophysics is to control and vary the properties of an external stimulus and then ask a participant to report what they have experienced – with as simple a question as possible. For example, one may be interested in the *detection* of whether a sound is present or not (i.e. did you hear that stimulus?) or, further, in *identifying* what kind of stimulus characteristic is present (i.e. where was the stimulus). As such we can translate detection into the sensing of a stimulus – and identification as a higher-level process that can sometimes result in a failure to identify a stimulus. For example, if a stimulus is weak and noisy, it may be sensed but a participant may be unable to identify or report a characteristic associated with it. To understand this sort of effect, Psychophysicists may ask a participant to report when they can hear an auditory stimulus under different levels of background noise. This task is one of *discrimination* –where participants must determine whether what they are experiencing is noise, or whether there is a signal present (W. H. Ehrenstein & Ehrenstein, 1999).

The method of comparison is the psychophysical methodology that is most suitable to assess the perceived timing of events. By presenting two signals separated by varying

stimulus onset asynchronies (SOAs) and consequently asking a participant to make a judgement about some characteristic or relationship between the two stimuli, such as, for example, 'were the stimuli simultaneous?' or, 'which stimulus was first?', it is possible to determine the point at which stimuli are perceived as being simultaneous – that is, their relative perceived timing. I will now briefly introduce the notion of synchrony and temporal order between the senses before introducing contemporary models of how the brain may deal with time perception.

1.3.2 Intersensory Synchrony and Temporal Order

We live in a multisensory environment where perception is not simultaneous – it takes time. The perception of synchrony or temporal order is not straightforward, as differences in neural and physical transmission times can cause synchronous events to be perceived as asynchronous, and vice versa. When a distant bolt of lightning illuminates the sky at night and sends out thunderous sound waves, we see the light first and then hear the sound even though both signals were emitted simultaneously. The discrepancy in the perception of a simultaneous multisensory event is due to the relative differences in sensory registration to the eyes and ears as light travels much quicker than sound (300,000,000 vs. 330 metres per second). To complicate matters further, the processing time for visual stimuli (approx. 50ms) is longer than auditory stimuli (approx. 10ms) as the chemical transduction of light in the retina is slower than the mechanical transduction of sound waves in the ear (Allison, Matsumiya, Goff, & Goff, 1977; A. J. King, 2005; Spence & Squire, 2003; Vroomen & Keetels, 2010). The distance at which the differences in neural and physical transmission times are negated and signals arrive at the primary sensory cortices synchronously is around 10-15 metres away from the observer and has been called the horizon of simultaneity (Spence & Squire, 2003; Vroomen & Keetels, 2010). However, in interactions between a human observer and a sound/light emitting device at a close

distance (~1-3 metres), it has been commonly reported that visual signals have to precede auditory signals for the perception of simultaneity (Vroomen & Keetels, 2010; Zampini et al., 2003; Zampini, Guest, Shore, & Spence, 2005a; Zampini, Shore, & Spence, 2005b).

The temporal difference between the senses is measured by finding the asynchrony necessary to perceive simultaneity, which is defined as the *Point of Subjective Simultaneity* (PSS). To measure this difference, one can use the psychophysical methodology. An extension of simply discriminating whether a signal is present or not, is to present two stimuli (*X* and *Y*) with varying SOAs (*X-Y*) and force participants to report whether the two stimuli are simultaneous (Exner, 1875; Fujisaki, Shimojo, Kashino, & Nishida, 2004; Spence et al., 2001; Zampini, Guest, Shore, & Spence, 2005a; Zampini, Shore, & Spence, 2005b), or to report the temporal order of the pair (Boenke, Deliano, & Ohl, 2009; Gibbon & Rutschmann, 1969; Jaśkowski, 1992; Yamamoto & Kitazawa, 2001; Zampini et al., 2003).

In the *Simultaneity Judgment* (SJ) task, participants judge whether *X* and *Y* appear to be simultaneous – or not. Here, the proportions of 'simultaneous' responses are plotted as a function of SOA, which usually results in a Gaussian/Bell shaped curve where the PSS is taken as the peak of the distribution. Here, the assumption is that the peak represents perceived simultaneity, as this is the point at which participants are maximally sure that *X* and *Y* are synchronous. A further measure than can be derived from such a function is the standard deviation (SD) of the distribution of responses. As such, it characterizes a window of temporal integration, as it represents the range of SOAs at which *X* and *Y* are considered as belonging to the same event (Vroomen & Keetels, 2010).

In temporal order judgments (TOJs), the proportion of '*Y* first' responses is generally an increasing function of SOA. One usually obtains a sigmoid function where the PSS corresponds to the SOA at which an observer is maximally unsure about the

temporal order of the pair of stimuli (50% point). The steepness of the curve at the PSS reflects an observers' sensitivity to temporal order and is expressed as the *Just-Noticeable Difference* (JND). Generally this measure is taken as half of the difference between the SOA at the 25% and 75% points, however other methods such as the Spearman-Kärber may calculate this based on the 14% and 86% points (two sigma; see Miller & Ulrich, 2001). As such, the JND represents the smallest SOA an observer can reliably judge the temporal order thereof. A flat curve would result in a relatively larger JND and as such reflect an observer that has low temporal sensitivity whereas a steep curve would constitute a smaller JND and thus implies an observer has higher temporal sensitivity.

1.3.3 Estimating Perceived Timing using Psychophysics

I have discussed the psychophysical method and how one can measure the relative timing between two sensory events. In this thesis, I will use psychophysics to estimate the perceived timing of an event through the PSS. However, the PSS is only really a measure of the relative asynchrony in the time it takes it process two signals to be perceived as simultaneous – not *when* an event happened. To measure the perceived timing of a stimulus, we present a sequence of regularly timed stimuli and pair the last stimulus with a stimulus from another modality (which is unaffected by the sequence) to compare the PSS for stimuli presented on time, earlier than and later than expected. Presently, models of time perception do not predict that the PSS should change regardless of when a stimulus is presented. In the next section I will discuss such models and their predictions before introducing our *Bayesian model of perceived timing* that makes explicit predictions that will be tested in Chapters 2-4.

1.4   Contemporary Models of Time Perception

The aim of this thesis is to understand the computational mechanisms of how the brain may estimate the perceived timing of events – that is, how can the brain know when is now, when was then and when is next? Extant models of time perception are fundamentally based on the notion of perceived duration – that is, how the brain may represent and encode the time between two signals. I will now introduce and discuss such contemporary models of interval timing. Firstly, it should be addressed however, that there exists a great literature on different taxonomies of timing models – where some researchers have conceptualised models of time in terms of having a dedicated neural mechanism for the perception of time (Creelman, 1962; Gibbon, 1977; Gibbon et al., 1984; Treisman, 1963; Wing & Kristofferson, 1973), in contrast to time being an intrinsic product of sensory information processing, where recurrent spatial or activity patterns read out duration without the need of an internal clock (Buonomano, 2009; Buonomano & Merzenich, 1995; Karmarkar & Buonomano, 2007; Mauk & Buonomano, 2004). Further, dedicated models assume that there are specialized brain regions involved in the representation of temporal information, whilst intrinsic models primarily argue for a distributed timing mechanism over the brain (Ivry & Schlerf, 2008). The most notable neural instantiation of such a specialized-timing view, is the cerebellar timing hypothesis (Ivry, Spencer, Zelaznik, & Diedrichsen, 2002). The cerebellum has been implicated in the representation of time, as patients with cerebellar pathology have relatively poor performance on tasks that involve simply judging the duration between two auditory tones (Ivry & Keele, 1989; Mangels, Ivry, & Shimizu, 1998). In contrast, temporal information can be understood as the result of activity across a network of cortical regions (Lewis & Miall, 2003). However, this thesis is concerned with two popular classes of dedicated models for the perception of time: *Entrainment* and *interval* models (Gibbon et al., 1984;

Large & Jones, 1999; McAuley & Jones, 2003), and as such, I will now introduce both before showing how they can be formulated to make predictions about the timing of individual stimuli.

1.4.1 Interval Models

When one is asked 'what time is it?' or 'how long have you been waiting?' – it is quite likely that this person will glance at their watch and use it to estimate what the present time is – or how long the wait has been. As such, it is intuitive to think that the brain may use a clock-like mechanism in order to deal with the perception of time. Interval models of timing are born out of this analogy and they conceive time as a triad of clock, memory and decision processes (Creelman, 1962; Treisman, 1963). The most notable, and influential interval model is *Scalar Expectancy Theory* (SET; Church, Meck, & Gibbon, 1994; Gibbon, 1977; Gibbon et al., 1984). In the SET model, the internal clock is considered as a pacemaker-accumulator mechanism, where a dedicated pacemaker emits pulses continuously. To represent duration, the accumulator counts the amount of pulses between two signals and then stores them in memory (Figure 1.1). The hallmark of the SET model is that as the mean duration of an interval increases, the associated standard deviation of the duration estimate increases linearly also – this is often called the 'scalar property' of interval timing. The scalar property is synonymous with the Weber-Fechner Law (Fechner, 1860), which asserts a logarithmic relationship between physical magnitudes and the representation in the perceptual system, and as such, the JND between two physical magnitudes is proportional to the absolute physical magnitude. Each interval is maintained in working memory before being passed to a more robust representation in long-term memory. The key point here is that time, in these accounts, is represented as discrete interval durations that are subsequently compared with other intervals at a decision stage (Allman, Teki, Griffiths, & Meck, 2013; Church & Broadbent, 1990; Gibbon et al., 1984).

intervals. The interval models are also symmetric in the sense that they do not predict any differences in the detection of temporal irregularities due to the anisochrony a stimulus is presented. For example, if a stimulus is presented earlier than expected – then there should be no difference in the discrimination of its irregularity compared to a stimulus presented later than expected (Drake & Botte, 1993; ten Hoopen et al., 2011; Miller & McAuley, 2005). We test the predictions of interval models in Chapter 2, where we find that, contrary to current belief, that there is a difference in performance of the detection of irregularity due to the anisochrony at which a stimulus is presented. Further, we find differences in the perceived timing of stimuli as a function of their relation to expectation – early stimuli are perceptually delayed whilst late stimuli are perceptually accelerated in order to appear closer to expectation. Interestingly, we find that stimuli presented isochronously (on-time) are also perceptually accelerated. Interval models of time perception cannot account for these patterns of results however entrainment models can be formulated to explain only the acceleration of stimuli presented isochronously.

1.4.2 Entrainment Models

Entrainment models offer an alternative realisation of interval timing. Similar to interval models, the basic tenet of these models is that a dedicated pacemaker is an entrainable oscillator that peaks in amplitude at the expected onset of future stimuli (Large & Jones, 1999; Large & Palmer, 2002; Large & Snyder, 2009; McAuley & Jones, 2003) – though phase coincidence (Miall, 1989), or the recurrence of activity patterns (Buonomano, 2009; Buonomano & Merzenich, 1995; Karmarkar & Buonomano, 2007) have also been proposed as alternative intrinsic entrainment models. Whilst interval models have mainly been formulated to explain interval timing and determining which of two intervals is longer (or shorter) – entrainment models are more conducive to also explain stimulus

timing in rhythmic sequences – as internal oscillations gradually adjust to the phase of external rhythms.

A prominent entrainment-based model of how the brain deals with rhythmic sequences is *Dynamic Attending Theory* (DAT) (Jones & Boltz, 1989; Large & Jones, 1999; Large & Palmer, 2002). Here, attention is not distributed evenly over time, but rather ebbs and flows with time's passing. Originally proposed as a model of rhythmic expectancy, DAT proposes that rhythm perception is induced by way of entrainment to external signals. Internal fluctuations in attentional energy (attentional 'peaks') generate temporal expectancies about the onset of future events that can acclimate to the period and phase of external events by way of an adaptive internal oscillator (Figure 1.2). At the neural level, the perception of regular events has been proposed to originate from neural oscillations that adjust and resonate with external signals (Henry & Herrmann, 2014; Large & Snyder, 2009; Zanto, Snyder, & Large, 2006). The framework of active sensing (Schroeder & Lakatos, 2009; Schroeder, Wilson, Radman, Scharfman, & Lakatos, 2010) – the fluctuation of excitation/inhibition cycles– can be tied directly to DAT. The high excitability phase of neural oscillations are thought to be associated with the peak of the attentional pulse and as such facilitate sensory selection and processing of stimuli that coincide with the peak of an oscillation (Henry & Herrmann, 2014; Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008). Therefore, one can reason that if a stimulus occurs at the peak of an oscillation and high excitability phase, then it should be given a perceptual boost and processed faster. This effect, is similar to *prior entry* (Spence & Parise, 2010; Sternberg, Knoll, & Gates, 1971), where attended stimuli are processed quicker than unattended ones. The idea of prioritized processing of attended stimuli exists in the visual cognition domain (Summerfield & Egner, 2009), and such attentional facilitation of perception has been highlighted in a number of studies in the temporal processing

literature (Spence et al., 2001; Sternberg & Knoll, 1973; Zampini, Shore, & Spence, 2005b) as well as at the neural level (J. J. J. McDonald, Teder-Sälejärvi, Di Russo, & Hillyard, 2005).

DAT accounts for perceived stimulus timing by considering that humans detect asynchronies between an expected stimulus onset time and the actual stimulus onset time (McAuley, 1995). If the stimulus onset occurs after the expected peak then a stimulus is perceived as being late whilst if it is before the expected peak then it is perceived as being early. Intuitively, when a stimulus onset time coincides with the peak of the expected time, then it is perceived as being on time – though as shown above, entrainment models could be formulated to predict an acceleration of attended-to stimuli that occur at the peak of an oscillation. As a consequence of increasing attentional expectancies due to entrainment, sensitivity to temporal deviations improves as a function of increasing sequence length (Barnes & Jones, 2000; Drake & Botte, 1993; McAuley & Kidd, 1998; Miller & McAuley, 2005).
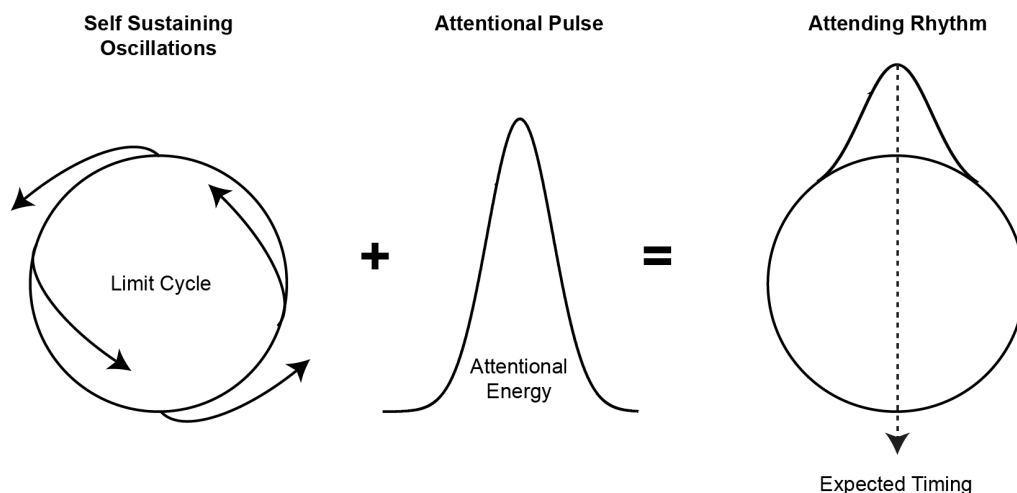
Figure 1.2. A schematic overview of Dynamic Attending Theory. An adaptive internal oscillator is a dynamic system that periodically generates temporal expectancies (Jones & Boltz, 1989; Large & Jones, 1999; Large & Palmer, 2002). The oscillations coupled with pulses of attentional energy at (recurrent) expected time points, given the phase of a rhythm, result in attention being allocated at the expected time-

point. Discrepancies between he onset times of a stimulus in relation to its expected onset gives rise to the detection of temporal irregularities.

Extant *Bayesian* models of time perception have been formulated (Jazayeri & Shadlen, 2010; Miyazaki, Nozaki, & Nakajima, 2005; Shi, Church, & Meck, 2013) – but only for the representation of interval duration. I will now introduce the idea of Bayesian time perception for duration perception before briefly giving an overview of our proposal of a Bayesian account of perceived timing.

1.4.3 A Bayesian Model of Interval Timing

As mentioned previously, time is subject to various contextual distortions. A seminal example of contextual calibration is Vierordt's law (Lejeune & Wearden, 2009; Vierordt, 1868). When observers are presented with various intervals of different lengths and subsequently asked to reproduce each interval – they tend to overestimate the duration of short intervals, and underestimate long ones. This is a type of 'central-tendency' effect – participants migrate their estimates of duration towards the mean of exposed intervals. A prevalent model of such an effect is that the perception of interval duration is derived from not only the perception of current sensory information, but also from the prior knowledge of the duration of previously exposed intervals (Jazayeri & Shadlen, 2010; Lejeune & Wearden, 2009; Petzschner & Glasauer, 2011; Petzschner, Glasauer, & Stephan, 2015; Taatgen & van Rijn, 2011). Prior knowledge of the temporal statistics of the environment, in this sense, biases temporal perception.

A suitable candidate to explain the central-tendency effect observed in time perception is the *Bayesian* framework (Bayes, 1763). Bayesian models of perception have been successfully used to model several perceptual domains (Ernst, 2006; Ernst & Banks, 2002; Ernst & Bülthoff, 2004; Knill, 2007; Knill & Richards, 1996; Maloney &

Mamassian, 2009; Mamassian, Landy, & Maloney, 2002) and have been applied to

duration estimation (Hartcher-O'Brien, Di Luca, & Ernst, 2014; Shi et al., 2013) and

reproduction (Jazayeri & Shadlen, 2010; Miyazaki et al., 2005). In the Bayesian

framework, a generative model combines current sensory information (likelihood) with *a*

priori knowledge of the world (prior) in order to give rise to a percept (posterior). The

likelihood and prior in this model are weighted by their relative uncertainties (Colas,

Diard, & Bessiere, 2010; Fernandes, Stevenson, Vilares, & Körding, 2014; Griffiths &

Tenenbaum, 2011; Lucas & Griffiths, 2009; Vilares & Körding, 2011).

The Bayesian framework has recently been applied to the SET model of interval

timing (Shi et al., 2013). The central tenet of such a Bayesian model of interval timing is

that the triad of components of the SET model are translated into the Bayesian framework:

the likelihood, prior and posterior are considered analogous to the clock, memory and

decision stages (Figure 1.1). The clock stage represents the *likelihood function*, and is

rendered as such: if an interval delimited by two stimuli is duration *D*, with an allied

internal clock count of *C*, which represents the number of 'ticks' accumulated by the time

the second stimulus has delimited the interval, then the likelihood function $P^l(C\,|D)$, is the

probability of acquiring the perceived duration *C*, given the external stimulation *D*. The

width of the likelihood probability distribution indicates the relative sensory uncertainty

given the measurement – a steep function, for instance, would give a likelihood function

with little uncertainty about the duration observed *D*, whilst a flatter function would

indicate a likelihood function with great uncertainty about *D*.

The memory stage is analogous to the *prior probability distribution $P^p(D)$*.

Sensory estimates from the likelihood update the prior so to represent the distribution of

learned intervals. In order to arrive at an estimate of perceived duration, according to

Bayes' rule, the prior is combined with the likelihood, in order to form the *posterior distribution $P^q(D|C)$*:

$$P^q(D \mid C) = P^l(C \mid D) \cdot P^p(D) \qquad \text{(Eq. 1.1)}$$

The posterior distribution is considered as synonymous to the memory stage of the SET model. Given the posterior, a Bayesian ideal observer chooses an action given a loss function that specifies the relative cost or success of a potential behavioural response (Acerbi, Vijayakumar, & Wolpert, 2014; Acerbi, Wolpert, & Vijayakumar, 2012; Körding & Wolpert, 2004; Wolpert, 2007). The model predicts noisy sensory estimates of duration are biased towards the mean of the prior probability distribution.

1.4.4 Summary of Models

In summary, interval models of duration perception are based on the idea that an internal clock keeps track of time by counting the amount of pulses between the onsets of one event to another. When considering the perceived timing of a single stimulus, these models make no predictions about changes in the timing of a stimulus due to its anisochrony. Entrainment models, on the other hand, can be formulated to predict that expected stimuli are processed faster and as such, perceived earlier. In contrast to these accounts of time perception, the Bayesian framework has been applied to several perceptual domains, and has recently been applied to duration estimation (Hartcher-O'Brien et al., 2014; Shi et al., 2013). The Bayesian framework has been used to show how the representation of duration is calibrated in order to make intervals appear more similar to the duration of previously exposed ones (a central tendency effect). The likelihood function is similar to the clock stage of the SET interval-based model – the clock is responsible for the measurement stage of inferring the duration of an external event. The prior is akin to the long-term reference and memory stages of the SET model and as such represents the learned knowledge of the average durations experienced. The posterior distribution represents a percept and an

observer chooses a response after a decision rule, which is similar to the decision stage of the SET model. The model is useful in connecting the computational principles of Bayesian modelling with the information-processing account of duration perception of interval models. However, as with other interval-based models – the Bayesian account of time perception (described above) only makes predictions about what happens to the representation of intervals, and as such, does not predict any changes to the perceived timing of stimuli in sequences.

1.5 Shifting Focus from Perceived Duration to Perceived Event Timing

Interval and entrainment models were born out of modelling the perception of duration. Both models predict that changes in the detection of irregularity should be symmetrical – whether a stimulus is earlier or later than expected, there should be no difference in the detection of irregularity. Neither model has been formulated to deal with changes in the perceived timing of individual stimuli in a sequence. In the following Chapters we test such predictions – but also formulate our own model that moves perceived timing away from the representation of duration – towards the representation of event timing. We propose that a Bayesian inferential process happens at every point in time, where current sensory evidence is combined iteratively with expectations of the future timing of stimuli. Time, in this sense, is characterized by an on-line sensory continuum that represents each point in time. Such a model predicts that perceived timing is biased in a way that regularises slightly irregular stimuli towards the expected time point. Extant *Bayesian* models of time perception have been formulated (see Section 1.4.3; Jazayeri & Shadlen, 2010; Miyazaki et al., 2005; Shi et al., 2013) – but only for the representation of intervals.

The following Chapter (2) lays the foundation of this thesis and introduces our Bayesian model of perceived timing.

Chapter 2 A Bayesian Model of Perceived Timing

In this paper, we tested if temporal expectations could bias the perceived timing of stimuli. We present four experiments that show how the brain may regularize stimuli presented slightly too early or too late. We introduce a Bayesian model of perceived timing with dynamic priors and asymmetric likelihood functions that explains the perceptual phenomena we observe.

2.1   Abstract

The environment has a temporal structure and knowing when a stimulus will appear translates into increased perceptual performance. Here, we investigated how the human brain exploits temporal regularity in stimulus sequences for perception. We find that the timing of stimuli that occasionally deviate from a regularly-paced sequence is perceptually distorted. Stimuli presented *earlier* than expected are perceptually delayed, whereas stimuli presented *on time* and *later* than expected are perceptually accelerated. This result suggests that the brain regularizes slightly deviant stimuli in a Bayesian-optimal fashion by combining expectations in the form of a-priori probability of encountering future stimuli with incoming sensory information. The asymmetry in the temporal shift, which leads to speedup in the processing of expected stimuli, is due to the shape of the prior probability distribution.

2.2   Introduction

Events in our perceptual world often have a predictable temporal structure. Exploiting temporal regularities can decrease metabolic consumption (VanRullen & Dubois, 2011) and automatize behaviour for rhythmic activities such as dance, locomotion, speech, and music production (McNeill, 1995; Repp, 2005). Predictable timing of events leads to improved stimulus detection and discrimination (Brochard, Tassin, & Zagar, 2013; Carnevale, de Lafuente, Romo, Barak, & Parga, 2015; Correa, Lupiáñez, & Tudela, 2005; Cravo, Rohenkohl, Wyart, & Nobre, 2013; Escoffier, Sheng, & Schirmer, 2010; Jazayeri & Shadlen, 2010; Rohenkohl & Nobre, 2011), perceptual changes (Kok, Brouwer, van Gerven, & de Lange, 2013), and faster responses (Jazayeri & Shadlen, 2010; Lakatos et al., 2008; Miyazaki et al., 2005). The neural mechanisms behind these perceptual phenomena are unclear. Here we propose to use one of the simplest types of stimulus regularities, the occurrence of stimuli after equal intervals of time (*isochrony*), to quickly generate the expectation for a successive stimulus. We investigate how perceived timing changes due to such an expectation.

There are several schools of thought about how the brain deals with the regularity of stimulus sequences. *Interval-based models* assert that the time between two stimuli is represented as a discrete interval duration that is compared with subsequent intervals (Church et al., 1994; Creelman, 1962; Gibbon, 1977; Gibbon et al., 1984; Treisman, 1963). The representation is refined with more stimuli presented, leading to increased performance (Drake & Botte, 1993; Miller & McAuley, 2005; Schulze, 1989). In contrast, *entrainment models* advocate that the phase and frequency of temporal patterns is the important aspect. The dynamics of attending to stimuli, for example, has been shown to adjusts to rhythmic external stimulation (Jones & Boltz, 1989; Large & Jones, 1999; Large & Palmer, 2002; Large & Snyder, 2009). At a neural level, phase coincidence (Miall, 1989)

and activity patterns (Buonomano & Merzenich, 1995; Karmarkar & Buonomano, 2007) progressively tune to the phase and frequency of rhythmic stimulus sequences. Exogenous attention is then deployed at the expected time (Henry & Herrmann, 2014; Rohenkohl, Coull, & Nobre, 2011). Interestingly, attention and expectation have opposite effects on neural responses, where expectation reduces neural responses (Auksztulewicz & Friston, 2015). Such reduction is accounted for by the free energy principle (Friston, 2005; 2008), according to which the brain continuously predicts stimulation and thus increases in neural activity represent deviations from expectations.

For all these approaches, sensitivity to temporal irregularities should increase as a function of the number of stimuli composing a sequence (Large & Palmer, 2002; Schulze, 1989). Moreover, according to interval-based models, the presence of a stimulus sequence should not have an influence on the perceived timing of individual stimuli. Entrainment (Lakatos et al., 2008) and predictive coding models which are based on time-point representations (Friston, 2005; 2008), on the other hand, predict that expected stimuli could be perceived earlier than unexpected ones, a phenomenon called prior entry (Spence & Parise, 2010). Here we reason that if a regular sequence leads to the expectation of a stimulus, not only stimuli presented when expected, but also stimuli presented later on should be also expected and thus should be perceptually accelerated (Nobre, Correa, & Coull, 2007). On the contrary, several approaches have advanced that there should be no difference in the perceived timing of early and late stimuli (Barnes & Jones, 2000; Henry & Herrmann, 2014; McAuley & Jones, 2003). Here we want to disambiguate such predictions and characterize in which way presenting stimuli in a sequence influences perceived timing.

Extant Bayesian accounts of time perception (Jazayeri & Shadlen, 2010; Shi et al., 2013) do not account for changes in perceived timing of individual stimuli as they are based on the representation of intervals. We propose that a Bayesian inferential process happens not only at the interval level, but also at the individual stimulus timing level. Perception is obtained at each point in time through the iterative combination of incoming sensory information with expectations of a stimulus based on previous intervals. A fundamental aspect of our model is that – differently from what happens at the interval level (Jazayeri & Shadlen, 2010; Shi et al., 2013) – probability distributions about timing are *asymmetric* due to the way time flows. This asymmetry explains perceptual acceleration of expected stimuli as a more tuned neural response with a sharp onset for stimuli presented at expectation.

2.3   Results

2.3.1 Behavioural Results: Asymmetric Temporal Deviation Detection

In Experiment 1 participants judged whether the timing of the last stimulus in a sequence was regular or irregular. As one would expect, Figure 2.1C,D shows that the proportion of sequences reported to be regular decreases with large anisochronies, and such a pattern is more marked for audio sequences, reflecting the higher reliability of temporal judgments with auditory stimuli. As described in the literature, sensitivity to temporal deviations increases with longer sequences (Drake & Botte, 1993; Large & Palmer, 2002), but here we find that changes are accentuated for stimuli presented earlier than expected – an asymmetry in anisochrony detection (Figure 2.1C,D).
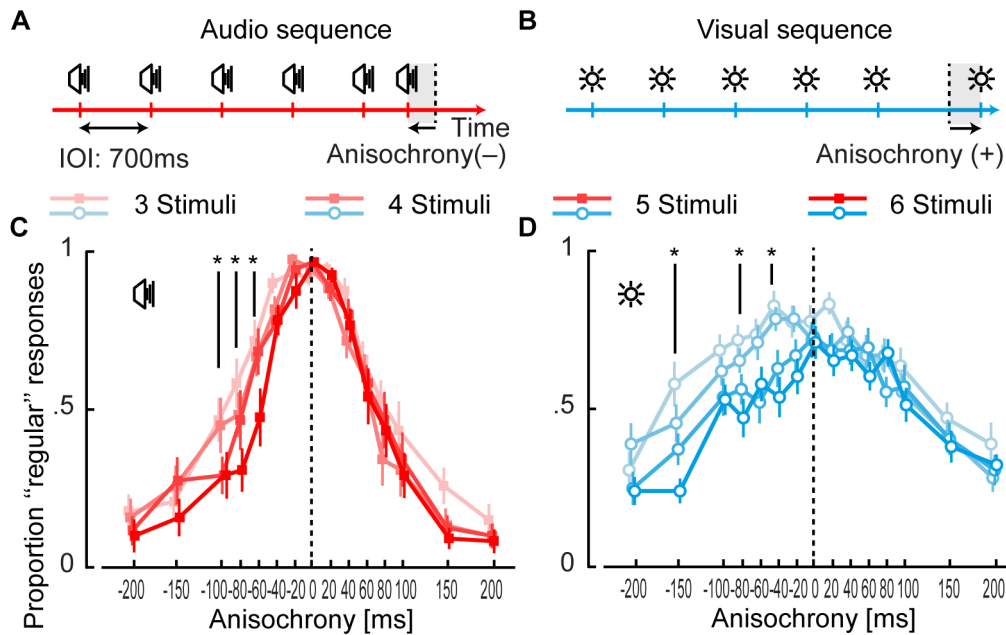
Figure 2.1. Stimuli and results of Experiment 1. Each participant is only presented with sequences of stimuli in one modality. (A) Example of an audio sequence with the final stimulus presented early. (B) Example of a visual sequence with the final stimulus late. (C) Proportion of "regular" responses as a function of the anisochrony of the last stimulus in an audio sequence and (D) in a visual sequence. Each line represents data obtained with a different sequence length. The distribution of responses is steeper with longer sequences (interaction term of a two-way repeated measures (r.m.) ANOVA on the inverse-normal proportion of "regular" responses bounded between .01 and .99; audio: $F(42,588)=1.8$, $p=.0016$, $\eta p^2=.11$; visual: $F(42,588)=1.5$, $p=.0135$, $\eta p^2=.10$). Asterisks denote anisochronies at which the proportions of responses significantly differ across the four sequence lengths (one-way r.m. ANOVA Bonferroni corrected, $p<.0033$). In all graphs, error bars represent the standard error of the mean. Supplementary Figure 2.1 shows the data of Experiment 1S where sequence duration was kept constant within a block of trials. In all graphs, error bars represent the standard error of the mean.

2.3.2 Behavioural Results: Changes in Perceived Timing

In order to measure whether the asymmetry in isochrony judgments is related to a change in the perceived timing of stimuli, in Experiment 2 we employed a novel experimental paradigm where the last stimulus in a 4-stimulus sequence is paired with a stimulus in another modality and participants reported the temporal order of this audiovisual pair

(Figure 2.2A). From participants' responses we could determine the audiovisual

asynchrony necessary for the Perception of Subjective Simultaneity (PSS). The last

stimulus in the sequence could be presented on time or anisochronously (earlier or later

than expected). If we assume that the sequence is more likely to exert an influence on

stimuli of the same modality than of the other modality, then *changes* in PSS due to the

anisochrony indicate a modification of the time at which the final stimulus is perceived

(Figure 2.2B); we call this effect *Bias by Expected Timing* (BET). In particular, the

comparison of PSS values obtained with audio and visual sequences suggests that if the

last stimulus in the sequence is presented slightly earlier than expected, the BET leads to a

later perception of the stimulus (delay). On the other hand, for the last stimulus presented

at the expected time-point or later than expected, the BET leads to an earlier perception of

the stimulus (acceleration). Physically synchronous audiovisual stimuli are differentially

reported as either "sound first" or "light first" dependent on their anisochrony with the

sequence as shown in Figure 2.3B. The BET is independent of the sequence modality

(Figure 2.3A). We find no difference in the JND (discriminability) across conditions or

between modalities (Figure 2.3C).
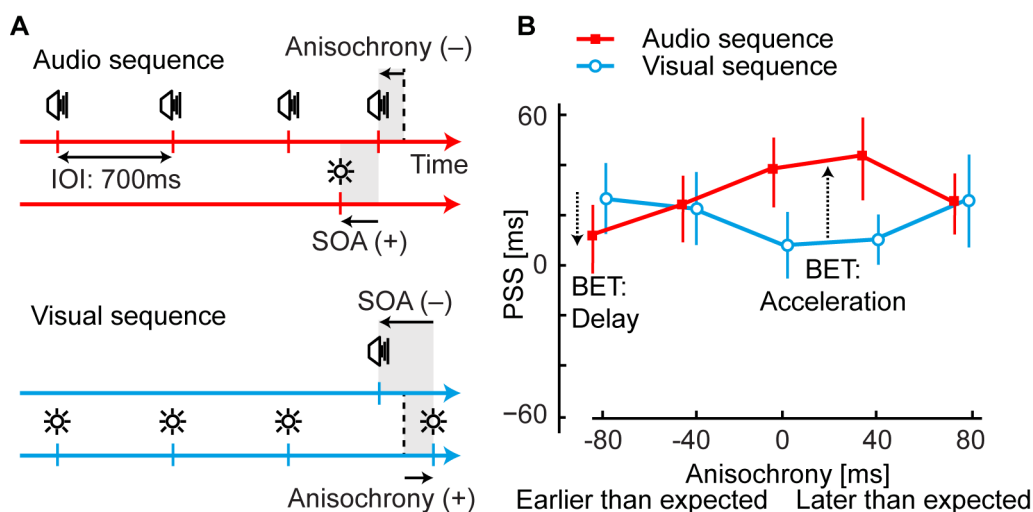
Figure 2.2. Stimuli and results of Experiment 2. (A) Examples of trial sequences where participants judged the temporal order of the audiovisual pair presented at the end of a sequence. Auditory and visual sequences were interleaved. Top: An audio sequence with the final stimulus presented earlier than expected (negative anisochrony) and with a light presented before the final audio stimulus (positive SOA). Bottom: A visual sequence with the final stimulus presented later than expected (positive anisochrony) and with sound presented before the final visual stimulus (negative SOA). (B) Average PSS values corresponding to the SOA at which audio and visual stimuli are perceived as being simultaneous. A positive PSS means that light has to be presented before the sound to be perceived as simultaneous. The difference between PSS values on the two curves indicates the bias by expected timing (BET): in this graph perceptual acceleration happens when the audio PSS is higher than the visual PSS. If there was no change in perceived timing across the presented anisochronies, the pattern of PSS values should be horizontal. The BET, instead, changes as a function of anisochrony (interaction term of a two-way r.m. ANOVA, $F(4,44)=4.8$, $p=.0026$, $\eta p^2=.30$) as stimuli presented at -80ms are perceptually delayed whereas stimuli presented at 0ms and +40ms are perceptually accelerated. The experiment has been replicated using simultaneity judgments (Supplementary Figure 2.2).
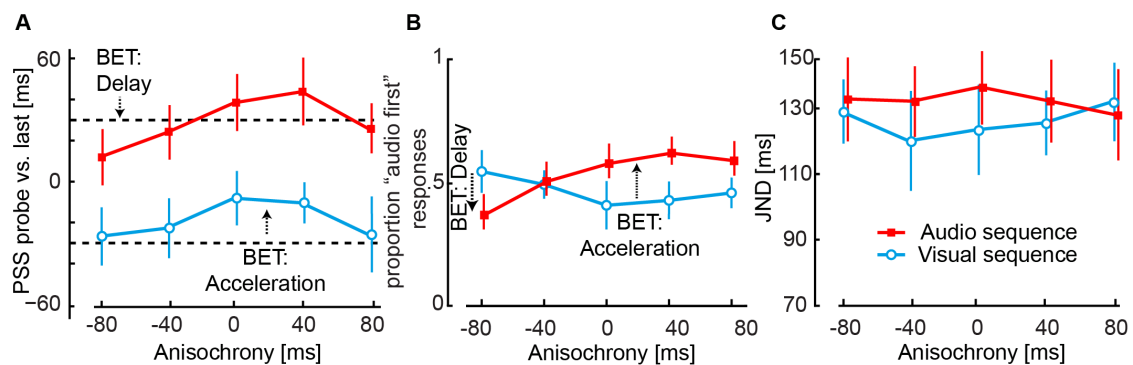
Figure 2.3. Additional analysis of Experiment 2 data. (A) PSS values presented as the asynchrony between probe vs. last repeated stimulus (and not light versus sound as depicted in Figure 2.2A and whose results are shown in Figure 2.2B). The dotted line represents a PSS of 30ms between lights and sounds, which is reversed in sign for the visual sequence condition. The BET is evidenced as an equal distortion of the two curves. For example, in an isochronous auditory sequence with a physically synchronous light stimulus there is an increase in audio first responses, which evidences the auditory stimulus being accelerated. A two-way r.m. ANOVA with factors sequence modality and anisochrony evidences the difference in latency between audio and visual stimuli ($F(1,11)=7.1$, $p=.022$, $\eta p^2=.39$) as well as the BET effect (factor anisochrony,

$F(4,44)=4.8$, $p=.0026$, $\eta p^2=.30$), but no interaction ($F(4,44)=.4$, $p=.80$). (B) Proportion of "audio first" responses for physically synchronous audio and visual stimuli. The BET is evidenced as the difference between the two curves. A two-way r.m. ANOVA with factors sequence modality and anisochrony yielded no main effects but a significant interaction $F(4,44)=4.6$, $p=.0033$, $\eta p^2=.30$. (C) JND values. A two-way r.m. ANOVA with factors sequence modality and anisochrony indicates that no main effects or interaction are significant (all $p >.3$).

2.3.3 Behavioural Results: Longer sequences and Different IOIs

To test whether the BET depends on the number of stimuli in the sequence, in Experiment 3 participants judged the temporal order of an audio and a visual stimulus following the presentation of audio sequences of different lengths (3,4, or 5 repeated stimuli presented in different blocks, Figure 2.4A). Results indicate that the BET increases as a function of sequence length (Figure 2.4B).

Furthermore, to test whether the observed effects are due to the repeated presentation of the same interval across all trials, in Experiment 4 we used trial sequences with four stimuli each, but varying inter-onset intervals (IOIs) interleaved within a block (Figure 2.5A). The BET is still present when stimuli having different periodicities are interleaved in the same experiment (Figure 2.5B).
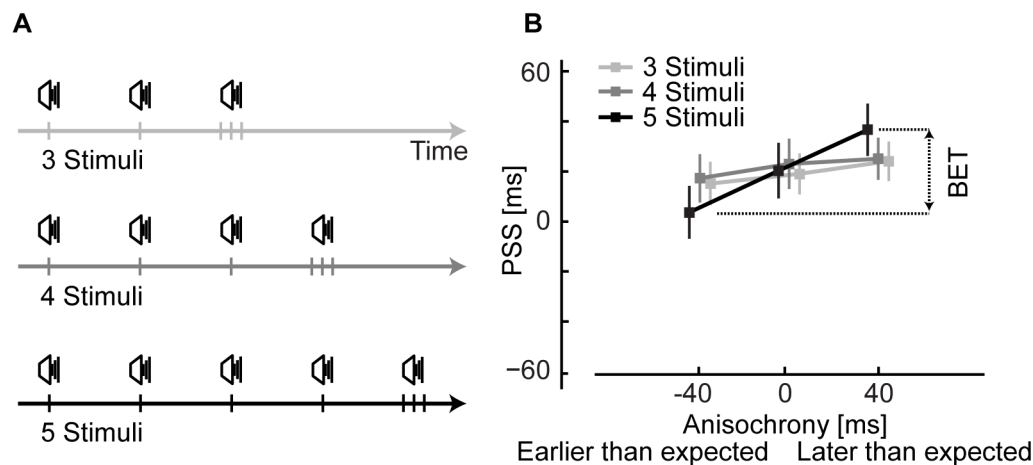
Figure 2.4. Stimuli and results of Experiment 3. (A) Examples of sequences of different lengths, where the last audio stimulus is paired with a visual stimulus (not shown). Differently from the results of Figures 2 and 3, here only auditory sequences have been presented. (B) PSS values for early, on time, and late stimuli differ significantly, confirming the BET found in Experiment 2 (factor anisochrony of a 2-way r.m. ANOVA, $F(2,46)=7.9$, $p=.001$, $\eta p^2=.25$). The magnitude of the BET increases with longer stimulus sequences (interaction of anisochrony and sequence length, $F(4,92)=2.5$, $p=.049$, $\eta p^2=.10$) and the effect is present with 5 stimuli (one-way r.m. ANOVA, $F(2,46)=10.4$, $p<.001$, $\eta p^2=.31$), but not with 4 and 3 stimuli ($F(2,46)=1.6$, $p=.22$; $F(2,46)=.60$, $p=.57$).
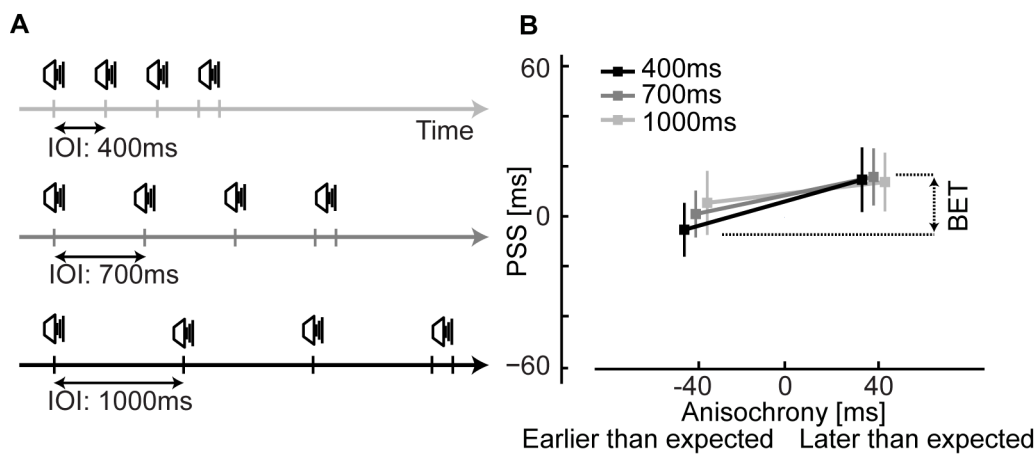
Figure 2.5. Stimuli and results of Experiment 4. (A) Examples of sequences with an IOI of 400ms, 700ms, and 1000ms where the last audio stimulus was anisochronous (±40ms) and paired with a visual stimulus (not shown). Only auditory sequences have been presented. (B) PSS values indicate a BET similar to the other experiments (factor anisochrony of a 2-way r.m. ANOVA, $F(1,23)=15.7$, $p=.0006$, $\eta p^2=.41$), which suggests that testing the same IOI throughout the experiment is not necessary to elicit the BET and that the effect is not limited to one IOI.

2.3.4 A Bayesian Model of Perceived Timing

We model the results collected using *Bayesian Decision Theory (BDT)*. Such a framework has been successfully applied to several perceptual domains (Kersten & Yuille, 2003; Knill & Richards, 1996; Mamassian et al., 2002; Petzschner & Glasauer, 2011; Wolpert & Ghahramani, 2000) including interval estimation (Hartcher-O'Brien et al., 2014; Shi et al.,

2013) and reproduction (Jazayeri & Shadlen, 2010; Miyazaki et al., 2005), but here we apply it for the first time to perceived timing of individual stimuli.

To do this, we hypothesize that the brain represents the probability of experiencing brief events at every time point. To give an example, we can represent the timing of clapping sounds as the probability of *perceiving* a clap at any point in time (past, present, and future). The probability of *sensing* a clap increases at time points immediately following the time at which the clap was produced (likelihood probability, Figure 2.6A). The probability of *encountering* a clap in the future increases at regular intervals for future time points due to the regular timing of applauses (prior probability; Figure 2.6B). If we extend BDT to the time domain, the likelihood and prior probabilities should be combined at each time point (Eq. 2.6) leading to the posterior probability (Figure 2.6C). We will now examine the three components of this process – the likelihood, prior, and posterior.

The *likelihood function* captures the probability of sensing a stimulus after it has occurred. It represents the temporal smearing due to the variability in the delay of sensory processing and, as such, it is akin to the impulse response function (Blommaert & Roufs, 1987). In other applications of BDT to temporal properties, the likelihood has been assumed to have a Gaussian distribution over time (Hartcher-O'Brien et al., 2014; Jazayeri & Shadlen, 2010; Miyazaki et al., 2005; Shi et al., 2013) but here we propose a different shape because of the intrinsic constraints of the representation of individual stimuli over time. First, time flows in one direction and thus the causality of sensory processing needs to be directional. As such, the probability of a stimulus being sensed is more than 0 only after a delay due to neural processing. Thereafter, because sensory processing can only last a finite amount of time, the likelihood probability should return to zero (unless the stimulus could be missed). As in models of reaction time (Ratcliff, 1978), we describe the probability of sensing a stimulus at time t as a bounded accumulation of sensory evidence

(Eq. 2.8). Figure 2.6A shows the shape of the distributions obtained for accumulation of sensory evidence from audio and visual signals that graphically resembles lognormal distributions. If we assume that exactly one perceived onset is associated with a stimulus, then the probability of perceiving the stimulus at any time point sums to 1 and in this way we can deal with two *likelihood probability distributions* (instead of likelihood functions).
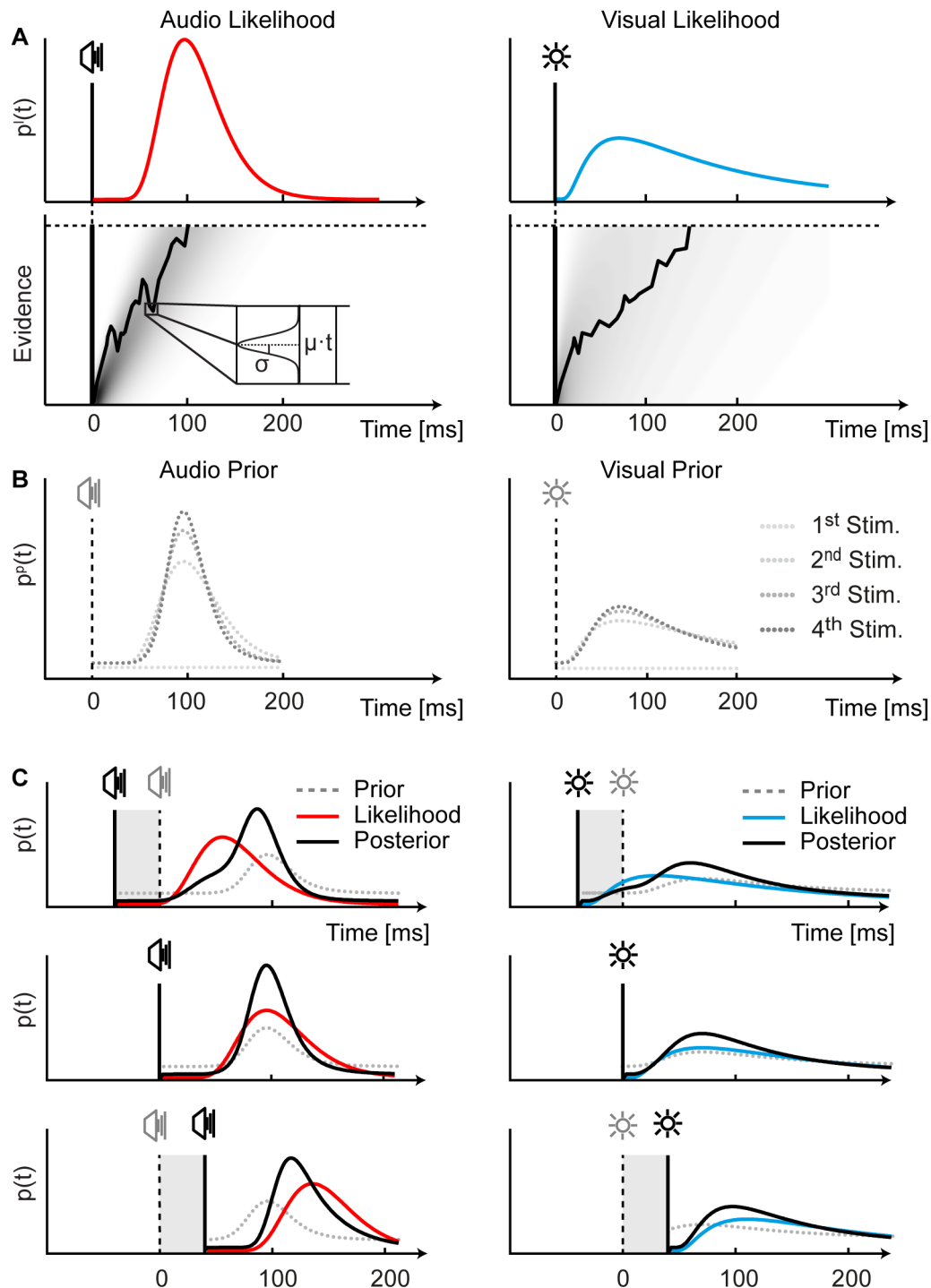
Figure 2.6. Bayesian model of perceived stimulus timing with asymmetric probability distributions (see Methods for details). (A) The top panel shows the likelihood probability distribution – the probability of sensing the stimulus presented at *t=0*. The distribution is obtained by the accumulation of noisy sensory evidence until a threshold has been reached (dotted horizontal line on lower panel) and is described by Equation 7. The parameters of the two distributions are obtained by fitting simple audiovisual temporal order judgments. It should be noted how the onset and the maximum point likelihood are not indicative of the processing latencies of the audio and visual signals as it would do for symmetric distributions, as the mean of the distribution is displaced from the mode. (B) The prior probability distribution for the next stimulus is obtained by combining the prior for the previous stimulus with the current posterior distribution, plus a constant (Methods, Eq. 2.7). (C) Combination of prior and likelihood to obtain the posterior distribution according to Eq. 6 for the fourth stimulus (last stimulus in Experiment 2) appearing -40 ms, 0 ms, or +40 ms with respect to the expected time (separate rows). Perceived timing is obtained retrospectively by considering the overall posterior distribution (i.e., by computing the mean of the distribution).

In bounded accumulation models of reaction time, the parameters for the accumulation of sensory evidence vary due to stimulus characteristics, but also due to context and task conditions (Ratcliff, 1978). Here instead, we constrain the likelihood function not to change with identical stimuli whilst the shape of the *prior probability distribution* changes with more stimuli presented. The *a-priori* probability of a stimulus over time is modelled to be flat when the first stimulus is presented (Figure 2.6B). After the first stimulus occurs, the prior should not be flat any more due to the knowledge of the temporal statistics of the environment. To understand why, consider that if someone claps once, when do you expect the next clap to occur? The probability of the *second* stimulus occurring *before* the first one is necessarily nil: the prior starts at 0 for when the first clap is heard and it increases in the future. The most probable time at which you expect a second clap corresponds to the most frequently experienced interval between claps, roughly ¼ of a second (Repp, 1987). The probability of hearing a second clap then decreases but doesn't quite reach 0, as hearing a clap even tomorrow is always a

possibility. Here we assume that the prior for the second stimulus peaks at the most frequent inter-onset interval used in the experiment (700 ms).

When the hands clap for the second time, the perceptual system has an estimate of the duration of the interval between two successive claps. From previous experience there is the knowledge that subsequent intervals are likely to be similar in duration (clapping variability is typically 2.5% of the intervals (Repp, 1987)). Here we assume that the prior linking two successive interval durations is a delta function because apart from the small anisochrony of the last stimulus most intervals presented in each trial have the same duration. The interaction between the sensed duration and the prior that links two successive intervals is what has been modelled as the likelihood probability distribution in previous models of interval estimation (Jazayeri & Shadlen, 2010; Miyazaki et al., 2005). Here instead, this prior is taking part in the formation of priors about points in time.

Temporal expectations build up as more information is acquired. To model this, we update the prior probability in a way similar to a Kalman filter, by recursively integrating the posterior distribution of the previous stimulus into the prior (Eq. 2.7). The prior distribution becomes more and more similar to the asymmetric likelihood, whilst its maximum value does not deviate from the previously experienced intervals (Figure 2.6B). As the posterior is produced by the asymmetric prior and likelihood, its right side is also longer than the left but this asymmetry decreases at every stimulus. To model how the posterior probability distribution leads to perception, here we consider the whole shape of the probability distribution over time (i.e., as opposed to only using its maximum). The pattern of BETs is due to the combination of the asymmetric likelihood with the asymmetric prior: there is an attraction of the posterior towards the prior, but the larger reduction of the posterior's right tail leads to perceptual acceleration of expected stimuli compared to the likelihood taken alone (Figure 2.6C middle).

Figure 2.7 shows the results of a simulation for the Bayesian model we propose, as well as the interval-based and entrainment models we have discussed in the introduction (see Methods for details about their implementation). The predictions of each model for the conditions of Experiments 1 and 2 should be compared to the experimental data summarized in Figure 2.1 and Figure 2.2B respectively.
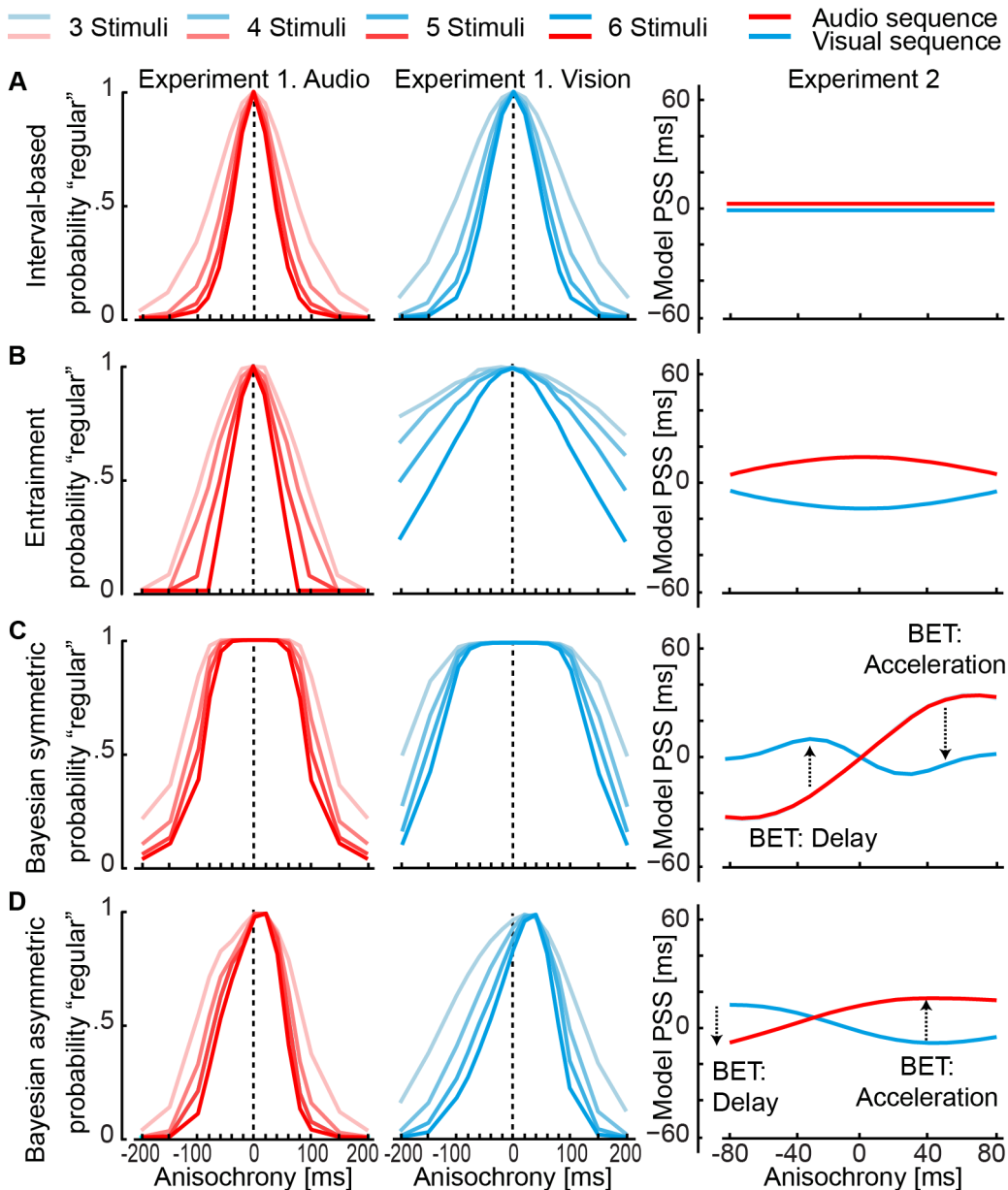
Figure 2.7. Comparison of the model predictions for Experiments 1 and 2 (see Methods for details about their implementations). For Experiment 1 the maximum point of each curve is normalized to 1. (A)

Predictions of an interval-based model obtained by fitting the multiple look model (Drake & Botte, 1993) to the distribution of responses. The model does not predict asymmetries in Experiment 1 or changes in perceived timing in Experiment 2 as it is based on representation of durations. (B) Prediction of an entrainment model (Large & Palmer, 2002). Detection of irregularities has been hypothesized to be symmetrical (Barnes & Jones, 2000; Jazayeri & Shadlen, 2010; McAuley & Jones, 2003; Miyazaki et al., 2005) and the predictions for Experiment 1 suggest only a slight asymmetry (that is more evident in the visual condition). The model does not make explicit predictions about changes in perceived timing of stimuli, but here we show how a 15 ms prior-entry effect (Spence & Parise, 2010) proportional to the attentional pulse would affect the results of Experiment 2. (C) Predictions of a Bayesian model with Gaussian likelihood distributions. The posterior is obtained by combining likelihood and prior while the likelihood of the last stimulus is shifted according to the presentation anisochrony. The predictions for a distortion in perceived timing of stimuli obtained in Experiment 2 are used to modify the conditions of the interval-based model in Experiment 1. It should be noted that there is no BET for isochronous stimuli due to the symmetry of the distributions. (D) Prediction of a Bayesian model with asymmetric distributions (Figure 2.6). PSS for Experiment 2 are obtained from the posterior of the audio and visual stimuli (see Figure 2.9). The pattern indicates that there is no BET for stimuli presented roughly 40 ms earlier than expected. Predictions for Experiment 1 are obtained by modifying the timing of stimuli in the interval-based model.

2.4   Discussion

Our psychophysical experiments show that temporal regularity can change the perceived timing of stimuli – the *bias by expected timing* effect (BET) – without requiring participants to perform speeded responses (that can be affected by motor preparation), nor magnitude estimation (that can be subject to behavioural optimization (Petzschner et al., 2015)). The results of Experiment 2 indicate that the BET leads to an acceleration of stimuli presented at the expected time point or later. In addition, the BET for stimuli presented earlier than expected induces a perceptual delay. It has been reported that longer sequences lead to better discrimination of anisochrony (Drake & Botte, 1993; Miller & McAuley, 2005; Schulze, 1989) and current accounts of temporal sensitivity predict

symmetric performance for early and late stimuli (Barnes & Jones, 2000; Henry &
Herrmann, 2014; Large & Palmer, 2002; McAuley & Jones, 2003). However, there have
been no studies that have tested this prediction, although "slight asymmetries" in the
profile of data have been previously described (Barnes & Jones, 2000). Our data show that
this is not the case and it is more in line with the asymmetric distortion in perceived
duration for early and late stimuli.

The BET counteracts the improved detectability of stimuli presented later than
expected, that is, stimuli following a long sequence that are presented later than expected
are perceptually accelerated (leading to an increase of "regular" responses) against the
detectability of the asynchrony (which should lead to an increase of "irregular" responses).
On the other hand, from the results of Experiment 2 we see that perceptual delay is only
present at large anisochronies for stimuli presented earlier than expected (larger than the
point where the two curves cross in Experiment 2, around 40ms as in Figure 2.2B, which
is captured by the model in Figure 2.7D). Thus the BET for early stimuli is insufficient to
counteract the effect of the improved detectability leading to an asymmetric distribution of
responses as also shown in Figure 2.7D.

Figure 2.7 allows us to qualitatively compare the predictions of extant models of
time perception with the ones of our proposed model. The difference between the
predictions are striking considering the number of fitted parameters:
- JND_2 obtained for three stimuli sequences used to predict the values with more stimuli,
and fitted weight l for the multiple look (interval) model;
- period coupling q, period coupling η, and focusing κ all fitted for the entrainment model;
- standard deviation of accumulated noise σ calculated with pre-test data, and fitted
constant ω to the prior for the asymmetric Bayesian model.

Interval-based models explain perceptual effects related to the presence of rhythmic sequences through a modification of the representation of the interval duration (Church et al., 1994; Drake & Botte, 1993). Whilst the model accounts for an increase in the sensitivity to temporal deviations (Figure 2.7A), such a predicted increase is necessarily symmetrical and thus cannot account for the experimental data. Furthermore, the model does not predict changes in the perceived timing of stimuli at different anisochronies as it is based on the representation of interval durations.

To quantify the predictions of *entrainment models*, we simulated an eminent model tailored to the experimental paradigm employed in Experiment 1 (Large & Palmer, 2002). We find that the detection of irregularity does not follow the asymmetric pattern of Experiment 1 (Figure 2.7B). Entrainment models could be formulated to predict changes in perceived timing of stimuli by appealing to the prior-entry effect (Spence & Parise, 2010) as a function of temporal attention (Barnes & Jones, 2000; Breska & Deouell, 2014; Large & Palmer, 2002). The outcome is a symmetric acceleration that decreases with deviant stimuli (Figure 2.7B). It should be noted that in the original formulation, the detection of irregularity has been thought to be unaffected by this temporal distortion (Large & Palmer, 2002).

The *Bayesian* model with *symmetric* distributions predicts that the perceived timing of irregular stimuli should be biased so to make any deviant intervals to be more similar to previously experienced ones (Shi et al., 2013). The magnitude of the bias decreases with large anisochronies and the effect is identical for stimuli presented too early and too late leading to a symmetric pattern in Experiment 2 (Figure 2.7C). The distortion in perceived timing towards isochrony should make the detection of anisochronies more difficult, leading to a wider (and symmetric) distribution of responses in Experiment 1.

The *Bayesian* model with *asymmetric* distributions is based on the relaxation of the normality assumption often employed in BDT accounts (Jazayeri & Shadlen, 2010; Miyazaki et al., 2005; Sciutti, Burr, Saracco, Sandini, & Gori, 2014; Shi et al., 2013). As for the prior entry phenomena (J. J. J. McDonald et al., 2005), perceptual acceleration for on-time and late stimuli and perceptual delay for early stimuli are explained through a changes in the shape of the posterior – not by a shift of the distribution. In this way, the absence of a BET is predicted for stimuli presented approximately 40 ms earlier than expected, not for isochronous stimuli. The asymmetry in the BET makes the predicted pattern of perceived timing of stimuli shown in Figure 2.7D qualitatively match the pattern of results visible in Figure 2.2B. The model predicts a temporal regularization as in recently-proposed models of interval estimation and reproduction (Jazayeri & Shadlen, 2010; Miyazaki et al., 2005; Sciutti et al., 2014), as interval duration estimates are computationally successive to the estimate of individual stimulus timings (Schwartze & Kotz, 2013). We propose that such regularization could be seen as a modulation of the prior-entry effect as a function of the survival probability (Janssen & Shadlen, 2005; Nobre et al., 2007). The distortion in perceived timing also generates better discrimination of temporal irregularities for early than for late stimuli in long sequences, which resembles the pattern found in Experiment 1 (Figure 2.1B). In sum, the asymmetric Bayesian model provides predictions that qualitatively resemble the data of the two experiments.

The model, perhaps counterintuitively, predicts that the BET should not vary substantially if the sequence is composed of different stimuli (i.e. sounds vs. lights) as we assume common processing mechanisms across modalities (Petzschner et al., 2015). Because the prior resembles the likelihood and the BET is due to the ratio between the width of the prior and the width of likelihood, the ratio between the two widths remains roughly constant. For this reason, BET curves have similar vertical deformations for

different stimulus types (Figure 2.8). The difference between stimuli becomes evident as a modulation in the tuning of the effect (i.e. the spread of the BET across anisochronies). As we find no effect of anisochrony on PSS with 4 stimuli in Experiment 3, but the same condition leads to a BET effect in Experiment 2, we hypothesize that the difference can be accounted for by the added constant ω, whose value is higher because of the use of only audio sequences. Because of the shape of the prior over time, the model also naturally accounts for perceptual phenomena related to the scalar property of interval timing (the estimation error of an interval increasing as the IOI increases (Gibbon et al., 1984)) as with longer intervals the prior becomes flatter leading to a smaller BET (Figure 2.5).
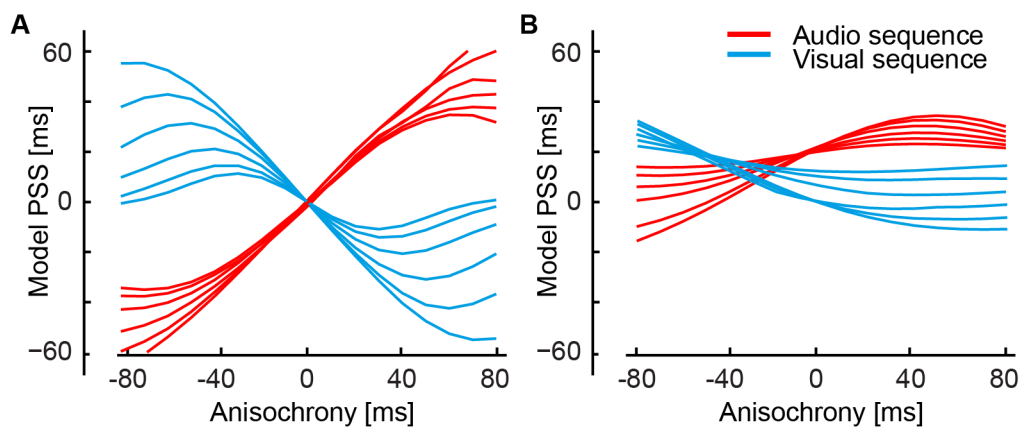
Figure 2.8. Predictions of the Bayesian models for Experiment 2 with different values of the added constant ω: higher values lead to flatter curves as the prior has effectively less and less effect and the BET is smaller. (A) Bayesian model of perceived timing with symmetrical distributions and (B) with asymmetrical distributions.

In our experimental paradigms we exploit a prior based on the natural tendency of intervals to be similar to the ones previously experienced. The effect of such a prior on the posterior probability can be quickly modulated *without* a lifetime experience of the natural environmental statistics (Weiss, Simoncelli, & Adelson, 2002) or without the need to present the same property throughout the experiment (Jazayeri & Shadlen, 2010). Such

rapid modulations are in line with the bottom up influence of regular sequences on perception (Breska & Deouell, 2014) and have also been found in simultaneity judgments after exposure to only one audiovisual stimulus (Van der Burg, Alais, & Cass, 2013). Changes in simultaneity perception have also been related to changes in perceived timing of individual stimuli (Di Luca, Machulla, & Ernst, 2009) that in some accounts have been explained by changes of the likelihood function (Sato & Aihara, 2009), rather than by the presence of an asymmetric prior as proposed here.

The iterative formation of the prior within a trial can be seen as the progressive entrainment of cortical activity leading to tuned attentional deployment (Henry & Herrmann, 2014; Lakatos et al., 2008; Rohenkohl & Nobre, 2011). The phase of delta-theta activity could be a plausible neurophysiological mechanism for representing a-priori probabilities of encountering a stimulus (Arnal & Giraud, 2012), and recent work supports our view that facilitation of sensory processing is shaped by a-priori probability of stimulus timing (Carnevale et al., 2015). The results of their simulations (Carnevale et al., 2015) show a pattern with a very similar shape as the prior shown in Figure 2.6b. Temporal expectations have been shown to lead to a desynchronization of alpha-band activity (Rohenkohl & Nobre, 2011), where the neural response to stimuli is amplified at the expected time point leading to the modulation of perceived timing (J. J. J. McDonald et al., 2005). Stimuli presented too early are not amplified, but the slow offset makes stimuli that are presented too late to be amplified, leading to the asymmetric effect on perceived timing found here.

2.4.1 Conclusion

In this paper we have proposed a quantitative model for the bias by expected timing (BET) effect, where temporal expectations are modelled as Bayesian priors. Such priors come into place after only two identical stimuli and they act as predictors for when a future

stimulus should occur. Our results suggests that the brain processes stimuli based on a probabilistic assessment of whether they will be presented at every point in time and that the resultant temporal expectations can lead to marked distortions in perceived timing.

2.5   Methods

2.5.1 Ethics Statement

The study and all experimental protocols were approved by the STEM Ethics Committee of the University of Birmingham. The methods were carried out in accordance with the approved guidelines.

2.5.2 Participants

Participants were 90 undergraduate students with an average age of 20.83 (SD: 2.20). In total there were 90 undergraduate student participants with an average age of 20.83 (SD: 2.20). In Experiment 1 there were 15 participants (10 females, M_{age} = 21.07, SD_{age} = 1.87) in the auditory experiment, and 15 in the visual experiment (9 females, M_{age} = 20.27, SD_{age} = 1.83); In Experiment 2 there were 12 participants (10 females, M_{age} = 20.67, SD_{age} = 2.50); In Experiment 3, 24 participants (18 females, M_{age} = 21.17, SD_{age} = 2.53); whilst in Experiment 4 there were 24 participants (16 females, M_{age} = 20.67, SD_{age} = 2.16). Informed consent was taken prior to the experiment from all participants, who were either compensated £6 per hour or were given course credits. All participants reported normal or corrected-to-normal hearing and vision, and they were all naïve to the purpose of the experiment.

2.5.3 Experimental Setup

Participants sat in a quiet, well-lit room at a distance of approximately 50 cm from the light- and sound-producing apparatus. Visual stimuli were flashes produced by a red 5 mm LED positioned in front of the participant (20 ms with 5 ms linear ramp, 91 Cd/m^2). Audio stimuli were beeps produced by a speaker 50 cm to the left of the participant (20 ms with 5 ms linear ramp, 1 kHz, 75.1 dBA). Signals were generated using a computer audio card

connected to two identical audio amplifiers. Signals for the whole trial were loaded on the audio card before the trial started to ensure accurate timing.

2.5.4 Psychophysical Procedures

Experiment 1 – Isochrony Judgments. The aim of Experiment 1 was to test whether there is an increase in sensitivity to temporal deviations as a function of how many stimuli there are in a sequence. Fifteen participants took part in the audio experiment and another fifteen in the visual experiment. Sequences of 3, 4, 5 or 6, unimodal stimuli (either audio or visual) were presented with a regular Inter-onset Interval (*IOI*) of 700 ms except the last stimulus, which had a deviation of 0, ±20, ±40, ±60, ±80, ±100, ±150, or ±200 ms. Each trial type was repeated 8 times. The participant's task was to report whether the last stimulus appeared to be regular or not with the rest of the isochronous sequence. Participants responded by pressing one of two keys and the next stimulus would appear 1.5 to 2s after they had been released. For each participant, we computed the proportion of responses for each anisochrony and sequence length. Individual trials for different conditions were randomly interleaved in all experiments. The data of all participants were analysed.

Experiment 2 – Audiovisual Temporal Order Judgments. The goal of Experiment 2 was to understand whether the anisochrony at which a stimulus is presented affected the perceived timing of a stimulus in a sequence. Twelve participants took part in the experiment. Participants completed the experiment in two phases: the *practice phase* and *test phase*. The goal of the practice phase was to familiarize participants with the audiovisual TOJ task, assess performance, and provide baseline data for the creation of the Bayesian models. Participants were presented with a single audiovisual stimulus pair separated by a stimulus onset asynchrony (*SOA*) of 0, ±20, ±90, ±170, ±250, or ±350 ms.

Each *SOA* was repeated 6 times, totalling 66 trials. The participant's task was to report whether the audio or visual stimulus appeared first in time. Participants responded by pressing one of two keys and the next stimulus would appear 1.5 to 2s after they had been released.

During the test phase, participants were presented with a unimodal (either audio or visual) sequence of 4 stimuli having an *IOI* of 700 ms except the last stimulus, which deviated by either 0, ±40, ±80 ms. The last stimulus in the sequence was presented together with a stimulus in the other modality (e.g., a visual stimulus paired with a sequence of sound stimuli) with an *SOA* of 0, ±40, ±80, ±120, or ±200 ms with respect to the anisochrony of the last stimulus presented. Each trial type was repeated 8 times. The participant's task was to report which of the two stimuli presented at the 4th timepoint appeared first in time – "audio first" or "visual first". Participants responded by pressing one of two keys and the next stimulus would appear 1.5 to 2s after they had been released.

For each participant, we computed the proportion of responses for each presented *SOA*. Of particular interest to our hypotheses was the point of subjective simultaneity (*PSS*): the *SOA* at which an individual participant was equally likely to respond that either of the two stimuli was first. Positive *PSS* values mean that the light had to be presented before the sound to be perceived as synchronous and negative values indicate that the sound had to be presented before the light for perceived synchrony. Changes in PSS as a function of anisochrony indicate a modification of the perceived timing of stimuli due to expectation. Also of interest was the just-noticeable difference (*JND*), the asynchrony necessary so that participants report the correct order of the stimuli at a proportion of .84 (which corresponds to 2σ). The *PSS* and *JND* were estimated as the first and second moments of the distribution underlying the psychometric function by using the Spearman-

Kärber method (Miller & Ulrich, 2001). This method provides non-parametric estimates that avoid assumptions about the distributions underlying the psychometric functions. A mathematical derivation of the method follows. First we define SOA_i with $i=\{1, ... 15\}$ as the 15 values of audiovisual SOA used in the experiments and p_i with $i=\{1, ... 15\}$ as the associated proportion of "light first" responses. We further set two SOAs outside of the range tested, $SOA_0=-250$ ms, $SOA_{16}=+250$ ms to be able to compute the intermediate SOA between two successive ones

$$s_i = \frac{SOA_{i+1}+SOA_i}{2}, \text{ with } i=\{0, ... 15\} \tag{Eq. 2.1}$$

We then define two associated proportions to these extreme SOAs $p_0=0$, and $p_{16}=1$, and we calculated the associated values of the difference in proportion, taken at and above 0 to monotonize the proportion of responses

$$dp_i = max\,(0,\ p_{i+1} - p_i)\,, \text{ with } i=\{0, ... 15\} \tag{Eq. 2.2}$$

With these indexes we can express *PSS* and *JND* analytically as such:

$$PSS = \frac{1}{\sum_{i=0}^{15} dp_i} \sum_{i=0}^{15} s_i dp_i \tag{Eq. 2.3}$$

and

$$JND = \sqrt{\sum_{i=0}^{15} dp_i(s_i - PSS)^2} \tag{Eq. 2.4}$$

Values of *PSS* and *JND* in the test phase of the experiment were used to assess participant performance. If *JND* was below 200 ms and if *PSS* did not exceed ±175 ms, participants data was not included in the analyses. The data of 4 participants were not analysed. The test phase data was used to determine the likelihood distribution parameters of both the symmetric and asymmetric Bayesian models (detailed below) as this simple TOJ task was not biased by temporal expectations and thus reflected likelihood probabilities alone.

Experiment 3 – Number of Stimuli in a Sequence. Experiment 3 was aimed at measuring whether the changes in *PSS* found in Experiment 2 increase as a function of the number of stimuli in a sequence. Twenty-four participants took part in the experiment. Only one sequence length was presented in each of four blocks (the order was counterbalanced across participants). Sequences of 3, 4, or 5 audio stimuli were presented with an IOI of 700 ms except the last stimulus, which had a deviation of 0ms or ±40 ms. The last stimulus was presented together with a visual stimulus with an SOA of 0, ±40, ±80, ±120, or ±200 ms. Each trial type was presented 12 times. The data of all participants were analysed.

Experiment 4 – Sequences with Different Periods. The goal of Experiment 4 was to check whether changes in *PSS* still occur if sequences don't have the exact same period. Twenty-four participants took part in this experiment. Four types of audio sequences were presented with an *IOI* of 400, 700, or 1000 ms, except the last stimulus, which had a deviation of ±40 ms. The last stimulus was presented together with a visual stimulus with an SOA of 0, ±40, ±80, ±120, or ±200 ms. Each trial type was presented 12 times. The data of all participants were analysed.

2.5.5 Model Predictions

Interval-Based Model. We adapted the formula of the multiple-look model (Drake & Botte, 1993) to the conditions of Experiment 1 by formulating the participant's task as the comparison between a variable number of intervals with one final interval. In this way we could estimate the JND obtained with intervals of $N=\{3, 4, 5\}$ (JND'_N) from the individual subjects' value of JND with the sequence of 2 intervals (JND_2). JND is calculated from Equation 2.4 by substituting the proportion of "regular" responses to *dp*.

$$JND'_N = \sqrt{JND_2{}^2 \left(\frac{l}{N} + (1 - l)\right)}, \text{ with } N = \{3,4,5\} \qquad \text{(Eq. 2.5)}$$

The weight parameter *l* was tuned by minimizing the sum of the squared differences between the observed data and the model for the audio and visual modalities. As such, the *l* parameter was 0.964 for audio and 0.958 vision. Predicted JND_N' were used as parameters of Gaussian distributions of the responses (the maximum point of the curves was normalized to 1 for better comparison across the models). The mean response distributions across participants for each sequence length are shown in Figure 2.7A. Interval-based models predict no changes in perceived timing of stimuli, leading to constant PSS values as a function of anisochrony.

Entrainment Model. We implemented the entrainment model for perceived temporal regularities (Large & Palmer, 2002) and simulated 1000 stimuli for each of the temporal deviations and sequence lengths used in Experiment 1. For greater detail please refer to Large & Palmer (2002). We tuned individual parameters by minimizing the sum of the squared error between the observed data and model output to audio (period coupling q=.2, oscillation coupling η=3.5, amplitude of the stimulus X=1, focusing parameter κ=1.9) and visual sequences (q=1, η=3, X=1, κ=1.5), both with an oscillator period of p=700 ms. The probability distribution of a regular sequence obtained is shown in Figure 2.7B (maximum point normalized to 1). Entrainment models do not make an explicit prediction about a change in the perceived timing of stimuli, however to show how entrainment models could be modified by temporal attention, we included a 15 ms prior entry effect (Spence & Parise, 2010) that is proportional to an attentional pulse (Barnes & Jones, 2000; Jones & Boltz, 1989; Large & Palmer, 2002).

Bayesian Symmetric Model. Perception is obtained from the posterior distribution, the combination of the on-line sensory evidence (likelihood) with the a-priori knowledge of when a stimulus is expected to be sensed (prior). We propose that expectations are not

static, but they are obtained by iteratively updating the probability of encountering a stimulus at each point in the future.

The likelihood probability distribution $p^l(t)$ is the probability of sensing a stimulus at time t given that the stimulus is produced in the environment. A Gaussian distribution with mean μ and variance σ^2 is used to describe the noise in sensory latency. We constrain $\mu = k\,\sigma$ and determine the value of the parameters k_A, σ_A, k_V and σ_V (subscripts A and V denote audio and vision respectively) that give most similar values of *PSS* and *JND* (via grid search of the minimum sum of squared differences) obtained as described in Figure 2.9 to test phase data of Experiment 2 (*PSS*=25 ms visual first and *JND*=130 ms). This leads to values of k_A=50, σ_A=.038, k_V=30 and σ_V=.071. The posterior probability distribution $p^q(t)$ is obtained by multiplying the probabilities of the likelihood $p^l(t)$ and the prior $p^p(t)$

$$p^q(t) \propto p^l(t) \cdot p^p(t) \qquad \text{(Eq. 2.6)}$$

The prior probability distribution $p^p(t)$ is obtained by using the posterior probability $p^q(t)$ for the previous stimulus (i.e., $p^p(t)$ for the time *t-IOI*). The added constant ω leads to a prior with heavy tails (Roach, Heron, & McGraw, 2006) that allows sudden changes in IOI and thus decreases the tendency of fully incorporating the posterior into a new prior (thus mitigating the increase in False Alarms, Carnevale et al., 2015). This is expressed by:

$$p^p(t + IOI) \propto p^q(t) + \omega \qquad \text{(Eq. 2.7)}$$

The parameter ω changes the predictions of the model as shown in Figure 2.8A.

To obtain the predictions for Experiment 2 we calculated the values of the posterior probability distributions for the last stimulus in the sequence applying Equations 2.6 and 2.7 iteratively. Following previous empirical work (Boenke, Deliano, & Ohl, 2009), we assume that the brain does not only consider the onset of the stimulus to perform a TOJ. Although it is unclear what feature is considered for temporal order judgments (Miller &

Schwarz, 2006; Sternberg & Knoll, 1973), for computational simplicity, we adopt the mean of the distribution. At each trial, the response is determined by the sign of the difference in timing between the means of the distributions to be compared (Sternberg & Knoll, 1973). A similar, but computationally more tractable rule, would be to calculate the difference in timing between which the accumulation of the two distributions reaches 0.5 (i.e. the time of the median probability). To calculate the proportion of responses across trials, we then applied signal detection theory (Figure 2.9) to the audio and visual posterior distributions over time (Green & Swets, 1973) to obtain the predicted probability of response for each asynchrony tested. We then obtained a *PSS* using Equation 2.3. The value of the parameter ω modulates the amount of regularization as shown in Figure 2.8A. Figure 2.7C is obtained with the same ω as the asymmetric model.

To derive the predictions for Experiment 1, we used the *JND*s calculated in the interval-based model (Eq. 2.5) to determine the Gaussian curves for each sequence length. Before calculating the response probability using said Gaussians, we applied a temporal distortion to the sensed stimulus timing equal to the PSS obtained in Experiment 2.
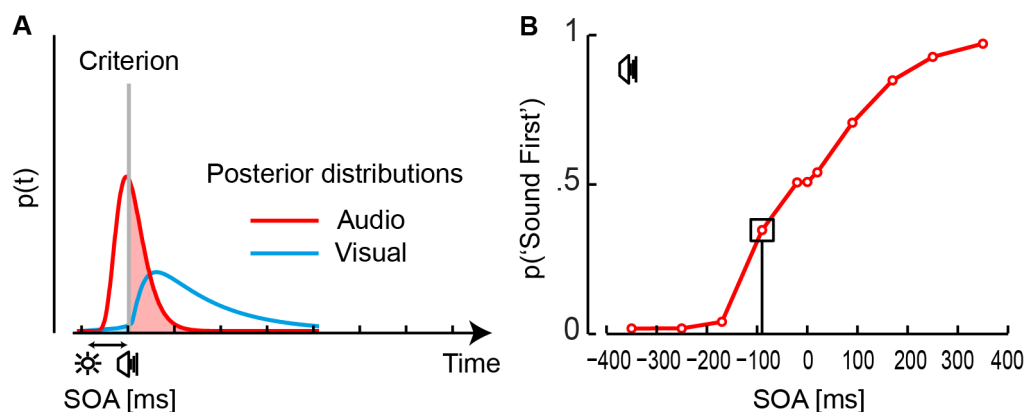
Figure 2.9. Example of how Signal Detection Theory is used to compute model responses. (A) This example represents what happens in TOJ tasks with an audio stimulus presented 90 ms before a visual stimulus. The posterior distributions of the audio and visual stimuli are considered to compute the proportion of sound first responses. The unbiased criterion is established by finding the point that gives the highest d' between the two

curves. The probability of "sound first" responses is calculated as the sum of the two areas below the visual

posterior and on the left of the criterion (Hits) and below the audio posterior on the right of the criterion (CR).

(B) The values of probability of "sound first" responses obtained from the model for different SOAs are

analyzed using the Spearman-Kärber method as for participants' responses (see Methods).

Bayesian Asymmetric Model. The likelihood probability distribution $p^l(t)$ is

modeled as the bounded accumulation of sensory information with rate μ and neural noise

that is Gaussian with variance σ^2 at each time point (Ratcliff, 1978). We fixed the starting

point of the accumulation (value 0) and the threshold level (value 1). As such, the

expression of the likelihood can be rendered analytically as a bounded Wiener process, i.e.

the accumulation of sensory evidence until a threshold is reached with additive Gaussian

noise. The probability of reaching the threshold over time is expressed by:

$$p^l(t) \propto e^{-\frac{dt}{2t\sigma^2}(1-\mu t)^2} \qquad\qquad (Eq.\ 2.8)$$

The proportional sign is due to the normalization across the whole distribution, that makes

the area under the curve equal to 1. We set the accumulation rate to μ=10 and, as for the

symmetric model, we determined the value of the parameter σ_A and σ_V that give most

similar values (through a grid search of the minimum sum of squared differences) of *PSS*

and *JND* to audiovisual data obtained in the test phase, obtaining σ_A=30 ms and σ_V=85 ms

(see Figure 2.6A). The prior probability distribution and posterior probability distribution

are obtained as described for the symmetric model (Eq. 2.6 and 2.7). The predictions for

the asymmetric Bayesian model are presented in Figure 2.7D where the parameter ω

modulates the BET as shown in Figure 2.8B. Parameters were fit for each subject in all

models.

2.6   Supplementary Information
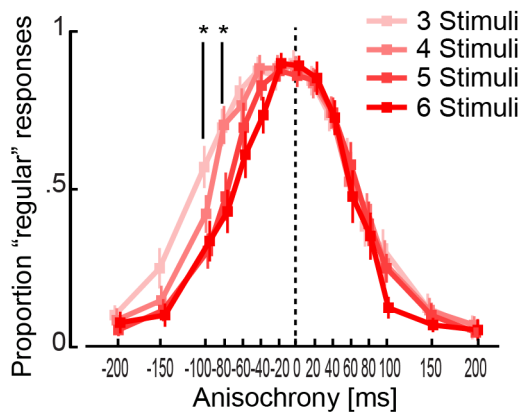
Supplementary Figure 2.1. Proportion of "regular" responses in Experiment 1S plotted as a function of the anisochrony of the last audio stimulus. Sequence lengths were presented in different blocks of trials (rather than interleaved as in Experiment 1). Fifteen new participants took part in this experiment. Asterisks denote asynchronies at which the proportion of responses differs across sequence lengths showing a change in responses for stimuli presented earlier than expected.
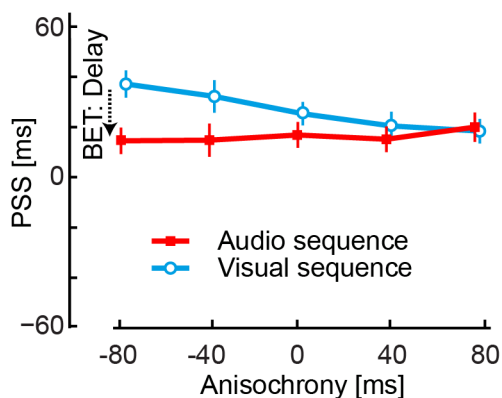
Supplementary Figure 2.2. PSS values obtained in Experiment 2S with simultaneity judgments performed by 12 new participants. PSS was obtained by substituting *dp* with *p* in SI Equations 3 and 4 in SI Materials and Methods. All other experimental conditions are identical to Experiment 2 whose data is shown in Figure 2.2 and 2.3. PSS values qualitatively replicate the change in BET as a function of anisochrony of the last stimulus (interaction term of a two-way r.m. ANOVA, $F(4,44)=5.2$, $p=.0016$, $\eta_p^2=.32$). The prominence of a delay in the BET could be due to a different response criterion with the two sequence types or to a decrease in perceptual acceleration. I.e., it has been previously established that prior-entry biases leading to perceptual acceleration have larger magnitude in TOJ than SJ data (Spence & Parise, 2010). It should be noted that this

pattern is inconsistent with the predictions of interval-based models (there should be no change) and of entrainment models (they cannot account for delay).

Chapter 3 Temporal Regularity of the Environment Drives Time Perception

In this paper, we tested if the temporal context biases the perception of regularity – as well as the perceived timing of stimuli. We present a parsimonious account of the phenomena we observed based on the Bayesian framework.

3.1   Abstract

What causes a sequence of isochronous stimuli be perceived as temporally regular? It's instinctive to assume that a regularly timed sequence should always be perceived as regular, but here we show that such a judgment depends on what environment the sequence is presented in. A regularly timed environment promotes perceptual regularisation of small irregularities by changing the perceived timing of stimuli. In contrast, an environment composed of irregularly timed sequences can make even perfectly regular sequences be perceived as irregular due to the untamed variability in neural processing. These results suggest that the brain learns the temporal statistics of the environment and uses them to generate expectations about future stimuli. The combination of temporal expectations with sensory information has the effect of modulating the time at which stimuli are perceived and of improving the reliability of temporal estimates.

3.2   Introduction

Time is a physical dimension that pervades numerous aspects of human perception, yet humans do not perceive time veridically. For example, the subjective experience of duration can be modulated by non-temporal characteristics such as stimulus properties (Horr & Di Luca, 2015a; Thomas & Brown, 1974; Wearden et al., 2007), complexity (Schiffman & Bobko, 1977), stimulus regularity [3], sensory modality (Goldstone & Lhamon, 1974; Wearden et al., 1998; 2006), and context (Dyjas & Ulrich, 2013). As subjective duration can be biased by the characteristics of the immediate environment, here we ask whether the perception of temporal regularity is subject to similar influences. In fact, we observe that context plays a role in the perception of rhythmic stimuli, as humans effortlessly learn the temporal structure of events (Grahn & Rowe, 2013; Kösem & van Wassenhove, 2012; Large & Palmer, 2002; Povel, 1981; 1984; Su & Pöppel, 2012), an ability that is present even in new-borns and infants (Hannon & Trehub, 2005; Winkler, Haden, Ladinig, Sziller, & Honing, 2009). It has been shown that humans, among other animals (Bateson & Kacelnik, 1997; Buhusi et al., 2002; Gallistel et al., 2004; Gribova et al., 2002; Henderson et al., 2006; Janssen & Shadlen, 2005), can improve perceptual judgments (Brochard et al., 2013; Correa et al., 2005; Cravo et al., 2013; Doherty, Rao, Mesulam, & Nobre, 2005; Escoffier et al., 2010; Grill-Spector, Henson, & Martin, 2006; Rohenkohl & Nobre, 2011), reduce neural metabolic consumption (VanRullen & Dubois, 2011) and automatize behaviour (McNeill, 1995; Repp, 2005) by entraining to regular rhythms. Whilst there are several accounts of how the brain deals with rhythmic stimuli (Buhusi & Meck, 2005; Buonomano & Merzenich, 1995; Creelman, 1962; Gibbon, 1977; Gibbon et al., 1984; Karmarkar & Buonomano, 2007; Meck, 2005; Shi et al., 2013; Teki, Grube, & Griffiths, 2011; Treisman, 1963; van Rijn, Gu, & Meck, 2014; Wackermann & Ehm, 2006), all of them assume that a stimulus repeated after the same interval is actually

perceived to be a regular sequence. But as noted previously, time is subject to perceptual distortions so why should a temporal property like regularity be immune to these too?

In order to test whether perceived regularity of a sequence can change due to the influence of the environment, we present sequences of stimuli that are mostly regularly timed to one group of participants, whereas to a second group we present mostly irregularly timed sequences. For both groups, 25% of trials contained a perfectly regular sequence i.e. with the same inter-onset interval (isochrony). We find that an environment of mostly irregularly-timed stimuli makes regular sequences appear irregular. We rule out explanations based on response biases or changes in the criterion for judging regularity and we propose an explanation of these effects that is based on the influence of expectations on the perceived timing of individual stimuli, i.e. the brain makes predictions about the timing of future events that modify perception. An ideal candidate to conceptualise this phenomenon is *Bayesian Decision Theory* (BDT), which has been successfully applied to various perceptual domains (Kersten & Yuille, 2003; Knill & Richards, 1996; Mamassian et al., 2002; Petzschner & Glasauer, 2011; Wolpert & Ghahramani, 2000). Bayesian accounts of time perception have been also formulated (Jazayeri & Shadlen, 2010; Petzschner et al., 2015; Shi et al., 2013), but they do not account for changes in the perceived timing of individual stimuli as they are contingent on the representation of the interval durations that stimuli delimit. In contrast, we suggest that the brain represents the timing of events separately from duration for the purpose of estimating temporal properties of the stimulus sequence, i.e. regularity. Such a scheme relies on previous knowledge about the timing of stimuli. In a regular environment, perceived timing of stimuli is obtained by the combination of incoming sensory information with expectations of when such stimuli should appear. Expected timing should thus bias perception in a way that regularises slightly anisochronous stimuli. To test this,

we measured when stimuli at the end of a sequence are perceived and we find a temporal regularisation effect – stimuli are delayed or accelerated shifting them closer to isochrony points. Our results give more credence to the contemporary idea (Acerbi et al., 2012; Jazayeri & Shadlen, 2010) that the brain keeps track of the temporal statistics of the environment and uses them for perception in a way consistent with Bayesian inference.

3.3   Results

3.3.1 The Temporal Environment Modifies Perceived Regularity

To test whether the regularity of the environment changes how temporal regularity of stimulus sequences is actually perceived, we exposed participants either to an environment where the majority of sequences were perceived to be temporally regular (regular environment, Figure 3.1a upper panel) or in an environment where most of the sequences were perceived to be irregular (irregular environment, Figure 3.1a lower panel). To assess the sensitivity of participants to discriminate whether a sequence is regular (ten Hoopen et al., 2011; 1995; Schulze, 1989), for both groups we presented 25% of sequences composed of 5 beeps where only the timing of the last stimulus deviated from isochrony. The participants' task was to report whether the sequence as a whole was regular – or not. By parametrically varying the timing of the last stimulus we could determine the deviations at which sequences started to appear irregular.

Participants did not change their responses in the two contexts, as the overall proportion of 'regular' responses did not differ between groups (t_{18}=1.3, p=0.22). On the other hand, the distribution of responses for sequences composed of 4 initial regularly-timed stimuli as a function of the deviation of the last stimulus follows a different pattern in the two groups (Figure 3.1b, see Supplementary Figure 3.1 and Supplementary Table 3.1 for responses with jittered sequences). For participants exposed to an irregular

environment, the number of 'regular' responses for slight deviations in the timing of the

last stimulus appears to be much lower than the same sequence embedded in the regular

environment (significant between -40 and 20ms, see Supplementary Table 3.2 for t-tests).

Crucially, participants are also more likely to report perfectly regular sequences (0ms

anisochrony) as being irregular when embedded in an irregular environment. The

difference between the two groups disappears for larger deviations of the last stimulus in

the sequence. Further, when the last stimulus had 0ms anisochrony, participants' perceive

regularity the most when the previous stimuli were isochronous – but perceived regularity

declines almost linearly as a function of the temporal jitter (Supplementary Figure 3.2). To

rule out a change in the criterion used to determine regularity, we applied the two-noisy

criteria simultaneity judgment model (Yarrow, Jahn, Durant, & Arnold, 2011) to the data

shown in Figure 3.1. The estimates of the early (negative anisochronies) and late (positive

anisochronies) criteria for judging regularity and the slopes at the threshold point do not

differ between regular and irregular environments (paired sample t-test on threshold for

early t_{18}=0.2, p=0.82; late t_{18}=-0.7, p=0.51; on slope for early t_{18}=-1.7, p=0.20; late t_{18}=-

2.3, p=0.060). By ruling out the presence of response biases we can increase our

confidence in concluding that the environment had a genuine influence on the perception

of sequence regularity.
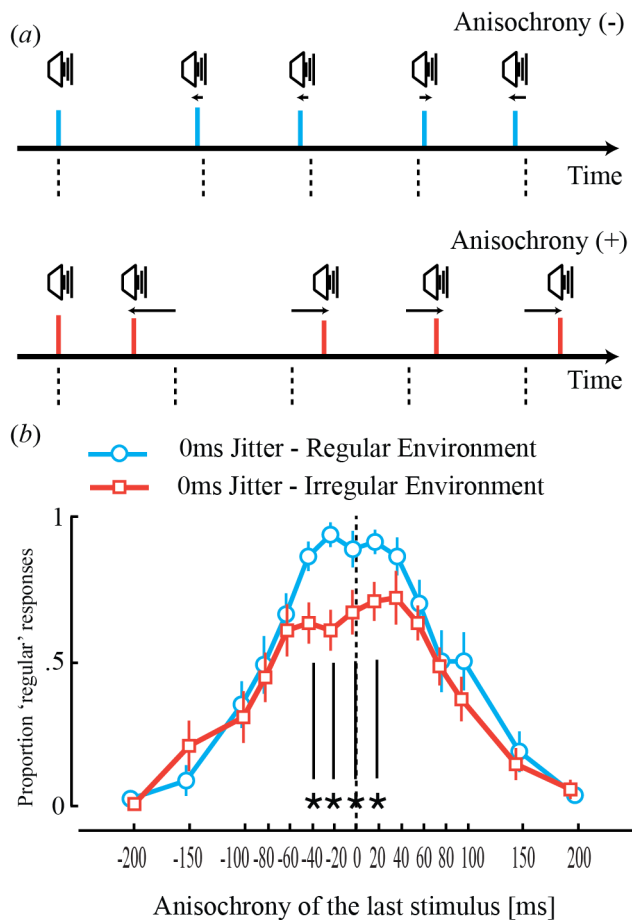
Figure 3.1. Experimental stimuli, design and results for Experiment 1. (a) Top: Example of a sequence with small jitter with the final stimulus earlier than expected. Bottom: Example of a sequence with large jitter with the final stimulus later than expected. (b) Proportion of 'regular' responses as a function of the timing of the final auditory stimulus in the sequence. Each line represents data obtained with 0 ms temporal jitter for the small and large jitter groups. The distribution of responses becomes steeper with sequences with small jitter compared to large jitter (interaction term of two-way r.m. ANOVA on the inverse-normal proportion of 'regular' responses bounded between .01-.99; for the regular context group: $F_{42,378}=3.6$, $p<.0001$, $\eta_p^2=.28$; and irregular context group $F_{42,378}=5.0$, $p<.0001$, $\eta_p^2=.36$) . When comparing 0ms jitter from the group exposed to small and large amounts of jitter, the distribution of responses is significantly lower for the large jitter group at the points denoted by an asterisk (Supplementary Table 3.2). In all graphs, the error bars represent the standard error of the mean.

3.3.2 The Temporal Environment Changes Perceived Timing

Here we posit that since perceived regularity of isochronous sequences could be decreased if sequences are embedded in a temporally irregular environment, the opposite could be

also true – the brain could have a mechanism that makes slightly anisochronous sequences appear more regular if embedded in a temporally-regular environment. This could be achieved if the perceived timing of individual stimuli is modified so that they appear closer in time to the regular time point. To test whether the perceived timing of a temporally deviant stimulus is modified, we asked participants to report whether the final stimulus in an auditory sequence was presented before or after a temporal probe in the visual modality (Figure 3.2a). To measure the perceived timing of the last stimulus, we compared the physical asynchrony at which audio and visual stimuli are perceived to be simultaneous (PSS, Point of Subjective Simultaneity) for stimuli presented slightly earlier and later than isochrony. If the brain regularises these stimuli, we should find that the perceived timing is delayed if the stimulus is presented too early and accelerated if the stimulus is presented too late. For sequences with no temporal jitter (the 4 initial stimuli were isochronous) we find a change in PSS between early and late stimuli only for the group exposed to a regular environment (t_{11}=2.5, p=.030; all others p >.14; see Supplementary Figure 3.3 for a representation of this effect on PSS decomposed into early and late presentation). Stimuli presented early appear perceptually delayed, whereas stimuli presented late are perceptually accelerated. The magnitude of such an effect is highest for stimuli at the end of an isochronous sequence of four stimuli (40% of the anisochrony +/-15.3% SEM), but the effect is in the same direction for irregular sequences (10, 20 and 30ms temporal jitter of the first four stimuli) in the regular environment. No difference in PSS is registered for the group exposed to an *irregular* environment (Figure 3.2c). In other words, participants exposed to a regular environment change the perceived timing of slightly deviant stimuli at the end of the sequence so that the sequence appears more regular than it actually is – a temporal regularisation. Notably, sensitivity in judging the timing of the audio and visual stimuli does not differ in the two temporal environments (Supplementary Figure 3.4).
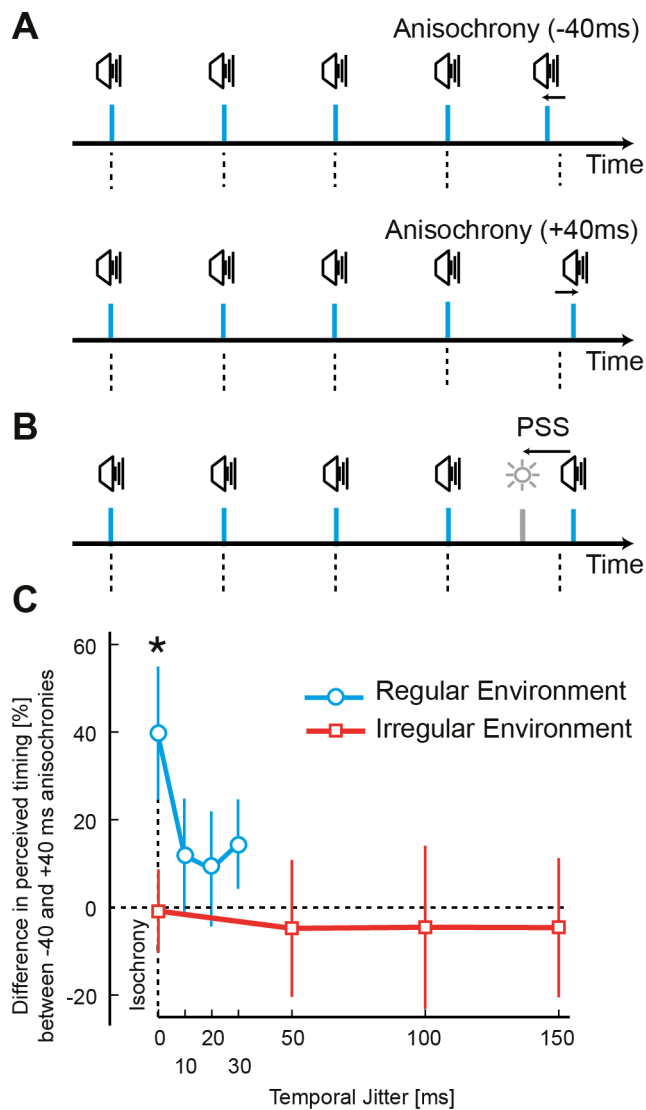
Figure 3.2. Experimental stimuli, design and results for Experiment 2. (a) Examples of a trial sequences in regular environments when the final stimulus was earlier (top panel) or later (bottom panel) than expected. (b) Participants judged the temporal order of the final auditory stimulus and a visual probe. (c) Difference in perceived timing (as a percentage) between stimuli presented -40 ms earlier than expected and +40 ms later than expected as a function of the temporal jitter. Each line represents the group with small jitter, and the group with large jitter. Only in the instance where the temporal jitter is 0ms in the small jitter group is the difference in perceived timing effect significantly different from 0, $t_{11}=2.6$, $p=.02$. Error bars represent the standard error of the mean.

3.4   Discussion

This study demonstrates that the degree of temporal regularity experienced in the environment underpins our perception of the timing of stimuli. A temporally regular environment causes sequences to be reported as being regular more often than when the same sequences are presented in an irregular environment. Such a difference is present also for perfectly regular sequences. In principle, this result could be explained in three different ways: (a) A response bias that leads participants to respond 'irregular' more often in the irregular environment; (b) A more lenient perceptual criterion for perceiving sequences to be regular in the irregular environment which tolerates perceptual anisochronies caused by neural noise, or (c) There could be an influence of the regular environment on the perception of the timing of individual stimuli, so that they are perceptually moved towards the expected time points: earlier than regular stimuli are perceptually delayed whilst later than regular stimuli are perceptually accelerated.

       Our data are not consistent with a response bias effect (a) as here one would expect to see a general increase in 'irregular' responses over all levels of anisochrony. However, as Figure 3.1b illustrates, there is only an increase in 'irregular' responses for slightly irregular sequences suggesting that participants did not change their responses to suit the environment they were exposed to. Secondly, we can dismiss explanations pertaining to changes in the criterion at which regularity is judged (b), as we found no difference in sensitivity to discriminate regularity between groups after implementing the two-noisy criteria simultaneity judgment model (Yarrow et al., 2011). This model, which was developed in the context of temporal order judgments (Ulrich, 1987), captures the probabilistic placement of low and high criteria for perceptual judgments. Values that fall within the two criteria are judged differently than either side. In Experiment 1, we apply this model to regularity judgments as a function of anisochrony. We find no difference

between the high (positive anisochronies) and low (negative anisochronies) criteria for judging regularity between regularly- and irregularly-timed environments. The lack of a difference between groups suggests no differential sensory noise between regular and irregular environments nor differences in the decision uncertainty for judging regularity. Our results point to the third explanation (c): the perceived timing of stimuli in a sequence is affected by the presence of a regular environment – irregular stimuli in a regular environment are perceived closer in time to the expected time point (Figure 3.1, 3.2). This effect can be understood as the brain attempting to estimate time from imprecise sensory information as detailed below.

3.4.1 Bayesian Time Perception

We interpret our results using a Bayesian framework, considering that the brain is trying to obtain an estimate of the timing for each stimulus from noisy sensory information. To improve the reliability of the estimates, the brain incorporates *a-priori* knowledge (prior probability) of when a future stimulus *may* occur with incoming sensory evidence (likelihood function) of when a stimulus *actually* occurred in the world (Figure 3.3). We consider that the brain dynamically updates the prior probability of experiencing a stimulus after each presentation. Temporal expectations rapidly increase and bias the timing of stimuli in a regular environment so to make them appear more regular (Figure 3.3a). In a temporally irregular environment, however, temporal statistics are less precise and thus expected timing is more uncertain (Figure 3.3b). An estimate of perceived timing is based on the posterior distribution, which is the combination of the prior and likelihood. In an irregular environment, the posterior distribution coincides with the likelihood distribution (as the there is no regularity in the sequences, the prior does not contribute to improve precision). Lower precision means a wider distribution, and thus at each encounter of a stimulus, even if physically regular, there is a possibility that the stimulus is

actually anisochronous. For stimuli embedded in a regular environment, instead, expectations improve the precision of timing estimates making stimuli appear closer to the presented time point but they also influence the timing of slightly deviant stimuli. Such an effect is evidenced by the results of Experiment 2 – where participants were asked to judge whether the final stimulus in a sequence was before or after the visual probe. In a regularly-timed environment, slightly anisochronous stimuli are biased to appear closer in time to the expected time point (Figure 3.2c) due to the effect of the prior.

(*a*) *Regular Sequence in a Regular Environment*
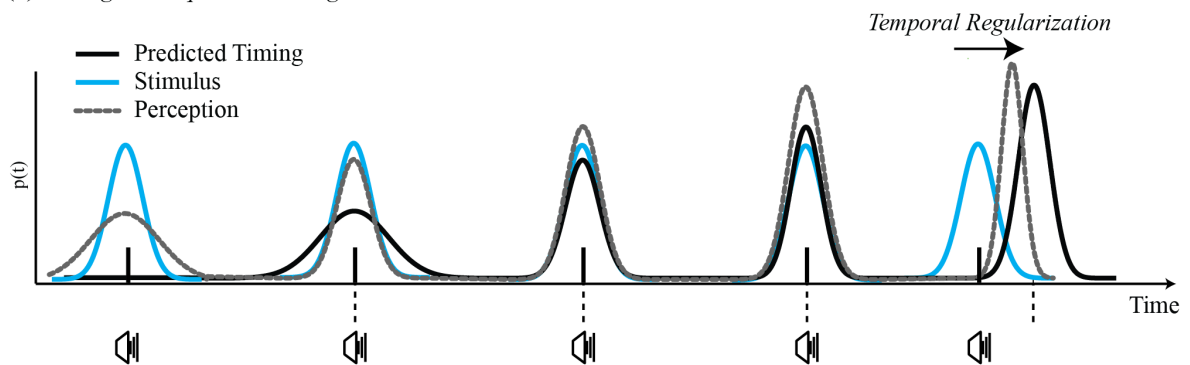
(*b*) *Regular Sequence in an Irregular Environment*
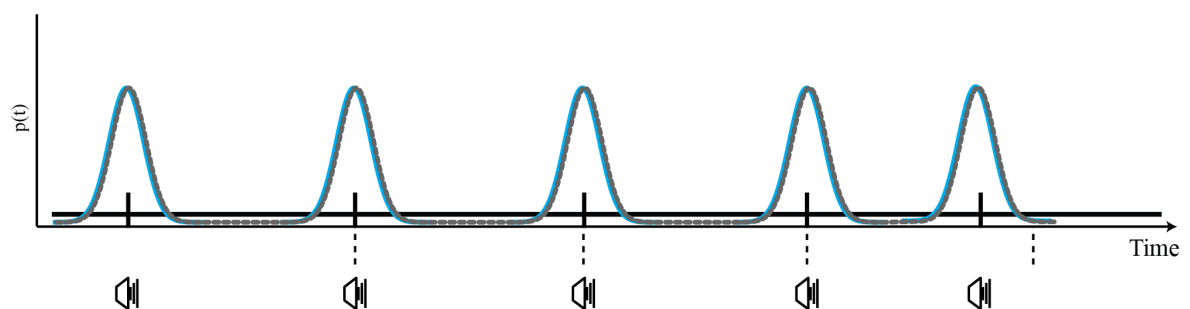
Figure 3.3. A Bayesian Model of Perceived Timing. (a) A regular sequence in a regular environment builds up temporal expectations (prior predicted timing; black line) after each presentation of a stimulus (likelihood; blue line); in order to form an estimate of perceived timing from the posterior (perception; grey dotted line). In this way, the posterior becomes the predicted timing for the next stimulus in the sequence. If a final stimulus is presented slightly earlier than expected – then it is perceptually delayed towards the predicted timing – as well as having greater temporal precision. (b) A regular sequence in an irregular environment on the other hand, the posterior distribution coincides with the likelihood, as a flat prior does not contribute to improve precision or generate greater temporal expectations. Lower precision results in a

wider posterior distribution and as such, after each stimulus presentation, a physically regular stimulus could on chance be reported as being irregular.

We reason that temporal regularisation reflects the tendency of the brain to obtain a compact representation of the environment that can be generally applied for perception (Baddeley, 1992). In a regular environment, small temporal irregularities are likely to be due to sensory noise, and thus representing each individual irregularity would constitute a disadvantageous memory burden. Modifying the perceived timing of stimuli, so to appear more regular, reduces neural metabolic consumption (VanRullen & Dubois, 2011) and improves the perceptual processing of regularly timed stimuli (Brochard et al., 2013; Correa et al., 2005; Cravo et al., 2013; Doherty et al., 2005; Escoffier et al., 2010; Grill-Spector et al., 2006; Jazayeri & Shadlen, 2010; Rohenkohl & Nobre, 2011).

Current models of temporal perception (Gibbon et al., 1984; Karmarkar & Buonomano, 2007; Large & Palmer, 2002; C. Miall, 1989; Schulze, 1989; Treisman, 1963) are mainly concerned with a representation of interval duration and do not make explicit predictions about whether the temporal environment should influence the perceived timing of temporally deviant stimuli. In fact, in these models, the perceived timing of a stimulus is mostly ignored, because they assume that stimuli are simply perceived after a constant perceptual delay. More importantly, no model relates the change in the detection of temporal deviations to changes in the perceived timing of stimuli. Conversely, we find that in a regularly-timed environment stimuli presented earlier than expected are perceptually accelerated whilst late stimuli are perceptually delayed.

3.4.2 *Monitoring Temporal Statistics*

Temporal regularity has important ecological functions, i.e. in sensorimotor synchronization (Elliott, Wing, & Welchman, 2014; Manning & Schutz, 2013; Merker,

Madison, & Eckerdal, 2009; Repp, 2005; Su & Pöppel, 2012), musical performance (Honing, 2013; Large & Palmer, 2002; Merchant, Grahn, Trainor, Rohrmeier, & Fitch, 2015), discrimination (Kusnir, Chica, Mitsumasu, & Bartolomeo, 2011; Rohenkohl, Gould, Pessoa, & Nobre, 2014) and causal learning (Greville & Buehner, 2010). Thus it is not surprising that the brain monitors regularity across stimuli and uses the regularity of the context to modify perception as we present in this study. Exposure to regularities results in the expansion of visual working memory (Brady, Konkle, & Alvarez, 2009; Umemoto, Scolari, Vogel, & Awh, 2010), improved object labelling (Estes, Evans, Alibali, & Saffran, 2007) and object classification (Turk-Browne, Scholl, Johnson, & Chun, 2010). The phenomenon we report in this paper is similar to long-term adaption effects where the context of the experiment influences our perception – and as such, perception is not absolute but relative to some internal reference (Dyjas & Ulrich, 2014). Such an effect has been shown in other temporal contexts, i.e. perceived duration (Jazayeri & Shadlen, 2010) and simultaneity (Miyazaki, Yamamoto, Uchida, & Kitazawa, 2006). Human observers have been shown to use the statistics of the environment to adjust their perception – similar to the phenomena we describe in this paper. Adaptation level theory (Helson, 1947; 1964) has been proposed to account for such instances as the computational problem can be phrased in terms of a null point of a phenomenological dimension. When a physical stimulus corresponds to the null point, a 'neutral' perception is evoked. The null point, however, is subject to adaptation. Applied to our experiments, the null point should move toward being more irregular when in an irregular environment, which should have resulted in a bias to report more regular responses. This is in contradiction to the results of Experiment 1, as we find the reverse – participants reported more irregular responses in an irregular environment.

The temporal regularisation effect that we report here is consistent with a similar tendency found in the literature about perceived duration (Hartcher-O'Brien et al., 2014; Jazayeri & Shadlen, 2010; Miyazaki et al., 2005; Shi et al., 2013). When asked to reproduce intervals, participants do not only employ the information provided by the current interval, but they consider the distribution of intervals they have previously been exposed to making their response similar to the mean duration (Jazayeri & Shadlen, 2010). A similar phenomenon is also found in perceived simultaneity, where participants exposed to an environment comprising a biased distribution of tactile asynchronies, respond about temporal order in the direction of the most frequently presented asynchrony (Miyazaki et al., 2006). Such tendencies to assimilate the context in the judgments is in competition with contrast effects occurring at a 'local' level, where just the previous trial can influence the temporal perception of the next (Van der Burg et al., 2013). It has been argued that the long-term assimilation and short term contrast effects interact with or precede one another (Chopin & Mamassian, 2012; Yamamoto, Miyazaki, Iwano, & Kitazawa, 2012). According to such accounts (Chopin & Mamassian, 2012), the perceptual system encodes both the prior probabilities of the stimulus distribution over time whilst also representing the recent history and redistributing resources in order to efficiently code incoming sensory signals (Barlow, 1961; Wainwright, 1999). Such expectation biases perception in a way that resolves any perceptual conflict with current sensory information.

3.4.3 Re-focusing Time Perception

The Bayesian model of perceived timing we present in this paper advances a re-focusing of investigation towards event-based perception in the temporal domain. Historical and seminal accounts of time perception are focused on modelling time from the perspective of duration (Creelman, 1962; Gibbon et al., 1984; Shi et al., 2013; Treisman, 1963). The model we present here is not in competition with such accounts of duration perception

(Jazayeri & Shadlen, 2010; Petzschner et al., 2015), but it considers a dimension that has been overlooked: the perceived timing of stimuli. We reason that expectations about the absolute timing of stimuli need to be continuously updated during a sequence and thus they must have a fast time-course. On the other hand, long-term priors that enable expectations are interval-based and are updated after each stimulus presentation. For this reason they have a slow time-course that can represent the statistics of the environment. The combination of the two types of information (fast and slow time course) with sensory information improves the brain's efficiency in producing sensory estimates about temporal properties of the environment (Friston, 2005; VanRullen & Dubois, 2011; Wei & Stocker, 2012; 2015).

3.4.4 Conclusion

In this study, we have established that the temporal regularity of the environment drives time perception. Sequences embedded in regular and irregular environments perceptually change, that is: (1) perceived regularity decreases in an irregular environment and, (2) the perceived timing of slightly irregular stimuli is modified to appear more regular if embedded in a regular environment. Using a Bayesian framework, we propose that temporal expectations about timing of stimuli are generated based on the statistics of the environment. Such expectations bias the perceived timing of stimuli in such a way that results in temporal regularisation and thus a more convincing percept of having experienced a regular sequence.

3.5   Methods

3.5.1 Experiment 1

Twenty Undergraduate students from the University of Birmingham took part in this experiment. Participants were recruited via the online participant recruitment system and received course credits as reimbursement. Participants were naïve to the purpose of the study and all had normal or corrected-to-normal hearing or vision.

The auditory stimuli were identical tones produced by a speaker positioned 50 cm to the left of the participant (20 ms with 5 ms linear ramp, 1 kHz, 75.1 dBA). To ensure reliable timing, the signals of the whole trial were sent to the audio card before presentation.

Participants sat in a quiet, well-lit room.. After being given instructions and informed consent, participants were presented with sequences of five auditory beeps with an inter-onset interval (IOI) of 700 ms that could be highly irregular (high jitter group: 0, 50, 100 or 150 ms jitter) or perceptually regular (small jitter group: 0, 10, 20, 30 ms jitter). Sequences were not blocked but were randomly interleaved such that the participant could not know the jitter of the next stimulus sequence. The participants' two-alternative forced choice (2AFC) task was to simply report whether the stimulus sequence was regular or not by pressing one of two keys. The final stimulus was randomly temporally deviant by 0 ±20 ±40 ±60 ±80 ±100 ±150 or ±200 ms. Participants were asked to respond as accurately as possible with no time limit on the response window. After pressing the regular or irregular key, the next trial would start after 1.5 - 2 seconds. Each condition was presented 8 times. The experiment lasted approximately 30 minutes and participants were debriefed upon completion of the task. The data of all participants were analysed.

3.5.2 Experiment 2

Twenty-four undergraduate students were recruited via the University of Birmingham research participation system and received course credits for their participation. All participants reported normal or corrected-to-normal hearing and vision and were all naïve to the purpose of the study.

Participants sat in a well-lit and quiet room. Participants performed the experiment in 2 phases. In the *practice phase* participant completed a 2AFC temporal-order (did the light or sound come first?) judgment on an auditory and visual stimulus with SOAs of 0, ± 20, ± 90, ± 170, ± 250, or ± 350 ms, that were repeated 6 times in order to extract a psychometric function. Participants responded by pressing one of two keys in order to indicate whether the light or sound was perceived first. The next trial would begin 1.5 or 2 s after the response key had been pressed. The purpose of the practice phase was to identify whether the participant could successfully discriminate temporal order. Data was not analysed if the ability to discriminate the temporal order (JND) was above 250 ms in the practice phase or 200 ms in the test phase. In the *test phase*, one group of participants was exposed to sequences of stimuli that were perceptually isochronous (jitter = 0, 10, 20 or 30 ms) whilst a second group was exposed to perceptually anisochronous sequences (jitter = 0, 50, 100 or 150 ms). In each trial, four auditory stimuli were presented with an

3.5.2 Experiment 2

Twenty-four undergraduate students were recruited via the University of Birmingham research participation system and received course credits for their participation. All participants reported normal or corrected-to-normal hearing and vision and were all naïve to the purpose of the study.

The auditory stimuli were five identical tones produced by a speaker approximately 50 cm to the left of the participant (20 ms with 5 ms linear ramp, 1 kHz, 75.1 dBA). Visual stimuli were red flashes of light produced by a 5 mm LED (20 ms with 5 ms linear ramp, 91 Cd/m[2)] positioned directly in front of the participant and 20 cm to the right of the auditory stimuli. The signals were generated on MATLAB using the psychophysics toolbox extension (Brainard, 1997; Pelli, 1997) and sent to the audio card before presentation to ensure reliable timing.

Participants sat in a well-lit and quiet room. Participants performed the experiment in 2 phases. In the *practice phase* participant completed a 2AFC temporal-order (did the light or sound come first?) judgment on an auditory and visual stimulus with SOAs of 0, ± 20, ± 90, ± 170, ± 250, or ± 350 ms, that were repeated 6 times in order to extract a psychometric function. Participants responded by pressing one of two keys in order to indicate whether the light or sound was perceived first. The next trial would begin 1.5 or 2 s after the response key had been pressed. The purpose of the practice phase was to identify whether the participant could successfully discriminate temporal order. Data was not analysed if the ability to discriminate the temporal order (JND) was above 250 ms in the practice phase or 200 ms in the test phase. In the *test phase*, one group of participants was exposed to sequences of stimuli that were perceptually isochronous (jitter = 0, 10, 20 or 30 ms) whilst a second group was exposed to perceptually anisochronous sequences (jitter = 0, 50, 100 or 150 ms). In each trial, four auditory stimuli were presented with an

IOI of 700 ms. For both groups, 25% of the trials comprised perfectly isochronous sequences (no temporal jitter). The participant's task was to decide whether the last auditory stimulus in the sequence was before or after a visual probe stimulus (2AFC: did the sound or light come first?). The audio and visual stimuli were presented with one of 9 SOAs (0, ±40, ±80, ±120, or ±200 ms). Critically, the last auditory stimulus in the sequence was presented ± 40 ms earlier or later than expected (negative values mean the stimulus was presented earlier than expected). Each condition was repeated 10 times in order to construct a psychometric function.

3.5.3 Psychophysical Analysis

In order to see a change in the perceived timing of stimuli, the point of subjective simultaneity (*PSS*) from a psychometric function was obtained. Each psychometric function was established for each participant by calculating the cumulative proportion of responses for each *SOA*. The proportions of responses in each psychometric function were monotonized prior to analysis. We define the *PSS* as the *SOA* at which each subject was equally likely to report that the visual probe or last repeated stimulus was first. Negative *PSS* values mean that sound had to be presented before the light in order to be perceived as synchronous whilst positive values mean that light had to be presented before sound to be perceived as synchronous. We use the Spearman-Kärber method to obtain the *PSS* and *JND*: estimated as the first and second moments of the psychometric function differential across *SOAs* (see Miller & Ulrich, 2001) for further description of method). This *PSS* estimation is a non-parametric method that does not make assumptions about the distributions that underlie the obtained psychometric functions. A mathematical derivation of the method can be expressed as such: First, we define SOA_i with $i=\{1, ... 15\}$ as the 15 values of SOA used in the experiments and p_i with $\{i=1, ... 15\}$ as the associated proportion of 'light first' responses. We further define $SOA_0 = -250$ ms, $SOA_{16} = +250$ ms,

p_0=0, and p_{16}=1, so that we can compute the intermediate *SOA* between two successive ones

$$s_i = \frac{SOA_{i+1} + SOA_i}{2}, \text{ with } i=\{0, \dots 15\} \quad \text{(Eq. 3.1)}$$

and the associated values of the difference in proportion, taken at and above 0 to monotonize the proportion of responses

$$dp_i = max\ (0,\ p_{i+1} - p_i)\ , \text{ with } i=\{0, \dots 15\} \quad \text{(Eq. 3.2)}$$

With these indexes we can express *PSS* analytically as:

$$PSS = \frac{1}{\sum_{i=0}^{15} dp_i} \sum_{i=0}^{15} s_i dp_i \quad \text{(Eq. 3.3)}$$

And the *JND* as:

$$JND = \sqrt{\sum_{i=0}^{15} dp_i (s_i - PSS)^2} \quad \text{(Eq. 3.4)}$$

Data were not analyzed if participants had a *JND* above 250 ms or if *PSS* exceeded ± 175 ms when data were collapsed across conditions. The data of all participants were analysed.

3.6   Supplementary Information
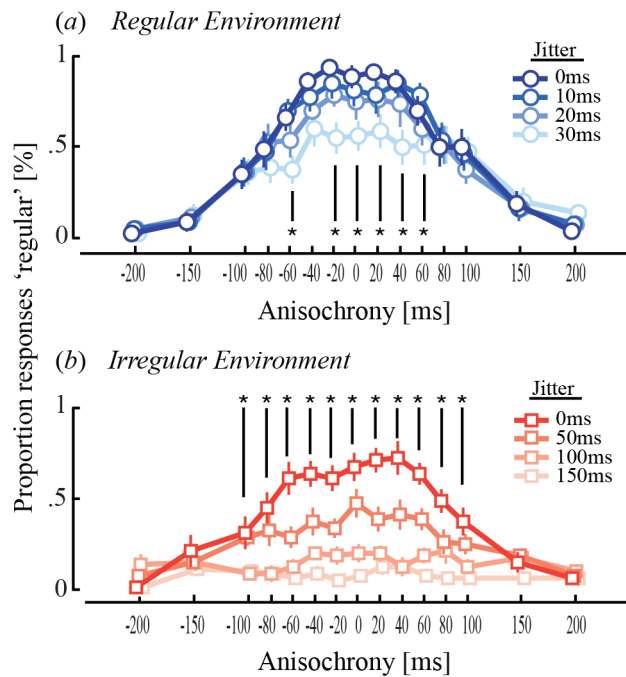
Supplementary Figure 3.1. Additional analysis of Experiment 1. Proportion of 'regular' responses as a function of the timing of the final stimulus in the sequence for (a) sequences of small jitter and (b) sequences of large jitter. Each line represents a different temporal jitter applied to the stimuli around an IOI of 700 ms. The distribution of responses becomes steeper with less relative jitter for both (A) small jitter and (b) large jitter sequences (interaction term of two-way repeated measures ANOVA on the inverse-normal proportion of 'regular' responses bounded between .01 and .99: Small jitter: $F_{42,378}$=3.5, p<.0001, η_p^2=.28; Large jitter: $F_{42,378}$=5.0, p<.0001, η_p^2=.36). Asterisks denote anisochronies at which the proportion of responses differs significantly across temporal jitters evidencing more irregular responses when a stimulus sequence is more jittered. See Supplementary Table 3.1 and 3.2 for statistical tests). In all graphs, error bars represent the standard error of the mean.

Supplementary Table 3.1. Analysis of the 'regular' responses at each anisochrony for different magnitudes of jitter in Experiment 1. *F* values and Bonferroni-corrected *p* values (*p* values are multiplied by 15) for one-way r.m. ANOVA on the proportion of responses in the four conditions at each level of anisochrony. Asterisks denote significant p values at 5% alpha level.

Anisochrony [ms]	Small Jitter			Large Jitter		
	$F_{3,27} =$	$p=$	$\eta_p{}^2=$	$F_{3,27} =$	$p=$	$\eta_p{}^2=$
-200	0.7	>0.99	0.07	4.2	0.228	0.32
-150	0.3	>0.99	0.04	0.7	>0.99	0.07
-100	0.1	>0.99	0.02	7.1	*0.017	0.44
-80	1.2	>0.99	0.12	11.9	*0.001	0.57
-60	12.3	*0.001	0.58	25.5	*0.001	0.74
-40	4.3	0.198	0.32	37.0	*0.001	0.80
-20	14.0	*0.001	0.61	25.2	*0.001	0.74
0	7.1	*0.017	0.44	24.4	*0.001	0.73
20	6.1	*0.039	0.40	32.7	*0.001	0.78
40	12.1	*0.001	0.57	26.0	*0.001	0.74
60	6.7	*0.023	0.43	38.8	*0.001	0.81
80	0.1	>0.99	0.01	12.8	*0.001	0.59
100	1.4	>0.99	0.14	9.4	*0.003	0.51
150	0.2	>0.99	0.02	4.3	0.208	0.32
200	3.4	0.473	0.28	0.2	>.99	0.02

Supplementary Table 3.2. Analysis of the "regular" responses at each anisochrony comparing 0ms (isochronous) jitter between the groups exposed to small and large jitter Experiment 1. *T* values and Bonferroni-corrected *p* values (values are multiplied by 15) for independent-sample t-tests on the proportion of responses between the two groups at each level of anisochrony. Asterisks denote significant p values at 5% alpha level.

Anisochrony [ms]	0ms Jitter – Difference between Small and Large Groups	
	$T_{18} =$	$p=$
-200	-0.6	0.556
-150	-0.1	0.237
-100	-0.3	0.760
-80	-0.3	0.778
-60	-0.4	0.671
-40	-2.6	*0.019
-20	-3.9	*0.001
0	-2.1	*0.047
20	-2.5	*0.021
40	-1.2	0.242
60	-0.6	0.546
80	-0.1	0.922
100	-1.0	0.336
150	-0.4	0.680
200	0.5	0.567
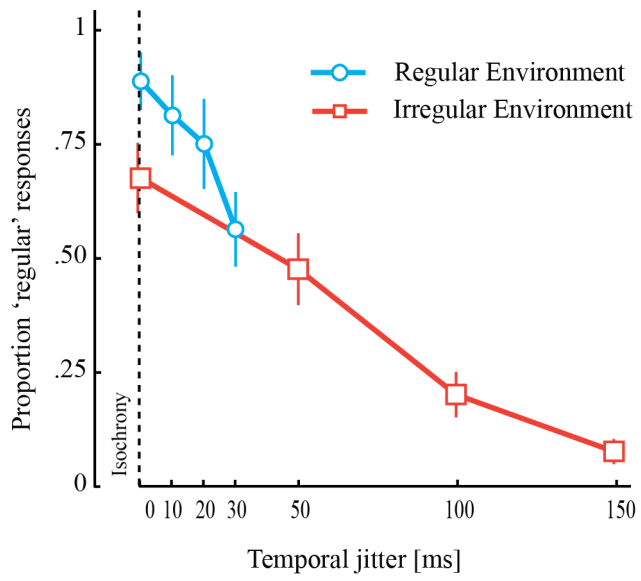
Supplementary Figure 3.2. Illustration of the proportion of regular responses without additional anisochrony of the last stimulus as a function of the temporal jitter in the sequence for both the regular environment (blue line) and irregular environment (red line).
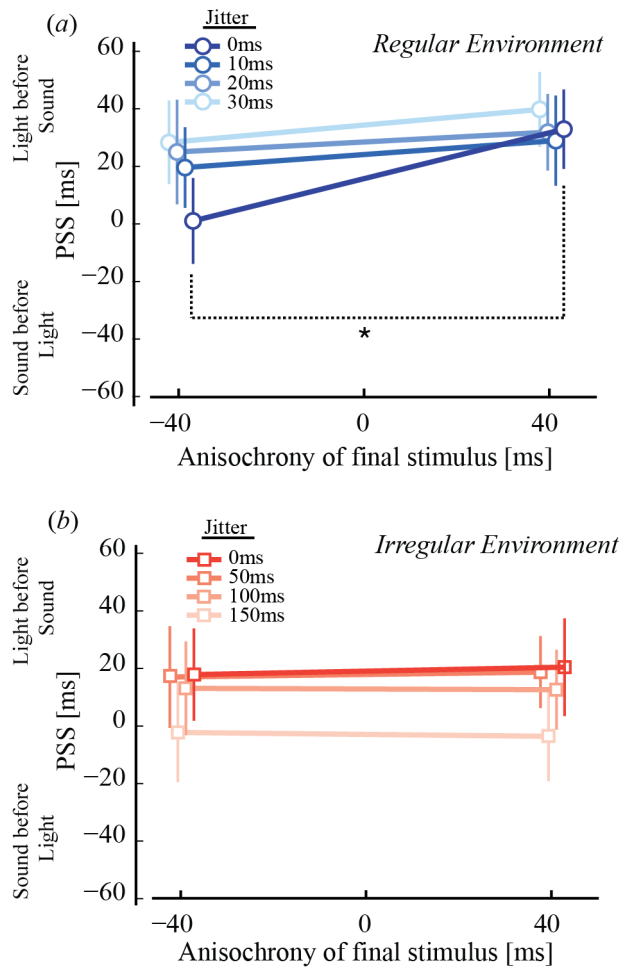
Supplementary Figure 3.3. Additional analysis for the results of Experiment 2. Average *PSS* values corresponding to the SOA at which both the audio and visual stimulus were perceived as simultaneous. Positive *PSS* values indicate that the light needs to be presented before the sound to be perceived as simultaneous whilst negative values indicate that the sound has to be presented before the light to be perceived as simultaneous. *PSS* values should be horizontal and not change if the stimulus anisochrony does not affect the perceived timing, but instead, the *PSS* values only change for 0ms jitter in the (a) small jitter group whilst no such change is exhibited in the (b) large jitter group (Interaction term of two-way repeated measures ANOVA for factors of anisochrony and jitter level in the small jitter group: $F_{3,33}=1.2$, $p=.33$, $\eta_p^2=.10$. Main effect Anisochrony: $F_{1,11}=8.5$, $p=.01$, $\eta_p^2=.43$, Main effect Jitter: $F_{3,33}=3.6$, $p<.02$, $\eta_p^2=.25$.
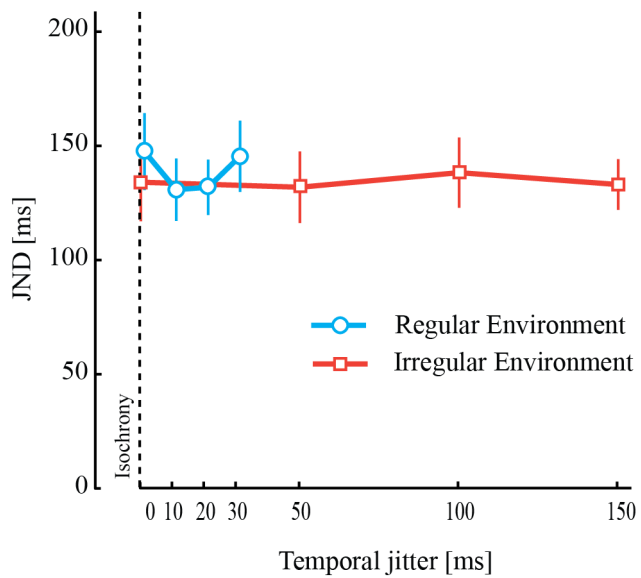
Supplementary Figure 3.4. Additional analysis of Experiment 2. JND values for small and large jitter group as a function of each anisochrony. The JND represents how well participants can discriminate between the audio and visual stimulus. A change in JND would suggest that the jitter was hampering their ability to discriminate between the stimuli and thus possibly cause changes in the PSS but we find no such effects as a two-way repeated measures ANOVA on anisochrony and temporal jitter for both the regular and irregular context groups revealed no main effects of an interaction (all $p > .1$).

Chapter 4 Bayesian Time Perception: Prior Expectations and Noisy Sensory Estimates Bias Perceived Timing

Here, we tested further intrinsic properties of the Bayesian model. In particular, we test how noise characteristics of stimuli can affect the perceived timing of stimuli.

4.1   Abstract

We propose a model of perceived timing where the brain represents stimuli as probability distributions over time. An inherent characteristic of this model is that perceived timing of low-intensity signals should be influenced more by a-priori expectations about the timing of stimuli more than signals of high intensity. To test this model prediction, we presented sequences of regularly timed auditory stimuli (isochronous sequences) with alternating amplitudes (low–high–low–…). Such a sequence generates an expectation that stimuli should be regular, but here the final stimulus of the sequence was instead presented slightly earlier or later than expected. The participant's task was to report the temporal order of the final stimulus compared to a visual probe presented with varying stimulus onset asynchronies. We find that perceived simultaneity of stimuli changes suggesting that early stimuli are delayed and late stimuli are accelerated. Furthermore, we find that perceived timing of low intensity stimuli are regularized more so than for high intensity stimuli, in line with the idea that the brain uses Bayesian priors about stimulus timing for perception.

4.2   Author Summary

How does the brain work out the timing of events? Sequences of regularly timed stimuli can inform us about the probable timing of future events. We hypothesized that the brain uses a Bayesian inference process that tracks such regularities with dynamic prior expectations to predict the timing of future events. Perception is determined from the combination of temporal expectations with current sensory evidence. This framework suggests that stimuli with slight deviations from temporal regularity should be biased towards expectations and further, that stimuli whose timing is uncertain should be regularized even more. In order to test this prediction, we introduce a paradigm that measures the perceived timing of a stimulus in a sequence. We find that earlier- and later-than expected stimuli are temporally regularized and as predicted, we find that less reliable stimuli are more regularized. Together, the data and model suggest a parsimonious explanation of time perception and provide further evidence that perception is the result of Bayesian inference.

4.3   Introduction

The environment is constantly evolving around us, providing a stream of sensory information that the brain has to process, analyse and respond to. The task of efficiently managing this information is important to survival. One way the brain can increase efficiency is to take advantage of temporal regularities in the world, which leads to decreased neural metabolic consumption (VanRullen & Dubois, 2011), and in automatizing rhythmic behaviour for activities like speech, music production, tapping and dance (McNeill, 1995; Repp, 2005). The simplest type of temporal regularity in the environment is the occurrence of repeated stimuli with the same inter-onset interval (IOI). The brain can use the regularity derived from regular sequences to improve stimulus detection and discrimination (Brochard et al., 2013; Carnevale et al., 2015; Correa et al., 2005; Cravo et al., 2013; Escoffier et al., 2010; Jazayeri & Shadlen, 2010; Rohenkohl & Nobre, 2011) and to drive faster responses (Jazayeri & Shadlen, 2010; Lakatos et al., 2008; Miyazaki et al., 2005). However, the neural mechanisms that underpin how the brain can estimate event timing remain unknown. Here we propose that the brain rapidly generates temporal expectations about future stimuli and uses such expectations to bias perception in a way that leads to the temporal regularization of slightly irregular stimulus sequences.

Humans make decisions about temporal properties all the time. We can estimate how much time has passed since a movie started, what the current time is, and how fast objects are moving. However, humans do not perceive time veridically (Allan, 1979; Grondin, 2010), and as such there is inherent inaccuracy in temporal judgments. The perceived timing of a stimulus – the point at which an event is thought to have occurred – has a limited resolution. Temporal processing is subdivided into discrete epochs: two stimuli occurring within a processing epoch will be judged as simultaneous, whilst reported as sequential if occurred within two or more epochs (Hirsh & Sherrick, 1961;

Kristofferson, 1967; VanRullen & Koch, 2003). Given the limits of perceptual accuracy in the dimension of time, here we consider such distortions in the perceived timing of stimuli in isochronous sequences and we propose a Bayesian model that explains how the brain may deal with such an intrinsic problem in arriving at a percept of time.

Whilst temporal accuracy is limited, the uncertainty of the time at which a signal arrives after being produced adds another layer of complexity into forming a percept of time. Uncertainty about the timing of a signal arises from two sources: External noise, i.e. from random variability in the compression and rarefaction of longitudinal airwaves; and secondly, internal sensory noise from neural transmission and subsequent encoding of temporal information (Green & Swets, 1973). Thus, it has been argued that a fundamental purpose of the brain is to minimize uncertainty and derive a 'best guess' of the state of the world (Helmholtz, 1963).

To arrive at an optimal estimate, it has been proposed that the brain employs sensory mechanisms that work in a way consistent with *Bayesian Decision Theory* (Bayes, 1763). In Bayesian decision theory an optimal observer for estimating when an event occurred can be defined by two probability distributions (Fig. 1a). First, the *likelihood function $p^l(t)$* is the probability of sensing a stimulus at time *t* given that the stimulus is produced in the environment. Second, the *prior probability distribution $p^p(t)$* is the knowledge of encountering a stimulus at a time in the future that increases rapidly with successive stimulus presentation (Fig. 4.1b). According to Bayes' rule (Bayes, 1763), the multiplication of the prior and likelihood function leads to the formulation of the *posterior distribution (*of which the observer selects a value of *t,* which commonly equates to the mean or mode of the posterior, Fig. 4.1cde). This distribution is then employed to obtain the estimate of perceived stimulus timing. Such a framework has been proven consistent with human performance in several perceptual domains (Kersten & Yuille, 2003; Knill &

Richards, 1996; Mamassian et al., 2002; Petzschner & Glasauer, 2011; Wolpert & Ghahramani, 2000), but here we apply it to the temporal dimension, and specifically, the perceived timing of stimuli.
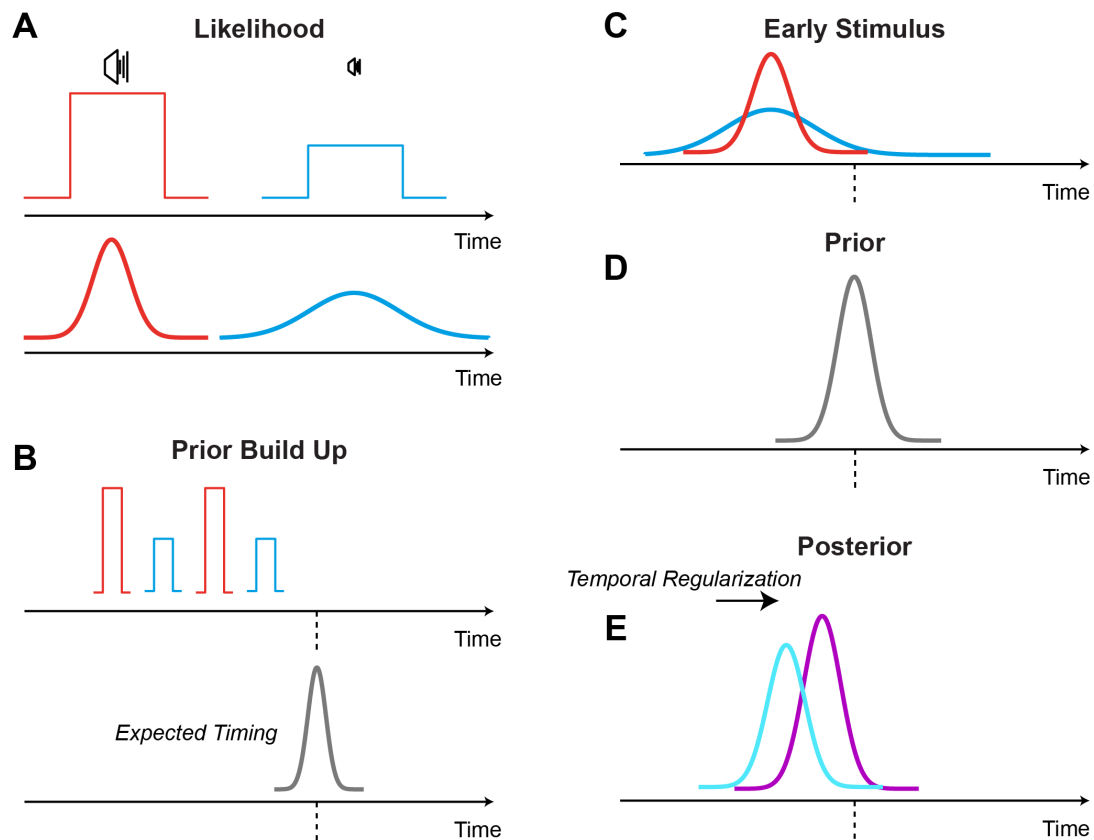
Figure 4.1. A computational model of perceived timing. (a) Illustration of two likelihood functions that describe sensory estimates with (left; red) low noise, and (right; blue) high noise. (b) Dynamic prior probability distributions that dynamically increase with each presentation of a stimulus. (c) High and low noise stimuli with corresponding likelihood functions that are presented slightly earlier than expected and combined with the prior (d) that is biased towards the expected timing (dotted line). (e) Posterior distributions show the temporal regularization effect and as such the perceived timing of a high noise stimulus is delayed closer to expectation (purple distribution) than a stimulus with low noise (cyan distribution).

From this framework we expect that the prior should bias temporal perception so that slightly irregularly timed stimuli are perceived closer to the expected time.

Furthermore, an intrinsic property of this model is that the regularization should be modulated by the precision of the temporal estimate that can be obtained from a signal. Low intensity stimuli have lower signal-to-noise ratio, which should lead to a likelihood with a broader distribution and thus a less precise estimate of temporal properties. As such, the combination of the likelihood with the prior should lead low intensity stimuli to be captured more by the prior and thus to be perceived closer in time to the expected time point than high intensity stimuli (Fig. 4.1e; Fig. 4.2c).
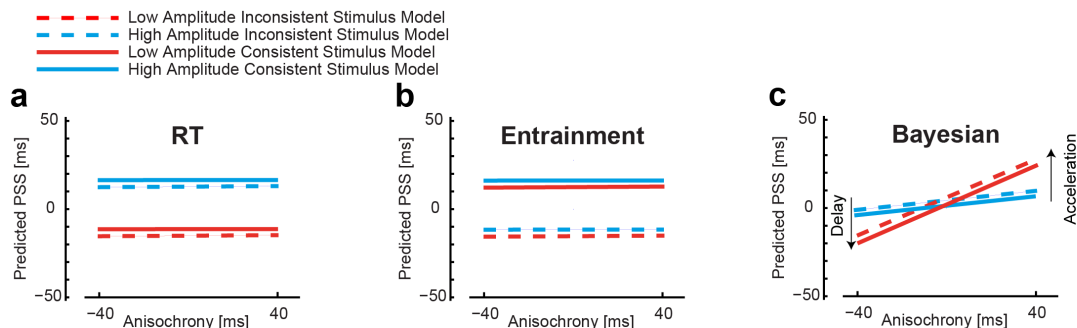
Figure 4.2. Comparison of predicted models PSS estimates. (a) A Reaction Time (RT) account of temporal perception predicts that stimuli that are of high amplitude are processed quicker than stimuli of low amplitude whilst Entrainment models (b) predict that stimuli that are consistent with the temporal pattern are processed quicker. (c) A Bayesian account of temporal perception, predicts that stimuli of lower amplitude are captured more so by a prior distribution and as such early stimuli are delayed towards expectation whilst late stimuli are perceptually accelerated – an effect similar to temporal regularization.

In this paper, we aimed to validate such a prediction, showing that prior expectations about timing of stimuli bias perception in a way that has not been considered previously, as current accounts of time perception (Buhusi & Meck, 2005; Buonomano & Merzenich, 1995; Creelman, 1962; Gibbon, 1977; Gibbon et al., 1984; Karmarkar & Buonomano, 2007; Meck, 2005; Shi et al., 2013; Teki et al., 2011; Treisman, 1963; van Rijn et al., 2014; Wackermann & Ehm, 2006) do not predict changes in the perceived

timing of stimuli. Further, we hypothesize that stimuli that are less reliable, and as such

noisier, should be captured more by prior expectations and therefore temporal

regularization for them should be stronger. We devised a paradigm that measures the

perceived timing of a stimulus in a sequence and we find that stimuli with low amplitude

are perceived closer to regularity whereas the prior less affects stimuli with higher

amplitude. We were able to reverse engineer the parameters of a Bayesian estimator that

successfully describes the experimental data in our experiment and gives more credence to

the idea that the brain is 'Bayesian' and uses statistical inference to make a best estimate of

the world.

Extant models of time perception do not explicitly make predictions about changes

in the perceived timing of stimuli however they could be formulated to predict such

changes. Models based on reaction time (RT), predict that the RT to high intensity stimuli

is shorter than those of lower intensity (Kohfeld, 1971; Pins & Bonnet, 1996; Ulrich,

Rinkenauer, & Miller, 1998) – the rationale behind this is that stimuli of high intensity are

processed quicker. As such, one could predict that stimuli that are high intensity are

perceived earlier than relatively lower ones (Fig. 4.2a). Entrainment models of time

perception, on the other hand, assert that the brain adjusts its internal oscillations in a way

that the peak aligns to the phase of external stimulation (Large & Jones, 1999; Large &

Palmer, 2002; Large & Snyder, 2009) that is in keeping with an underlying temporal

pattern. Attentional resources are deployed at the expected time (Henry & Herrmann,

2014; Rohenkohl & Nobre, 2011) and as such it may be predicted that stimuli that are

presented consistently in a temporal structure are processed faster in comparison to stimuli

that are inconsistent with the temporal structure (Fig. 4.2b). Contrastingly, the Bayesian

model of perceived timing we present in this paper predicts that stimuli with low

amplitude should be captured more by the expected timing and as such – regardless of sequence consistency – should be perceptually regularized (Fig. 4.2c).

4.4   Results

The proposed Bayesian model of perceived timing predicts that slightly irregular stimuli should be delayed towards expectation, but stimuli that carry a broader likelihood function (more uncertain timing) are comparatively more regularized. We show in this section that subjects do change the time at which they perceive a stimulus before fitting a Bayesian model of perceived timing to show that changes in the shape of the likelihood function due to sensory uncertainty explain the pattern of experimental data.

4.4.1 Behavioural Experiment

To investigate both the effect of increasing temporal expectations and temporal noise on *perceived timing*, we presented regularly timed sequences of stimuli with alternating amplitude. The intensity of the last stimulus could either be consistent or inconsistent with the preceding alternating amplitude (Fig. 4.3). We presented consistent/inconsistent sequences in order to control against subjects forming contingent expectations that a low intensity stimulus always follows a high stimulus and vice versa. Moreover, the last stimulus was presented with a slight conflict compared to the expected timing, i.e. it could be earlier than or later than expected ($\pm$ 40ms). The assumption of this methodology is that temporal expectations increase with each successive presentation of each stimulus, analogous to the priors as presented in Figure 4.1a. Subjects judged whether the final stimulus in the series of seven alternating high/low amplitude stimuli was before or after a visual probe stimulus and subjects reported the temporal order of this pair (i.e. was the sound or light first?).
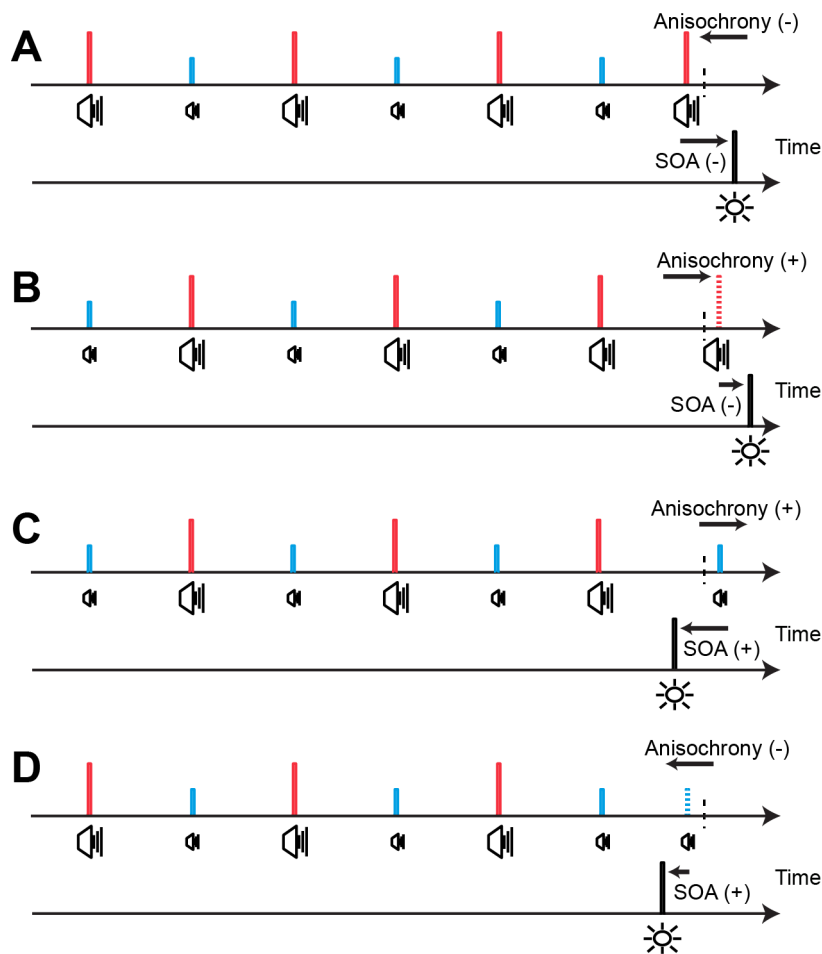
Figure 4.3. Schematic of the Experimental Procedure. Subjects were presented a sequence of isochronous stimuli that alternated in amplitude. The subjects' task was to report the temporal order of the audiovisual pair of stimuli at the end of the sequence. A Positive SOA indicates that the light was presented before the final auditory stimulus. The final stimulus in the sequence could be (a) a 'high' amplitude stimulus that was consistent with the sequence (example: negative anisochrony, negative SOA), (b) a 'high' amplitude stimulus that was inconsistent with the sequence (example: positive anisochrony, negative SOA), (c) a 'low' amplitude consistent with the sequence (example: positive anisochrony, positive SOA) or, (d) a 'low' amplitude stimulus that was inconsistent with the sequence (example: negative anisochrony, positive SOA).

We calculated the proportion of 'light first' responses as a function of the physical difference in timing between the final auditory stimulus and the probe (SOAs). We considered the combination of three independent variables: whether the intensity of the final stimulus was consistent or inconsistent with the sequence, whether the stimulus was

of low or high amplitude, and whether the final stimulus was presented earlier or later than expected. From these psychometric functions obtained in each of these 8 conditions we find the stimulus onset asynchrony (SOA) between the final stimulus and visual probe necessary for the *Perception of Subjective Simultaneity* (PSS; see Materials and Methods), that is, the SOA that corresponds to where subjects are maximally unsure about whether the auditory or visual stimulus was first. Consequently, the PSS allows us to capture the perceived timing of the repeated stimulus – and subsequent changes in the PSS as a function of the anisochrony at which it is presented tells us that the the brain may regularize temporally deviant stimuli.

Figure 4.4a presents the average PSS values of each condition as a function of whether the final stimulus was earlier or later than expected. The PSS values for each condition were assessed using repeated-measures ANOVA with factors anisochrony, sequence consistency, and amplitude of final stimulus. We found a difference in the perceived timing between early and late stimuli (main effect: Anisochrony; $F(1,17)=20.6$, $p=.0003$, $\eta p^2=.55$). The comparison of PSS values between early and late presentation for each condition suggests that late stimuli are perceived faster (accelerated) whilst early stimuli are perceived slower (delayed). Thus, the perceived timing of stimuli is changed to regularize small temporal deviations. Critically, this effect is modulated by the amplitude of the final stimulus (interaction term of anisochrony and stimulus amplitude; $F(1,17)=7.9$, $p=.012$, $\eta p^2=.33$; (Fig. 4.5) with no effect of sequence consistency (interaction term of sequence consistency and anisochrony, $F(1,17)=0.5$, $p=.49$, $\eta p^2=.01$ (Fig. 4).
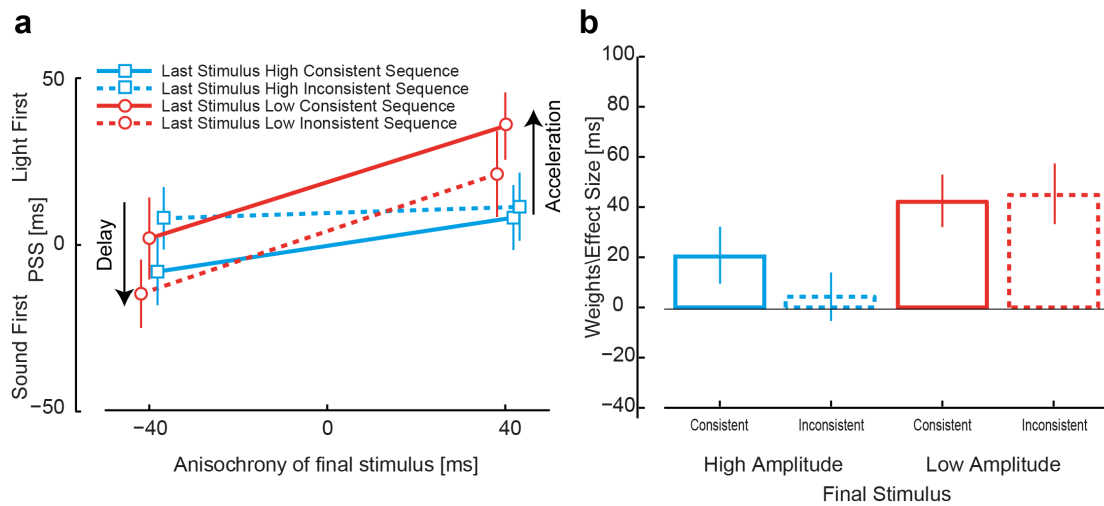
Figure 4.4. Subjects' Average PSS and Weights. (a) Subject's average PSS values corresponding to the Stimulus onset asynchrony (SOA) that visual and auditory stimuli are perceived as being simultaneous. A positive PSS indicates that light must be presented before the sound in order to be perceived as simultaneous. The difference between the PSS values of early (-40ms) and late (40ms) anisochronies on the X-axis indicates whether a change in perceived timing was evidenced. If there were not a change in perceived timing then the pattern of PSS for each condition would be horizontal. However, we find that only when the final stimulus is low intensity there is then a change in perceived timing (red lines). As the light has to be presented further in advance of the auditory stimulus to be perceived as simultaneous for late stimuli, this evidences a perceptual acceleration. When the asynchrony diminishes between the light and sound, then this evidences a perceptual delay – this is shown for early stimuli. The difference in the PSS values is reflected in (b) where the effect size (and as such the relative weights) show a strong change in perceived timing for when the final stimulus is of low amplitude.

To quantify the magnitude of the regularization effect, we calculated the effect size (E_i) given to the a-priori expectation with low and high intensity signals:

$$E_i = \frac{PSS_{early} \cdot PSS_{late}}{80} \qquad \text{(Eq. 4.1)}$$

where *i* indexes the effect size given to a sequence for each of the four conditions and PSS_{early} refers to the PSS when the final stimulus was -40ms earlier than expected whereas PSS_{late} is when the final stimulus was 40ms later than expected. The multiplication of these values is then divided by 80ms; which is the temporal difference

between both conditions. As such, the effect size represents how much the final stimulus likelihood is attracted by the prior. Figure 4.4B displays the calculated effect sizes for the expected time point in all conditions whereas Figure 4.5B illustrates the effect sizes when sequence consistency is collapsed. The effect sizes are significantly differed from 0 for the low amplitude stimuli, but they are not for high amplitude stimuli (single sample t-test, $t(17)=5.3$, $p<.001$; $t(17)=1.5$, $p=0.16$). Furthermore, the two weights differ from each other suggesting that perceived timing of signals with lower intensity are more biased towards the expected time point (paired sample t-test, $t(17)=-2.8$, $p=.012$).
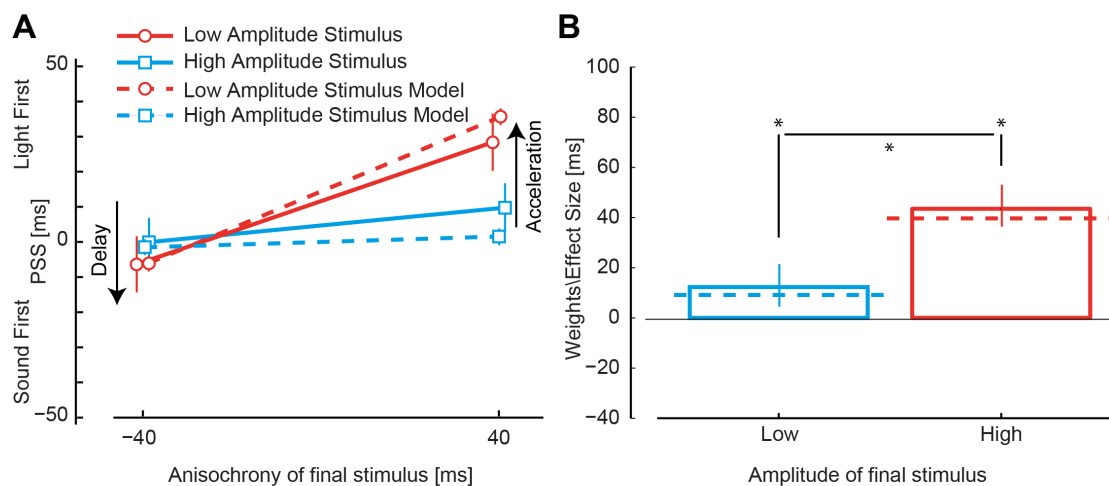
Figure 4.5. Collapsed PSS and Effect Sizes for High versus Low Intensity Final Stimuli and Model Predictions. (a) Subject's average PSS values corresponding to the Stimulus onset asynchrony (SOA) that visual and auditory stimuli are perceived as being simultaneous. A positive PSS indicates that light must be presented before the sound in order to be perceived as simultaneous. The difference between the PSS values of early (-40ms) and late (40ms) anisochronies on the X-axis indicates whether a change in perceived timing was evidenced. If there were not a change in perceived timing then the pattern of PSS for each condition would be horizontal. However, we find that only when the final stimulus is low intensity then there is a change in perceived timing (red lines). As the light has to be presented further in advance of the auditory stimulus to be perceived as simultaneous for late stimuli, this evidences a perceptual acceleration. When the asynchrony diminishes between the light and sound, then this evidences a perceptual delay – this is shown for early stimuli. The difference in the PSS values is reflected in (b) where the effect size show a strong

change in perceived timing for when the final stimulus is of low amplitude. Error bars represent the standard error of the mean.

4.4.2 Bayesian Decision Theory Model

Subjects' performance was modelled with *Bayesian Decision Theory* (BDT). This framework has been applied to different perceptual domains (Kersten & Yuille, 2003; Knill & Richards, 1996; Mamassian et al., 2002; Petzschner & Glasauer, 2011; Wolpert & Ghahramani, 2000) including temporal interval estimation (Hartcher-O'Brien et al., 2014; Shi et al., 2013) and reproduction (Jazayeri & Shadlen, 2010; Miyazaki et al., 2005). To conceptualize, we suggest that the brain represents the probability of experiencing stimuli at every time point. The likelihood $p^l(t)$ represents current sensory information and as such the time a stimulus was sensed at time t given it was produced in the environment. A Gaussian distribution with mean μ and variance σ^2 is used to describe the likelihood and noise (uncertainty) in sensory latency. The prior $p^p(t)$ represents the *a-priori* knowledge of the future timing of a stimulus. According to BDT, the likelihood $p^l(t)$ and prior probabilities $p^p(t)$ can be combined at each time point (Eq. 4.1), giving rise to the posterior probability from which a percept is obtained $p^q(t)$:

$$p^q(t) \propto p^l(t) \cdot p^p(t) \qquad\qquad \text{(Eq. 4.2)}$$

Similar to a Kalman filter, the prior is dynamic and recursively updates after each stimulus presentation by using the posterior probability $p^q(t)$ for the previous stimulus (i.e., $p^p(t)$ for the time *t-IOI'*). The mean of the prior corresponds to the point a stimulus is expected to be sensed. The *a-priori* probability of a stimulus is modelled to be flat when the first stimulus is presented but increases after each stimulus presentation. This is expressed by:

$$p^p(t + IOI') \propto p^q(t) + \omega \qquad\qquad \text{(Eq. 4.3)}$$

As more information is acquired, the prior builds and begins to approach the shape of the likelihood function after a few stimulus presentations until it eventually eclipses it and becomes less broad and steeper. The added constant ω leads to a heavy-tailed prior that allows for abrupt changes in IOI and as such decreases the tendency of entirely assimilating the posterior into a new prior (Roach et al., 2006). It follows that when the likelihood is displaced from the mean of the prior – the posterior is biased towards the prior. Further, if the likelihood function becomes less reliable (has a greater variance), then it is biased more towards the prior than a likelihood function with comparatively greater reliability.

4.4.3 Likelihood and Prior Estimation from Experimental Data

From the experimental data, we find that stimuli of low amplitude (i.e. less reliability / lower signal to noise ratio) are biased by the prior distribution more so than stimuli of higher amplitude (Fig. 4.4 and 4.5). We then determined the parameters of the proposed model that best fit such data. It should be observed that the model has only two parameters: ω and sigma, the variance of the likelihood function. We obtained ω and found σ_h^2 for the high intensity and σ_l^2 for the low intensity stimuli using a grid search over the parameter space that produce a shift in *PSS'* in the model that is similar to the one we registered experimentally.

To obtain the predicted *PSS'* for each participant in each condition, we simulated the probability distributions for each of the stimuli presented in the sequence. For this, we applied Equations 4.2 and 3 iteratively alternating σ_h^2 and σ_l^2 for the likelihoods of the high and low intensity stimuli, whilst adding the constant ω at each iteration. We then cumulate the posterior probability distributions for the last stimulus in the sequence. The 50% point is the *PSS'*. We jointly determine ω, σ_h^2 and σ_l^2 that minimize the difference with the empirical *PSS* for each participant. We obtain average values of $\sigma_h^2 =$

66.5000±0.6239ms for a high intensity stimulus and $\sigma_l^2 = 379.9444\pm0.5024$ms for a low intensity stimulus, and ω as 0.0118±0.0031. We present the model *PSS′* with the best-fit parameters in comparison to the empirical PSS in Figure 4.5A, and the resultant weights in Figure 4.5B. We find that the model PSS accurately captures the pattern of empirical data in a way consistent with predictions of the model – the best fitting parameters yield that low amplitude stimuli have an increased standard deviation of the likelihood function governing it, whilst the high amplitude likelihood function is best fit with a Gaussian distribution with a smaller standard deviation.

4.5   Discussion

Our results demonstrate that a Bayesian estimator model can accurately capture subjects' data in a human perceived timing task. We find that slightly irregularly timed stimuli are regularized to be perceived closer to expectation. Critically, we find that stimuli with low amplitude have strong regularization in comparison to stimuli with high amplitude. In accordance to the model detailed in the introduction and results section – we find the pattern of results best fits a Bayesian model of perceived timing interpretation in comparison to models based on reaction time or entrainment.

The proposed Bayesian model is based on the dynamic updating of temporal expectations after each presentation of a stimulus. Temporal expectations are combined iteratively with noisy internal representations of current sensory information, which results in the perceived timing of stimuli. The model can be utilized as a tool to fit human perceived timing data in perceptual decision tasks that are based on the method of constant stimuli as well as serving as an antecedent for future scientific inquiry into temporal perception.

4.5.1 Validation of Model Predictions

The model makes two primary predictions that we subsequently validated with empirical data. Firstly, the model implicitly predicts that stimuli that are presented slightly earlier or later than expected should be captured by prior expectations in a way that regularizes their perceived timing. Secondly, implicit in the model, is that shallower likelihood functions should be dominated more by the prior and as such we should see a strong effect of temporal regularization when a low intensity stimulus is the final temporally deviant stimulus.

To test these predictions, we devised a paradigm that allowed us to precisely measure the perceived timing of stimuli whilst varying the intensity of the final stimulus. To do this, we presented a sequence of stimuli with alternating amplitudes (low–high–low–…). The use of alternating amplitudes allows us to disentangle the signal reliability of the prior and likelihood functions – as the prior is based on the recursive updating of itself given the posterior of the previous stimulus. A sequence of identical stimuli cannot disentangle the prior from the likelihood function and as such it would be hard to conclude anything concrete about whether the prior or likelihood caused the changes in perceived timing reported in this paper.

To precisely measure perceived timing, the final stimulus was paired with a visual probe and subjects' were asked to report the temporal order of this pair (Figure 4.1). Such a method allows us to calculate the Point of Subjective Simultaneity (PSS) of the final repeated stimulus. In order to validate the predictions of our model, we altered the timing of the last stimulus so it could be earlier or later than expected. Such an alteration in the physical timing should result in a change in the PSS in a way that means that stimuli presented earlier than expected are perceptually delayed, whilst stimuli presented later than expected are perceptually accelerated. The objective synchrony of audiovisual stimuli

differs markedly to the perceived synchrony in the brain. Sound and light differ in their propagation velocity from source to sense organ, but also differ in their processing and transduction times in the brain (Allison et al., 1977; A. J. King, 2005; Spence & Squire, 2003; Sternberg & Knoll, 1973). As such, it is often reported that a visual stimulus should precede an auditory one in order to be perceived as simultaneous (Allison et al., 1977; A. J. King, 2005; Spence & Squire, 2003; Zampini et al., 2003) – the PSS. An increase in the time that a visual stimulus precedes an auditory stimulus to be perceived as simultaneous exemplifies a perceptual acceleration of the auditory stimulus, however a reduction in the asynchrony means that the auditory stimulus is perceptually delayed. Our data shows an effect consistent with temporal regularization and the predictions of our model – as final stimuli with low amplitude are perceptually delayed (reduced vision-audio PSS) if they are presented earlier than expected and perceptually accelerated (increased vision-audio PSS) if they are later than expected.

4.5.2 Interpretation to Existing Models of Temporal Perception

Timing can be conceptualized as meaning both *how long* an event lasted, or contrastingly, *when* an event occurred (Merchant & de Lafuente, 2014). Historically, both interval (Gibbon et al., 1984; Treisman, 1963) and entrainment (Buonomano & Merzenich, 1995; Karmarkar & Buonomano, 2007; Large & Jones, 1999; C. Miall, 1989) accounts of time perception are concerned with how long a discrete duration is perceived to have lasted, and as such are not concerned with changes in the perceived timing of the stimuli that delimit such intervals. Rather, they predict changes in the perceived duration between two events and that sensitivity to temporal irregularities increases as a function of the number of intervals in a sequence (Large & Palmer, 2002; Schulze, 1989).

In reaction time (RT) models, Faster RTs are elicited by stimuli with higher stimulus intensities (Cattell, 1886; Chocholle, 1940; Kohfeld, 1971). Stimuli with lower

amplitude are slower to respond to and harder to detect due to the increased noise in the estimate and as such, are perceived later than higher amplitude stimuli (Piéron, 1913; 1952; Pins & Bonnet, 1996). Interval-models, such as the extensively researched Scalar Expectancy Theory (SET), consider duration perception as the accumulation of pulses from a pacemaker per unit of time (Church et al., 1994; Gibbon, 1977; Gibbon et al., 1984). In this framework, more intense stimuli increase the rate of an internal pacemaker leading to longer duration estimates but faster reaction times. The prediction therefore is solely based on the fact that signal intensity should increase the processing speed of high intensity signals (Fig. 4.2a).

In the data of this experiment, however, we find that stimuli of lower amplitude are perceived faster than high amplitude stimuli if a stimulus is presented later than expected (Fig. 4.5). In concert with the RT accounts, we find that low amplitude stimuli are perceived slower than high amplitude stimuli when presented earlier than expected. A prior representing expected timing could explain such an effect – which has not been considered in RT or SET models. The prior regularizes slightly deviant stimuli but captures stimuli with low amplitude more and, as such, indicates that such stimuli may have a shallower likelihood function underlying them.

Entrainment models, on the other hand, assert that internal oscillators adjust to the frequency and phase of external stimulation (Jones & Boltz, 1989; Large & Palmer, 2002; Large & Snyder, 2009). Given the entrained pattern, exogenous attention is deployed at the expected time (Henry & Herrmann, 2014; Rohenkohl et al., 2011). The rhythmic deployment of attentional resources at the expected time results in a prior-entry style effect (Spence & Parise, 2010) resulting in perceptual acceleration of stimuli that occur consistently in the observed temporal pattern (Fig. 4.2b). We find, however, that weak signals irrespective of their sequence consistency are actually delayed when presented

slightly earlier than expected. Our data, therefore suggest that entrainment models need to incorporate the effect of temporal regularization into temporal processing as they only appeal to perceptual acceleration of stimuli.

4.5.3   Future Directions

Unlike previous Bayesian models investigating the perception of duration (Jazayeri & Shadlen, 2010; Miyazaki et al., 2005; Shi et al., 2013), this paper, elucidates how the brain can formulate an estimate of *when* an event occurred. As such, the Bayesian model of perceived timing makes explicit predictions that require further scientific inquiry and testing.

The model is based on the idea that temporal expectations (the priors) build iteratively based on the repeated presentation of isochronous intervals. As such, if sequences of anisochronous stimuli are presented, then the prior should not be built within a sequence and as such, should not bias the timing of stimuli. Further, we find in the natural environment that a variety of signals that can be expected – but don't arrive isochronously – i.e. the first three intervals of a sequence could are 500ms whilst the fourth is always 800ms. Are prior expectations limited to isochronous, rhythmic presentation or are they built hierarchically? Or does the temporal context influence the perception of such things as regularity or perceived timing. Further, do multiple priors develop over multiple temporal patterns? Neural oscillations in the brain offer an interesting way of representing such priors, as the phase of delta-theta band activity could be a plausible neurophysiological mechanism for their implementation (Arnal & Giraud, 2012; Arnal, Doelling, & Poeppel, 2014) as well thalamo-cortico-striatal circuits that have also been showed to encode interval timing as well as numerosity (Buhusi & Meck, 2005; Matell & Meck, 2004). Bayesian inference has been shown to operate at the neural level

(Ma, Beck, Latham, & Pouget, 2006), thus developing a neurophysiological model of temporal perception is paramount in understanding how the brain may perceive time.

In this paper, we used repeated auditory stimuli to build temporal expectations– to further characterize this model; one may present stimuli from different modalities. The temporal acuity of visual stimuli, for example, is known to be less than for auditory stimuli (Hartcher-O'Brien et al., 2014) – as such, one could posit that the likelihood function that represents visual information is broader than for auditory stimuli: does the prior still build within-modality or is it amodal? One could devise an experiment whereby presenting interleaved sequences of different modalities and measuring the perceived timing of stimuli may help clarify such a position.

To further characterize the temporal profile of the regularization effect for temporally deviant stimuli, the next key step would be to present a range of anisochronies of the final stimulus and understand at what point stimuli are not regularized, as well as seeing if such an effect is symmetrical around 0 anisochrony (on-time stimuli). One might consider that stimuli presented at the expected time are not regularized – as two Gaussian probability distributions i.e. a prior and likelihood as per the Bayesian model with identical means, result in a posterior distribution with the same mean – and as such, no change in perceived timing. However, if we consider the concepts of prior entry (Spence & Parise, 2010) and active sensing (Schroeder et al., 2010; Schroeder & Lakatos, 2009), one could posit that expected stimuli are actually perceptually accelerated. The waxing and waning of excitation/inhibition cycles in the active sensing framework, where high excitability phases of neural processing phase are representations of the peak of an attentional pulse, facilitate sensory selection and processing of stimuli that coincide within such excitation phases (Henry & Herrmann, 2014; Lakatos et al., 2008). Therefore, one can reason that if a stimulus occurs at the expected time, and as such the high excitability phase, then it should

be given a perceptual boost and processed faster. This type of effect, is similar to prior entry (Spence & Parise, 2010; Sternberg et al., 1971), where stimuli that are attended are processed faster than unattended ones. Similar 'Anti-Bayesian' percepts have recently been described with asymmetric likelihood functions (Wei & Stocker, 2012; 2015). As such, when applied to time perception and the anisotropic nature of temporal processing, such asymmetric functions may serve as a basis for the notion of the repulsion of a stimulus away from the expected time.

4.5.4 Conclusion

We have presented a novel Bayesian model for the perception of event-timing based on the iterative combination of prior expectations about the future timing of stimuli with noisy internal representations of current sensory information. We have also introduced a novel experimental paradigm that precisely measures the perceived timing of stimuli in repeating sequences. The model accurately captures the pattern of results reported in this paper, as we find that those stimuli that are presented slightly out of time are regularized so to appear more on time. The model makes further novel predictions that can be tested with future scientific inquiry as well as the possibility of being tied to plausible neurobiological mechanisms for the processing of temporal information.

4.6   Materials and Methods

4.6.1 Subjects

Eighteen undergraduate students were recruited via the University of Birmingham student research participation system. All subjects gave written consent before taking part in the experiment and all reported normal or corrected-to-normal hearing and vision. All subjects were naïve to the purpose of the experiment and were given course credits or were paid 6 pounds. The experiment was given ethical approval by the University of Birmingham Ethics Committee.

4.6.2 Stimuli

The auditory stimuli were tones produced by a speaker positioned 50 cm to the left of the participant (20 ms, with a 5 ms linear ramp). Stimuli were all the same pitch (frequency of 1 kHz) but differed between 'low' and 'high' amplitudes. A low stimulus (A) had an amplitude of 37.55 dBA, whilst a high amplitude stimulus (B) was 75.1 dBA. Visual stimuli were flashes produced by a red 5-mm LED positioned in front of the participant (20 ms with 5 ms linear ramp, 91 Cd/m^2).

4.6.3 Behavioural Task

Subjects sat approximately 50 cm from a custom-made light and sound device in a quiet, well-lit room. Subjects were presented with a sequence of seven auditory stimuli with an ISI of 400 ms except the last stimulus that had a temporal deviation of ±40 ms. The last auditory stimulus could be (1) consistent with the sequence/ low amplitude (ABABABA), (2) consistent with the sequence/ high amplitude (BABABAB), (3) inconsistent with the sequence/ low amplitude (BABABAA) or (4) inconsistent with the sequence / high amplitude (ABABABB). The final stimulus was presented together with a visual stimulus

with SOAs of 0, ±40, ±80, ±120, or ±200 ms (It is important to note that an SOA of 0 is locked to the temporal deviation of the final repeated stimulus). Subjects performed a two-alternative forced choice (2AFC) temporal-order judgment (TOJ; did the light or sound come first?) task on the audiovisual pair appearing at the end of the sequence. All conditions were interleaved and the next trial would begin 1.5 or 2 s after the response key was pressed. Each condition was presented 6 times in order to construct a psychometric function.

4.6.4 Psychometric Function

For each participant a psychometric function was established by relating the cumulative proportion of responses in the TOJ task. In order to see a shift in the perceived timing of stimuli, the point of subjective simultaneity (*PSS*) derived from a psychometric function was of great interest. We define the *PSS* as the *SOA* at which each subject was equally likely to respond that the final repeated stimulus or visual probe stimulus was first. Negative *PSS* values indicate that sound had to be presented before light for them to be perceived synchronous whilst positive values mean that light had to be presented before sound to be perceived as synchronous. Using the Spearman-Kärber method the *PSS* and *JND* were estimated as the first and second moments of the psychometric function differential across *SOAs* (see Miller & Ulrich, 2001 for more detail). This method is a non-parametric estimation that does not make assumptions about the distributions underlying the psychometric functions observed. A mathematical derivation of the method can be expressed as follows. First we define SOA_i with $i=\{1, \dots 15\}$ as the 15 values of SOA used in the experiments and p_i with $\{i=1, \dots 15\}$ as the associated proportion of 'light first' responses. We further define $SOA_0=-250$ ms, $SOA_{16}=+250$ ms, $p_0=0$, and $p_{16}=1$, to be able to compute the intermediate *SOA* between two successive ones:

$$s_i = \frac{SOA_{i+1} + SOA_i}{2}, \text{ with } i = \{0, \dots 15\} \qquad \text{(Eq. 4.5)}$$

and the corresponding values of the difference in proportion, taken at and above 0 to

monotonize the proportion of responses

$$dp_i = max\,(0,\ p_{i+1} - p_i)\,, \quad \text{with } i = \{0, \dots 15\} \qquad \text{(Eq. 4.6)}$$

With these indexes we can express *PSS* analytically as:

$$PSS = \frac{1}{\sum_{i=0}^{15} dp_i} \sum_{i=0}^{15} s_i dp_i \qquad \text{(Eq. 4.7)}$$

And the *JND* as:

$$JND = \sqrt{\sum_{i=0}^{15} dp_i (s_i - PSS)^2} \qquad \text{(Eq. 4.8)}$$

Data was not analyzed if subjects had a *JND* above 250 ms or if *PSS* exceeded $\pm$ 175 ms

when data was collapsed across conditions. The data of 5 participants were not analysed.

Chapter 5 General Discussion and Conclusions

The work presented in this thesis sought to understand the computational mechanisms underpinning the perceived timing of events in sequences. In particular, our goal was to shift the focus of time perception models away from perceived duration to perceived event timing. In this Chapter, I will summarize the main findings of each empirical Chapter, show how they contribute to the literature and explain how they improve our understanding of time perception.

5.1 Summary of Main Findings

5.1.1 Chapter 2: A Bayesian Model of Perceived Timing

The perceived timing of *when* an event occurred has often been overlooked in the literature of temporal perception. Numerous studies have sought to understand how discrimination performance to temporal irregularities increases as the amount of stimuli increases (Drake & Botte, 1993; Halpern & Darwin, 1982; ten Hoopen et al., 2011; Lunney, 1974; McAuley & Kidd, 1998; Miller & McAuley, 2005). In a first experiment, we tested such a prediction by asking participants to report whether the last stimulus (the stimulus could be earlier or later than expected) in a sequence of isochronous tones was 'on time' – or not. In contrast to the predictions made by the literature, we show that the increase in irregularity discrimination as a function of the amount of stimuli in a sequence is asymmetric – stimuli presented earlier than expected are better identified as irregular with increasing sequence length compared to stimuli appearing later than expected.

As a possible explanation for this asymmetry, we hypothesized that changes in the perceived timing of the final stimulus could account for the pattern of results. To measure the perceived timing of the final stimulus, we presented a sequence of isochronous tones where the final tone was paired with a stimulus in another modality. From the participants' responses, we calculated the PSS: the audiovisual asynchrony necessary to perceive both stimuli as simultaneous. We found that if the final stimulus was presented a little earlier than expected – then the perceived timing is changed in a way that delays the stimulus towards its expected timing. Conversely, we found that stimuli presented a little later than expected are perceptually accelerated towards expectation. The effect of stimuli being delayed towards the time they are expected can be understood as a *temporal regularization* – which is similar to central tendency effects in the time perception literature, such as Vierordt's Law (Lejeune & Wearden, 2009; Vierordt, 1868), where the duration of an interval is biased by the average duration of intervals previously experienced (Jazayeri & Shadlen, 2010; Petzschner et al., 2015). However, in opposition to a central tendency effect, we found asymmetries also in the perceived timing data of stimuli presented at their expected time (on time), as they are perceptually accelerated *away* from expectation.

We present a novel Bayesian model based on the dynamic updating of temporal expectations that can explain the asymmetries in the detection of irregularity *and* in the perceived timing of stimuli. In our model, perceived timing (the posterior distribution) is the result of combining the time you sense a stimulus (likelihood) with the time you expect to sense it (prior) – at each point in time. As opposed to current Bayesian accounts of time perception that use Gaussian probability distributions (Hartcher-O'Brien et al., 2014; Jazayeri & Shadlen, 2010; Miyazaki et al., 2006; Shi et al., 2013), the key tenet of our model is the relaxation of the assumption of normality in the probability distributions. We propose that probability distributions in the temporal domain are necessarily asymmetric

due to the way time flows. The anisotropic nature of time means that evidence accumulated about stimulus timing for the likelihood function can only start after a short delay due to neural processing – but although a stimulus cannot be sensed before a stimulus is presented – however there is always the chance it could be perceived a bit later than on average due to noise in the sensory system. Prior distributions about the expected timing of future events should also be asymmetric, as an organism cannot predict a second event to occur *before* the first event, and as such should start at 0 for when the first event occurs and the distribution continues to rise until the expected timing of a second event. However, due to the anisotropy of time, the second event could still be expected tomorrow, and as such the prior should have a long off tail.

Our Bayesian model of perceived timing makes intrinsic predictions that gave rise to experiments of Chapters 3 and 4. The first step was to reverse the idea of expectation and irregularity to hypothesize that stimuli that are not expected in a sequence i.e. are presented in a random sequence of irregular timings, should not have any temporal expectations built up. As such, we predicted that the perceived timing of stimuli in an environment where trials are isochronous should exhibit the temporal regularization effect we presented in Chapter 2 – early stimuli should be delayed towards expectation whilst late stimuli should be accelerated. However, stimuli that are presented in an irregular environment should not have any modulation of their perceived timing.

5.1.2 *Chapter 3: Temporal Regularity of the Environment Drives Time Perception*

The perception of regularity has historically been investigated in terms of deviations from its inverse: irregularity (Drake & Botte, 1993; Halpern & Darwin, 1982; Lunney, 1974; McAuley & Kidd, 1998; Repp, 1999; Schulze, 1978; 1989; Tanaka, Tsuzaki, Aiba, & Kato, 2008). But what makes a sequence of isochronous tones be perceived as regular?

Extant models of rhythm perception assume that if a stimulus is presented in an isochronous structure then it is simply perceived as such. Time, however, is a physical dimension that is often subject to distortion in human perception (Allman & Meck, 2012; Hellström & Rammsayer, 2015; ten Hoopen et al., 1995; Horr & Di Luca, 2015a; 2015b; Jazayeri & Shadlen, 2010; Lejeune & Wearden, 2009; Petzschner et al., 2015; van Wassenhove, Buonomano, Shimojo, & Shams, 2008; Wearden et al., 2007); so why should a temporal property such as regularity be taken for granted?

In this Chapter we aimed to test whether the temporal environment could influence the perception of regularity. Secondly, we tested a prediction of our Bayesian model of perceived timing – if a sequence has temporal irregular events, then the perceived timing of a stimulus should not be modulated – as the prior that biased perceived timing cannot be built. We find that a regularly-timed environment promotes the perception of regularity and changes the perceived timing of stimuli to make slightly irregular stimuli appear more regular. An irregular environment of jittered tones, on the other hand, makes perfectly regular tones embedded within it be perceived as slightly irregular.

We interpret these results within the context of our Bayesian model. We suggest that in a regular environment, temporal expectations dynamically build after each stimulus and subsequently bias the perception of slightly irregular stimuli to make them appear more regular. However, in an irregular environment, temporal expectations are less precise and as such do not build up, and therefore do not bias the perceived timing of stimuli. As the representations are less precise, the posterior distribution from which the perception of regularity is taken is wider, and as such there is a chance that an isochronous stimulus is perceived as being irregular. It is important to note that the model presented in this and the following Chapter (4) uses *symmetrical* Gaussian distributions. In this instance, we chose

such symmetrical distributions to reduce model complexity – but also to give a more parsimonious account of the findings in this Chapter.

5.1.3 Chapter 4: Bayesian Time Perception: Prior Expectations and Noisy Sensory Estimates Bias Perceived Timing

In Chapter 4, we tested an implicit assumption of the Bayesian model of perceived timing: noisy measurements should lead to broader likelihood functions that are captured more by the prior probability distributions. In order to test this, we presented a sequence of stimuli with alternating amplitudes where the final stimulus could be earlier or later than expected. By reporting whether the final stimulus was before or after a visual probe, we were able to calculate the perceived timing of stimuli. Concurrent with the data from Chapters 2 and 3, we find a temporal regularization effect – stimuli presented earlier than expected are perceptually delayed whilst late stimuli are perceptually accelerated. Importantly, addressing the motivation of this experiment, we find that the temporal regularization effect is strongest for stimuli when the final stimulus was of weak amplitude. The data promotes the idea that temporal expectations are dynamically updated at every point in time and used to bias perception in a way that regularizes perception.

5.2   Theoretical Implications and Contributions

Throughout this thesis the main aim was to add a new focus to time perception research that aims at understanding how the brain may perceive the timing of events, rather than duration. We wanted to understand whether and how the Bayesian framework could be used to characterize how temporal expectations could be built rapidly and used to modify perceived timing. In this section, I will discuss how the results and models from Chapters 2-4 impact the current theoretical accounts for the perception of time. Further, I will then

attempt to synthesize the findings and models into a broader theoretical framework for time perception.

5.2.1 Impact to Contemporary Models of Time Perception

The Bayesian model with asymmetric likelihood functions accurately captures the experimental data presented in the Experiments of Chapter 2. Previous timing models, for example interval-based and entrainment models of time perception, cannot account for the asymmetric patterns of results we observe. We show in Chapter 2 (Fig 2.1cd), that there is an asymmetry in temporal deviation detection: stimuli that are presented earlier than expected are better detected as off-time as the length of a sequence increases. Both interval and entrainment models predict a symmetric increase in temporal discrimination performance as the amount of stimuli in a sequence increases (Drake & Botte, 1993; ten Hoopen et al., 2011; Large & Jones, 1999; Large & Palmer, 2002). The Multiple-Look Model (MLM), an interval-based model of temporal discrimination, is based on the idea that as sequence length increases so does the precision of an estimate for each interval (Drake & Botte, 1993; Miller & McAuley, 2005). Similarly, the beat-averaging (Schulze, 1978; 1989), diminishing returns (ten Hoopen et al., 2011) and internal-reference model (Bausenhart, Dyjas, & Ulrich, 2014; Dyjas, Bausenhart, & Ulrich, 2012; Ulrich, 1987), are all based on similar premises. As the factor of change in such accounts is the better internal representation of an interval, interval-based models make no predictions about changes in the perceived timing of stimuli (Gibbon, 1977; Gibbon et al., 1984; Shi et al., 2013) – as stimuli simply delimit intervals[1].

A key interval-based model to explain such changes in representation mechanistically is SET (Gibbon et al., 1984). In this model, an internal pacemaker emits

---

[1] Our Bayesian model offers an alternative explanation, as temporal discrimination increases with sequence length through the increasing steepness of the posterior distribution resulting in greater temporal acuity (Li, Rhodes & Di Luca, In review; Rhodes, Li & Di Luca, In submission).

pulses that are accumulated and counted between two events – leading to a duration estimate. In order to account for the modulations in perceived timing that we report in Chapter 2-4, the SET model must be augmented to account for the changes in perceived timing we observe. Rather than being in competition with SET, the model we present represents a general issue in resolving how 'global' context effects can be reconciled with 'local' changes in perception – as it has been shown that the duration of just the previous stimulus can affect the perceived simultaneity of the next (Van der Burg et al., 2013); as well as the temporal regularization phenomena we report in this thesis. As such, a general model of time perception that both estimates perceived timing and duration is of paramount importance in order to reconcile such different ways of understanding how we perceive time.

Entrainment models of temporal perception similarly predict symmetrical performance in determining if earlier or later-than expected are irregular with respect to expectation (Henry & Herrmann, 2014; Large & Jones, 1999; Large & Palmer, 2002). Entrainment models are based on the idea that the phase and frequency of temporal patterns adjust to rhythmic events– where at the neural level, recurrent activity patterns (Buonomano, 2009; Buonomano & Merzenich, 1995; Karmarkar & Buonomano, 2007; Laje & Buonomano, 2013) or phase coincidence (C. Miall, 1989) progressively tune to the frequency and phase of external stimulation. Though not originally formulated to predict changes in perceived timing, entrainment models could be formulated to appeal to the rhythmic deployment of attention at an expected time-point to facilitate the processing of on-time stimuli to be perceived faster (Rohenkohl et al., 2011). However, we find that early stimuli are delayed towards expectation (Figure 2.2) and as such, current formulations of entrainment models cannot account for this finding (Buonomano & Merzenich, 1995; Karmarkar & Buonomano, 2007; Large & Jones, 1999; Large & Palmer,

2002; Large & Snyder, 2009; C. Miall, 1989) – as principally these models are based on phase correction for the next stimulus in a sequence, and not modifications of a stimulus at the present time, whilst it is also unclear how such models could account for perceptual delay. Similar to the implication for interval-models, entrainment accounts of temporal processing should consider the modulation of PSS that results in temporal regularization.

To summarize, the Bayesian model of perceived timing can explain the delay of early stimuli as well as the acceleration of on time and later than expected stimuli. Interval models do not make any predictions about changes in the perceived timing of stimuli and as such cannot account for our data, whilst entrainment accounts could be formulated to explain the acceleration of on time stimuli – however they cannot explain the delay towards expectation of early stimuli.

5.2.2 Impact to Sensory Processing Theories

Sensory processing involves three separate stages – (1) detecting incoming information, (2) representing incoming information and (3) interpreting that representation. Two distinct accounts exist to explain these processes: the efficient coding hypothesis explains how limited neural resources lead to efficient representations that are optimized with regard to the natural statistics in the environment (Barlow, 1961; Lewicki, 2002; Simoncelli, 2003; Wei & Stocker, 2015). The role of primary sensory processing is, as such, to reduce the inefficiency and redundancy in representing a raw image by recoding a representation into an efficient form (Huang & Rao, 2011). However, in this hypothesis, it is hard to determine how perceptual biases may arise. Built on such a theoretical bases, the predictive coding hypothesis suggests sensory processing is the result of combining current sensory information with prior knowledge about the world (Friston & Kiebel, 2009; Helmholtz, 1963; Kersten, Mamassian, & Yuille, 2004; Knill & Richards, 1996; Ma et al., 2006; Srinivasan, Laughlin, & Dubs, 1982) – according to Bayes' (1763) rule. Such

an information-processing approach can explain the myriad of data that shows consistent

perceptual biases (Ernst, 2006; Ernst & Banks, 2002; Knill & Richards, 1996; Körding &

Wolpert, 2004; Mamassian et al., 2002; Petzschner et al., 2015; Wolpert & Ghahramani,

2000). Recently, however, a unified model has been proposed that reconciles a predictive

coding (Bayesian) approach with efficient coding of a sensory representation (Wei &

Stocker, 2012; 2015), by constraining priors and likelihoods with natural stimulus

statistics.

In this thesis, however, we too show how sensory information may be represented

at the neural level – by constraining the likelihood function with the anisotropy of time.

We introduce the idea that the likelihood function is necessarily asymmetric in the

temporal dimension, with a steep onset and long-off tail (Chapter 2; Section 2.3.4). The

asymmetric likelihood function explains how stimuli that are presented on time are

perceptually accelerated – an anti-Bayesian effect. Interestingly, a recent article has shown

concurrent repulsions away from the peak of the prior through similarly asymmetric

likelihoods and priors (Wei & Stocker, 2012; 2015). The relaxation of the assumption of

normality is thus of theoretical importance as up until now, probability distributions have

been ubiquitously described as Gaussians in the Bayesian framework (Ernst, 2006; Ernst

& Banks, 2002; Jazayeri & Shadlen, 2010; Knill & Richards, 1996; Miyazaki et al., 2005;

Sciutti et al., 2014; Shi et al., 2013).

In Chapter 3, we report that the temporal context of an environment balances the

tendency of the regularisation effect. We find in concert with the temporal regularisation

effects found in Chapter 2 (Figure 2.2 and 4 (Figure 3.2) – however, we find no such

temporal regularization effect in perfectly regularly-timed sequences when the sequence is

embedded within an environment of mostly irregularly-timed sequences. The results of

Chapter 3 thus hint at the brain optimizing perception in order to process sensory

information more efficiently. Why regularize stimuli if most are actually irregular? Similarly, the exploitation of temporal regularities decreases neural metabolic consumption (VanRullen & Dubois, 2011). The predictable timing of future stimuli leads to improved stimulus discrimination and detection in a plethora of tasks (Brochard et al., 2013; Carnevale et al., 2015; Correa et al., 2005; Cravo et al., 2013; Escoffier et al., 2010; Jazayeri & Shadlen, 2010; Rohenkohl & Nobre, 2011), whilst the rhythmic entrainment of stimuli allows the automatizing of behaviour for activities such as dance, locomotion, speech, and music production (McNeill, 1995; Repp, 2005).

In Chapter 4 we further characterized how sensory processing may be affected by noisier measurements. We hypothesized that noisier signals should lead to shallower likelihood functions – and as such, be more captured by the prior compared to less noisy functions. This sort of effect has been found in the context of human speed perception, whereby a broader likelihood function results in speed estimates that are more dominated by the prior (Stocker & Simoncelli, 2006). As we have translated this type of effect into the domain of temporal perception – one could posit that this is applicable to other perceptual modalities and is, as such – perception-general.

5.3 Critique

From the three empirical Chapters presented in this thesis, I have strongly advocated the idea that the brain may be regularizing perceptual events. There is, however, still much cause for considering alternative interpretations of the data, as well as pointing out some facets of the model that may help increase clarity. In this next section, I will highlight such issues before moving onto discussing future directions of our work.

5.3.1 Psychophysical Interpretation

The logic behind the Bayesian model of perceived timing with asymmetric distributions is due to the asymmetry of distortions in perceived timing evidenced by Experiment 2 in Chapter 2 (Figure 2.2). The possibility remains, however, that differential attentional processing to one modality over another may cloud such an asymmetry. The phenomenon of prior entry, for example, dictates that attended stimuli are processed quicker than unattended ones (Spence & Parise, 2010; Sternberg et al., 1971). If we consider this with regards to our experimental paradigms, one may posit that auditory and visual trials are differentially attended to, and thus have their own associated prior entry effects. It may be argued, therefore, that prior entry may cause a baseline shift in the data, such that what appears to be a crossover point of 40ms (Figure 2.2) may actually be at 0 – as prior entry shifts stimuli earlier in time by 20-40ms. This sort of interpretation may render the asymmetric model as needless; however, there is evidence that advocates an asymmetric Bayesian model. For instance, when we asked participants' to perform simultaneity judgments (SJs) instead of temporal order judgements (Supplementary Figure 2.2), we find a crossover point that is positive. That is, the lines between auditory and visual sequences converge at around +40/80ms: in the opposite direction of the convergence point of Experiment 2 with temporal order judgments. The difference between audio and visual sequences with SJs thus evidences a generalised *delay*, in contrast to an effect such as prior entry that promotes the idea of perceptual *acceleration*.

We quantify the effect of delay and acceleration (for early and late stimuli respectively) by considering the difference in PSS estimates between sound and light sequences (Figure 2.2). However, in later Chapters (also Chapter 2, Experiments 3 and 4), we infer the effects of delay and acceleration by changes in PSS estimates across anisochronies in unimodal sequences (Figures 3.2 & 4.4). Given this inconsistency in

quantifying the effect, we highlight that the same pattern of results persists in Experiment 2 (Chapter 2). If we consider the auditory (or visual) sequences alone in Figure 2.2, then we see the same changes in PSS as a function of the final stimulus's anisochrony. Further, Experiment 2 was run with crossmodal sequences, and as such, has potentially higher noise in the measurements. Thus, comparing changes in perceived timing as a function of anisochrony within a sound (or light) sequence induces more risk of masking the BET effect. As such, we used a 'crossmodal' interpretation of these data to arrive at the conclusions we have expounded throughout this thesis.

One further issue to highlight, is that the TOJ task used in Experiment 2 is prone to decision-level biases based on the type of judgment being asked (Yarrow et al., 2011). Given that these response biases exist in the literature, it may be possible that participants differentially respond 'sound' or 'light first' to give rise to the pattern of results that we report. As our paradigm has not been used in the literature, further research should focus on how the type of task used to index changes in perceived timing changes the patterns of results found here.

5.3.2 Model Interpretation

The Bayesian model of perceived timing that we present throughout this thesis has an interesting inversion of the standard Bayesian framework (Kersten & Yuille, 2003; Knill & Richards, 1996). In the standard Bayesian inference model, the prior is modelled as the knowledge of the state of the world. The likelihood reflects the probability of sensing a stimulus given an external signal. Finally, the posterior probability reflects the probability of inferring the state of the world given sensory information. In our instantiation of a Bayesian model for time, however, the posterior and likelihood both reflect the probability of sensing a stimulus given the time a stimulus was presented. This interesting reversal of the posterior is caused by the inversion of the generative process used to give rise to the

likelihood functions used in our model (Chapter 2.3.4). Whilst the reversal of the posterior is in competition with the standard Bayesian framework, we point to the anisotropic nature of time as reason to why such a reversal may be reasonable. As time flows in one direction, one cannot sense the timing of an event before a stimulus is presented. Given this logic, we modelled the likelihood in such a way that allowed us generate a probability distribution that was in response to an external event.

A further issue to comment on is the framing of the model as changes in the low level processing of stimuli leading to perceptual delay or acceleration. The evidence for this sort of interpretation lies in oft-reported prior entry effects found in the literature (Spence & Parise, 2010; Sternberg et al., 1971), as well as in empirical work relating to the DAT theory of attentional entrainment. Both prior entry and DAT advocate that stimuli are processed faster if attended. However, our model may be viewed from the perspective of inference – rather than in changes in encoding. For example, it is entirely possible that a post-dictive estimation of the timing of an event happens sometime after the stimulus has been presented. Bayesian decision theory is inherently retrospective, thus in order to clarify whether the model we present is 'on-line' or post-dictive, future research could employ techniques to identify neurocorrelates of the model, and see if, for example, peaks of sensory registration change as a function of the anisochrony of the final stimulus of the sequence.

An alternative explanation of the BET effect that we report in Chapters 2-4 is to consider that subjects may learn to strategically allocate attention across time. In this way, as soon as the penultimate stimulus in a sequence offsets, a hazard function rises until the offset of the final stimulus in a sequence (Nobre, Correa, & Coull, 2007). As one can imagine, the effect of the hazard function would be smaller for early stimuli and stronger for later stimuli. If we further consider that if a prior entry style effects acts in concert with

a hazard function, then such an account may yield results similar to what we report in this thesis. That is, there is greater perceptual acceleration for stimuli the later they appear in time. However, if we consider this sort of interpretation further, there are a few issues that may be allayed by the BDT account. First, the hazard function has only been related to changes in motor preparation indexed by response times in humans (Nobre, Correa, & Coull, 2007) and monkeys (Janssen & Shadlen, 2005), and not changes in perception, as we present in our model. Second, we relate the prior as a mechanism to allocate attention at an expected time-point. In the hazard function approach, it is unclear how attention may be allocated to an expected time-point. Further experiments should elucidate if the hazard function could be used to describe changes in the perception of stimuli, rather than changes in motor preparation, as well as if the hazard function could be modified by temporal expectations.

5.4   Directions for Future Research

In order to continue to validate the proposed Bayesian model of perceived timing, the model must be tested and subsequently modified in order to reflect the findings of future work. In this section, I will make explicit predictions based on this model and from those formulate suggestions for future research.

5.4.1 Predictions

In this thesis, the main finding is that the perceived timing of stimuli is biased by dynamic prior expectations based on the iterative presentation of isochronous stimuli. As such, the characterization and limits of this finding require further investigation. I will now present some examples of how the model may be further explored.

To elicit the temporal regularization effect detailed in Chapters 2-4, I have presented sequences of isochronous events in order to build up prior expectations, yet in

the environment, sequences of repeated events are often *not* isochronous. In almost all forms of music around the world, there are rarely any instances of completely isochronous melodies – music has distinct and complex temporal patterns operating at different hierarchies and time signatures (Vuust & Witek, 2014). Syncopated rhythms, for example, carry expectations about the future timing of events – yet are not completely isochronous (Fitch & Rosenfeld, 2007). How can the brain predict such events in the context of the model we present – if it is based on the isochronous presentation of stimuli? The model, at present, would predict that a syncopated (as such), deviant stimulus would be biased towards the expected timing – yet it seems that when a stimulus is obviously earlier than expected – then we perceive it as such. To clarify this issue – we must first show the extent of the regularization effect over a whole range of anisochronies. One may predict that at a certain magnitude of anisochrony – the regularisation effect goes away. If this is the case, it may mean that a hierarchical prior takes over and modulates the tendency to regularize deviant stimuli. Further, one could also imagine another prior that is based on the rhythm and syncopation of a sequence, which also influences the lower-level regularisation and as such, the combination of the prior and likelihood.

Given that the prior is built after the presentation of isochronous events, sometimes events may not be sensed or not even occur. In the active sensing framework, entrained oscillations continue to be in phase consistency after the end of the external stimulation – yet decay after some time (Lakatos et al., 2008; 2005; Schroeder & Lakatos, 2009). In the same way, does the prior decay after time or does it stop influencing the moment a beat is missed? To test this, one could think of an experiment where the final stimulus is missed and presented at *T*+1, *T*+2, *T*+3 etc. where *T* is the *timing* the final stimulus. If the prior is still present (yet decayed) it should still modulate perceived timing – but the effect should diminish as the missed beats increases.

Moving away from the perception of audio or visual stimuli – the model could be extended to the realm of motor control. It has been consistently shown that humans synchronize to sensorimotor events such as finger tapping or dancing (Elliott et al., 2014; Elliott, Welchman, & Wing, 2009; Elliott, Wing, & Welchman, 2010; Repp, 1999; 2005; Repp & Su, 2013). A consistent finding in such studies is that the time of a tap (i.e. the time at which a finger touches a surface) is prior to the onset of an isochronous metronome. The model could account for such a negative error as it predicts that the perception of isochronous events is actually perceived earlier than expectation resulting in earlier taps. Further, how should an observer know when to initiate a tap? Due to the build of temporal expectations via the stimulation of a metronome observers can anticipate the timing of future taps and use this information to initiate a movement.

5.4.2 A Unified Model of Time Perception?

What should a unified model of time perception look like? A great deal of literature has been dedicated to the perception of time – and in particular, interval timing (Creelman, 1962; Gibbon et al., 1984; Matell & Meck, 2004; Meck, 2005; Merchant & de Lafuente, 2014; Treisman, 1963). The perception of duration has been described mechanistically with the SET model – though, recently it has been tied to thalamo-cortico-striatal circuitry (Matell & Meck, 2004). Contextual calibration effects on perceived duration have been modelled in the Bayesian framework – whereby duration estimates are biased towards the mean of previously experienced intervals (Jazayeri & Shadlen, 2010; Miyazaki et al., 2006; Shi et al., 2013). Context effects are bound by the fact they take a long course to learn the temporal statistics of the environment (Acerbi et al., 2012). The motivation for this thesis, however, was in re-focusing temporal perception from the duration dimension to perceived timing – as well as showing how the perception of time can be biased rapidly within just a single trial. Therefore, future work should seek to link together the existing

frameworks for perceived duration and perceived timing. As both perceived timing (Chapter 2-4) and contextual calibration of perceived duration (Jazayeri & Shadlen, 2010; Miyazaki et al., 2006; Shi et al., 2013) have been described in the Bayesian framework, a neural model of Bayesian inference to explain both perceived duration and timing could lead to a unified and neurophysiologically plausible account of time perception.

There are several theories of how the brain may represent probability distributions (Beck et al., 2008; Deneve, Latham, & Pouget, 1999; Fiser, Berkes, Orbán, & Lengyel, 2010; Hoyer & Hyvarinen, 2003; Pouget, Dayan, & Zemel, 2000; Zemel, Dayan, & Pouget, 1998). Whilst ultimately a computational framework to explain how prior expectations can be combined with current sensory evidence to arrive at a best estimate to the state of the world, Bayesian inference has been shown to operate at the neural level through probabilistic population coding (Ma et al., 2006). A whole raft of psychophysical experiments shows that humans perform to near Bayes-optimal inference (Beierholm, Quartz, & Shams, 2009; Ernst, 2006; Ernst & Banks, 2002; Kersten & Yuille, 2003; Knill & Richards, 1996; Körding & Wolpert, 2004; Ma et al., 2006; Petzschner & Glasauer, 2011; Shi et al., 2013; Stocker & Simoncelli, 2006; Vilares & Körding, 2011), but in this thesis we describe for the first time how subjects use Bayesian inference in the domain of perceived timing.

In order to translate our model to the neural level, we must first consider that our model is not in competition with interval-based accounts of time perception (Creelman, 1962; Gibbon et al., 1984; Jazayeri & Shadlen, 2010; Petzschner et al., 2015; Shi et al., 2013; Treisman, 1963) – but rather, our model should be synthesized with such models in order to arrive at a *general* model of time perception. A Bayesian neural inference model that is hierarchically organized such that at a low level population codes encode the perceived timing of stimuli but then feed-forward to a higher level that encodes the

duration between two stimuli may offer a way of harmonizing perceived duration and perceived timing.

5.5 Conclusions

During the last 150 years, great steps have been made in understanding how the human brain may perceive time. The advent of the psychophysical approach to studying perception has allowed researchers to precisely measure temporal properties of stimuli and as such, a large body of research has sought to understand the mechanisms underpinning temporal-perceptual phenomena. Contemporary models of time perception consider temporal processing from the perspective of duration. In this thesis, I have tried to add a new focus to time perception research away from the duration dimension towards an event-based approach. A Bayesian model of perceived timing re-focuses temporal perception research towards an event-based outlook. The model sets the scene to unify temporal processing accounts at neural, computational and behavioural levels, with the future goal of leading to a general model of time perception that is neurobiologically plausible and grounded in computational principles.

References

Acerbi, L., Vijayakumar, S., & Wolpert, D. M. (2014). On the Origins of Suboptimality in Human Probabilistic Inference. *PLoS Computational Biology*, *10*(6), e1003661. doi:10.1371/journal.pcbi.1003661.s004

Acerbi, L., Wolpert, D. M., & Vijayakumar, S. (2012). Internal Representations of Temporal Statistics and Feedback Calibrate Motor-Sensory Interval Timing. *PLoS Computational Biology*, *8*(11), e1002771. doi:10.1371/journal.pcbi.1002771.s003

Adams, R. D. (1977). Intervening stimulus effects on category judgments of duration. *Perception & Psychophysics*, *21*(6), 527–534.

Allan, L. G. (1979). The perception of time. *Perception & Psychophysics*, *26*(5), 340–354.

Allison, T., Matsumiya, Y., Goff, G. D., & Goff, W. R. (1977). The scalp topography of human visual evoked potentials. *Electroencephalography and Clinical Neurophysiology*, *42*(2), 185–197.

Allman, M. J., & Meck, W. H. (2012). Pathophysiological distortions in time perception and timed performance. *Brain*, *135*(3), 656–677. doi:10.1093/brain/awr210

Allman, M. J., Teki, S., Griffiths, T. D., & Meck, W. H. (2013). Properties of the Internal Clock: First- and Second-Order Principles of Subjective Time. *Annual Review of Psychology*, *65*(1), 130919205320001. doi:10.1146/annurev-psych-010213-115117

Arao, H., Suetomi, D., & Nakajima, Y. (2000). Does time-shrinking take place in visual temporal patterns? *Perception*, *29*(7), 819–830. doi:10.1068/p2853

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, 1–9. doi:10.1016/j.tics.2012.05.003

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2014). Delta-Beta Coupled Oscillations Underlie Temporal Prediction Accuracy. *Cerebral Cortex*. doi:10.1093/cercor/bhu103

Auksztulewicz, R., & Friston, K. (2015). Attentional Enhancement of Auditory Mismatch Responses: a DCM/MEG Study. *Cerebral Cortex*. doi:10.1093/cercor/bhu323

Baddeley, A. (1992). Working memory: The interface between memory and cognition. *Journal of Cognitive Neuroscience*, *4*(3), 281–288.

Bald, L., Berrien, F. K., Price, J. B., & Sprague, R. O. (1942). Errors in perceiving the temporal order of auditory and visual stimuli. *Journal of Applied Psychology*, *26*(3), 382–388.

Barlow, H. B. (1961). The coding of sensory messages. In W. H. Thorpe & O. L. Zangwill, *Current problems in animal behaviour* (pp. 330–360). Cambridge: Cambridge University Press.

Barnes, R., & Jones, M. R. (2000). Expectancy, Attention, and Time. *Cognitive Psychology*, *41*(3), 254–311. doi:10.1006/cogp.2000.0738

Bateson, M., & Kacelnik, A. (1997). Starlings' preferences for predictable and unpredictable delays to food. *Animal Behaviour*, *53*(6), 1129–1142. doi:10.1006/anbe.1996.0388

Bausenhart, K. M., Dyjas, O., & Ulrich, R. (2014). Acta Psychologica. *Actpsy*, *147*(C), 60–67. doi:10.1016/j.actpsy.2013.06.011

Bayes, T. (1763). An Essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, *53*(0), 370–418. doi:10.1098/rstl.1763.0053

Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., et al. (2008). Probabilistic Population Codes for Bayesian Decision Making. *Neuron*, *60*(6), 1142–1152. doi:10.1016/j.neuron.2008.09.021

Beierholm, U. R., Quartz, S. R., & Shams, L. (2009). Bayesian priors are encoded independently from likelihoods in human multisensory perception. *Journal of Vision, 9*(5), 23–23. doi:10.1167/9.5.23

Bindra, D., & Waksberg, H. (1956). Methods and terminology in studies of time estimation. *Psychological Bulletin, 53*(2), 155–159.

Block, R. A., & Gruber, R. P. (2014). Time perception, attention, and memory: A selective review. *Acta Psychologica, 149*(C), 129–133. doi:10.1016/j.actpsy.2013.11.003

Block, R. A., & Zakay, D. (1997). Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin & Review, 4*(2), 184–197. doi:10.3758/BF03209393

Blommaert, F. J., & Roufs, J. A. (1987). Prediction of thresholds and latency on the basis of experimentally determined impulse responses. *Biological Cybernetics, 56*(5-6), 329–344.

Boenke, L. T., Deliano, M., & Ohl, F. W. (2009). Stimulus duration influences perceived simultaneity in audiovisual temporal-order judgment. *Experimental Brain Research, 198*(2-3), 12. doi:10.1007/s00221-009-1917-z

Bortoletto, M., Cook, A., & Cunnington, R. (2011). Motor Timing and the Preparation for Sequential Actions
. *Brain and Cognition, 75*(2), 196–204. doi:10.1016/j.bandc.2010.11.016

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General, 138*(4), 487–502. doi:10.1037/a0016797

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

Brannon, E. M., Roussel, L. W., Meck, W. H., & Woldorff, M. (2004). Timing in the baby brain. *Cognitive Brain Research, 21*(2), 227–233. doi:10.1016/j.cogbrainres.2004.04.007

Breska, A., & Deouell, L. Y. (2014). Automatic Bias of Temporal Expectations following Temporally Regular Input Independently of High-level Temporal Expectation. *Journal of Cognitive Neuroscience, 26*(7), 1555–1571. doi:10.1037/0096-1523.16.1.121

Brochard, R., Tassin, M., & Zagar, D. (2013). Got rhythm... for better and for worse. Cross-modal effects of auditory rhythm on visual word recognition. *Cognition, 127*(2), 214–219. doi:10.1016/j.cognition.2013.01.007

Brody, C. D., Hernández, A., Zainos, A., & Romo, R. (2003). Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. *Cerebral Cortex, 13*(11), 1196–1207. doi:10.1093/cercor/bhg100

Brown, S. W., & Stubbs, D. A. (1988). The psychophysics of retrospective and prospective timing. *Perception, 17*(3), 297–310.

Brown, S. W., & Stubbs, D. A. (1992). Attention and interference in prospective and retrospective timing. *Perception, 21*(4), 545–557.

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

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews: Neuroscience, 6*(10), 755–765. doi:10.1038/nrn1764

Buhusi, C. V., Sasaki, A., & Meck, W. H. (2002). Temporal integration as a function of signal and gap intensity in rats (Rattus norvegicus) and pigeons (Columba livia). *Journal of Comparative Psychology, 116*(4), 381–390. doi:10.1037//0735-7036.116.4.381

Buonomano, D. V. (2009). Harnessing Chaos in Recurrent Neural Networks. *Neuron,*

63(4), 423–425. doi:10.1016/j.neuron.2009.08.003

Buonomano, D. V., & Merzenich, M. M. (1995). Temporal information transformed into a spatial code by a neural network with realistic properties. *Science*, *267*(5200), 1028–1030.

Callender, C. (2010). Is time an illusion? *Scientific American*, *302*(6), 58–65.

Carnevale, F., de Lafuente, V., Romo, R., Barak, O., & Parga, N. (2015). Dynamic Control of Response Criterion in Premotor Cortex during Perceptual Detection under Temporal Uncertainty. *Neuron*, *86*(4), 1067–1077. doi:10.1016/j.neuron.2015.04.014

Cattell, J. M. (1886). The influence of the intensity of the stimulus on the length of the reaction time. *Brain*.

Chocholle, R. (1940). Variation des temps de réaction auditifs en fonction de l'intensité à diverses fréquences. *L'année Psychologique*, *41*(1), 65–124. doi:10.3406/psy.1940.5877

Chopin, A., & Mamassian, P. (2012). Predictive Properties of Visual Adaptation. *Current Biology*, *22*(7), 622–626. doi:10.1016/j.cub.2012.02.021

Church, R. M., & Broadbent, H. A. (1990). Alternative representations of time, number, and rate. *Cognition*, *37*(1-2), 55–81.

Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology: Animal Behavior Processes*, *3*(3), 216–228.

Church, R. M., Meck, W. H., & Gibbon, J. (1994). Application of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes*, *20*(2), 135–155.

Colas, F., Diard, J., & Bessiere, P. (2010). Common Bayesian Models for Common Cognitive Issues. *Acta Biotheoretica*, *58*(2-3), 191–216. doi:10.1007/s10441-010-9101-1

Correa, A. A., Lupiáñez, J. J., & Tudela, P. P. (2005). Attentional preparation based on temporal expectancy modulates processing at the perceptual level. *Psychonomic Bulletin & Review*, *12*(2), 328–334.

Cravo, A. M., Rohenkohl, G., Wyart, V., & Nobre, A. C. (2013). Temporal Expectation Enhances Contrast Sensitivity by Phase Entrainment of Low-Frequency Oscillations in Visual Cortex. *Journal of Neuroscience*, *33*(9), 4002–4010. doi:10.1523/JNEUROSCI.4675-12.2013

Creelman, C. D. (1962). Human discrimination of auditory duration. *The Journal of the Acoustical Society of America*, *34*, 582.

Czeisler, C. A., Duffy, J. F., Shanahan, T. L., Brown, E. N., Mitchell, J. F., Rimmer, D. W., et al. (1999). Stability, Precision, and Near-24-Hour Period of the Human Circadian Pacemaker. *Science*, *284*(5423), 2177–2181. doi:10.2307/2898429?ref=no-x-route:14113f4d0d7c1990d96b2be3c82bd94b

Davies, P. (2002). That mysterious flow. *Scientific American*, *287*(3), 40–3– 46–7.

Deneve, S., Latham, P. E., & Pouget, A. (1999). Reading population codes: a neural implementation of ideal observers. *Nature Neuroscience*, *2*(8), 740–745. doi:10.1038/11205

Di Luca, M., Machulla, T. K., & Ernst, M. O. (2009). Recalibration of multisensory simultaneity: Cross-modal transfer coincides with a change in perceptual latency. *Journal of Vision*, *9*(12), 1–16. doi:10.1167/9.12.7

Dinnerstein, A. J., & Zlotogura, P. (1968). Intermodal perception of temporal order and motor skills: effects of age. *Perceptual and Motor Skills*, *26*(3), 987–1000.

Doherty, J. R., Rao, A., Mesulam, M. M., & Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *Journal of Neuroscience*, *25*(36), 8259–8266. doi:10.1523/JNEUROSCI.1821-05.2005

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

Drew, M. R., Zupan, B., Cooke, A., Couvillon, P. A., & Balsam, P. D. (2005). Temporal Control of Conditioned Responding in Goldfish. *Journal of Experimental Psychology: Animal Behavior Processes*, *31*(1), 31–39. doi:10.1037/0097-7403.31.1.31

Dyjas, O., & Ulrich, R. (2013). Effects of stimulus order on discrimination processes in comparative and equality judgements: Data and models. *The Quarterly Journal of Experimental Psychology*.

Dyjas, O., & Ulrich, R. (2014). Effects of stimulus order on discrimination processes in comparative and equality judgements: Data and models. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1121–1150. doi:10.1080/17470218.2013.847968

Dyjas, O., Bausenhart, K. M., & Ulrich, R. (2012). Trial-by-trial updating of an internal reference in discrimination tasks: evidence from effects of stimulus order and trial sequence. *Attention, Perception & Psychophysics*, *74*(8), 1819–1841. doi:10.3758/s13414-012-0362-4

Edwards, C. J., Alder, T. B., & Rose, G. J. (2002). Auditory midbrain neurons that count. *Nature Neuroscience*, *5*(10), 934–936. doi:10.1038/nn916

Ehrenstein, W. H., & Ehrenstein, A. (1999). Psychophysical Methods. In U. Windhorst & J. Håkan, *Modern Techniques in Neuroscience*. Heidelberg: Springer-Verlag.

Einstein, A. (1916). *Relativity: The special and the general theory*. New York: H. Holt and Company.

Elliott, M. T., Welchman, A. E., & Wing, A. M. (2009). Being discrete helps keep to the beat. *Experimental Brain Research*, *192*(4), 731–737. doi:10.1007/s00221-008-1646-8

Elliott, M. T., Wing, A. M., & Welchman, A. E. (2010). Multisensory cues improve sensorimotor synchronisation. *European Journal of Neuroscience*, *31*(10), 1828–1835. doi:10.1111/j.1460-9568.2010.07205.x

Elliott, M. T., Wing, A. M., & Welchman, A. E. (2014). Moving in time: Bayesian causal inference explains movement coordination to auditory beats. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1786), 20140751–20140751. doi:10.1523/JNEUROSCI.5561-10.2011

Ernst, M. O. (2006). A Bayesian view on multimodal cue integration. *Human Body Perception From the Inside Out*, 105–131.

Ernst, M. O., & Banks, M. S. M. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433. doi:10.1038/415429a

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*(4), 162–169. doi:10.1016/j.tics.2004.02.002

Escoffier, N., Sheng, D. Y. J., & Schirmer, A. (2010). Unattended musical beats enhance visual processing. *Acta Psychologica*, *135*(1), 12–16. doi:10.1016/j.actpsy.2010.04.005

Estes, K. G., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, *18*(3), 254–260.

Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Processe. *Archiv Für Die Gesamte Physiologie Des Menschen Und Der Tiere*, *11*(1), 403–432. doi:10.1007/BF01659311

Fechner, G. T. (1860). *Elemente dur psychophysik*. Leipzig, Germany: Breitkopf & Härtel.

Fernandes, H. L., Stevenson, I. H., Vilares, I., & Körding, K. P. (2014). The Generalization of Prior Uncertainty during Reaching. *Journal of Neuroscience*, *34*(34), 11470–11484. doi:10.1523/JNEUROSCI.3882-13.2014

Fiser, J., Berkes, P., Orbán, G., & Lengyel, M. (2010). Statistically optimal perception and

learning: from behavior to neural representations. *Trends in Cognitive Sciences*, *14*(3), 119–130. doi:10.1016/j.tics.2010.01.003

Fitch, W. T., & Rosenfeld, A. J. (2007). Perception and production of syncopated rhythms. *Music Perception*, *25*(1), 43–58. doi:10.1525/MP.2007.25.1.43

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836. doi:10.1038/335311a0

Friston, K. (2008). Hierarchical Models in the Brain. *PLoS Computational Biology*, *4*(11), e1000211. doi:10.1371/journal.pcbi.1000211.t001

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1211–1221. doi:10.1098/rstb.2008.0300

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience*, *7*(7), 773–778. doi:10.1038/nn1268

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review*, *107*(2), 289–344. doi:10.1037//0033-295X.107.2.289

Gallistel, C. R., King, A., & McDonald, R. (2004). Sources of variability and systematic error in mouse timing behavior. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*(1), 3–16. doi:10.1037/0097-7403.30.1.3

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*(3), 279.

Gibbon, J., & Rutschmann, R. (1969). Temporal order judgement and reaction time. *Science*, *165*(3891), 413–415. doi:10.1126/science.165.3891.413

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of Sciences*, *423*(1), 52–77.

Goldstone, S. (1968). Production and reproduction of duration: intersensory comparisons. *Perceptual and Motor Skills*, *26*(3), 755–760.

Goldstone, S., & Lhamon, W. T. (1974). Studies of auditory-visual differences in human time judgment. 1. Sounds are judged longer than lights. *Perceptual and Motor Skills*, *39*(1), 63–82. doi:10.2466/pms.1974.39.1.63

Grahn, J. A., & Rowe, J. B. (2013). Finding and Feeling the Musical Beat: Striatal Dissociations between Detection and Prediction of Regularity. *Cerebral Cortex*, *23*(4), 913–921. doi:10.1093/cercor/bhs083

Green, D. M., & Swets, J. A. (1973). *Signal Detection Theory and Psychophysics*. Huntingdon, NY: Krieger Publishing.

Greville, W. J., & Buehner, M. J. (2010). Temporal predictability facilitates causal learning. *Journal of Experimental Psychology: General*, *139*(4), 756–771. doi:10.1037/a0020976

Gribova, A., Donchin, O., Bergman, H., Vaadia, E., & de Oliveira, S. C. (2002). Timing of bimanual movements in human and non-human primates in relation to neuronal activity in primary motor cortex and supplementary motor area. *Experimental Brain Research*, *146*(3), 322–335. doi:10.1007/s00221-002-1174-x

Griffiths, T. L., & Tenenbaum, J. B. (2011). Predicting the future as Bayesian inference: People combine prior knowledge with observations when estimating duration and extent. *Journal of Experimental Psychology: General*, *140*(4), 725–743. doi:10.1037/a0024899

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23. doi:10.1016/j.tics.2005.11.006

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception &*

Psychophysics, *72*(3), 561–582. doi:10.3758/APP.72.3.561

Grondin, S., & McAuley, J. D. (2009). Duration discrimination in crossmodal sequences. *Perception*, *38*(10), 1542–1559. doi:10.1068/p6359

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

Hamlin, A. J. (1895). On the least observable interval between stimuli addressed to disparate senses and to different organs of the same sense. *The American Journal of Psychology*, *6*(4), 564–575.

Hancock, P. A., & Block, R. A. (2012). The psychology of time: a view backward and forward. *The American Journal of Psychology*, *125*(3), 267–274.

Hannon, E. E. E., & Trehub, S. E. S. (2005). Tuning in to musical rhythms: infants learn more readily than adults. *Pnas*, *102*(35), 12639–12643. doi:10.1073/pnas.0504254102

Hartcher-O'Brien, J., Di Luca, M., & Ernst, M. O. (2014). The duration of uncertain times: audiovisual information about intervals is integrated in a statistically optimal fashion. *PLoS ONE*, *9*(3), e89339. doi:10.1371/journal.pone.0089339

Hellström, Å., & Rammsayer, T. H. (2015). Time-order errors and standard-position effects in duration discrimination: An experimental study and an analysis by the sensation-weighting model. *Attention, Perception & Psychophysics*. doi:10.3758/s13414-015-0946-x

Helmholtz, Von, H. (1963). *Helmholtz's treatise on physiological optics*. New York: Dover.

Helson, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *The American Journal of Psychology*, *60*(1), 1–29.

Helson, H. (1964). *Adaptation-level theory: an experimental and systematic approach to behavior*. New York: Harper & Row.

Henderson, J., Hurly, T. A., Bateson, M., & Healy, S. D. (2006). Timing in Free-Living Rufous Hummingbirds, Selasphorus rufus. *Current Biology*, *16*(5), 512–515. doi:10.1016/j.cub.2006.01.054

Henry, M. J., & Herrmann, B. (2014). Low-Frequency Neural Oscillations Support Dynamic Attending in Temporal Context. *Timing & Time Perception*, *2*(1), 62–86. doi:10.1163/22134468-00002011

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

Honing, H. (2013). Structure and Interpretation of Rhythm in Music. In D. Deutsch, (3rd ed., pp. 369–404). London, UK: Academic Press. doi:10.1016/B978-0-12-381460-9.00009-2

Horr, N. K., & Di Luca, M. (2015a). Filling the blanks in temporal intervals: the type of filling influences perceived duration and discrimination performance. *Frontiers in Psychology*, *6*, 114–114. doi:10.3389/fpsyg.2015.00114

Horr, N. K., & Di Luca, M. (2015b). Taking a long look at isochrony: Perceived duration increases with temporal, but not stimulus regularity. *Attention, Perception & Psychophysics*, *77*(2), 592–602. doi:10.3758/s13414-014-0787-z

Hoyer, P. O., & Hyvarinen, A. (2003). Interpreting neural response variability as Monte Carlo sampling of the posterior. *Advances in Neural Information Processing Systems*, 293–300.

Höring, A. (1864). *Versuche über das Unterscheidungsvermögen des Hörsinnes für Zeitgrössen: Inaugural Dissertation*.

Huang, Y., & Rao, R. P. N. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 580–593. doi:10.1002/wcs.142

Ivry, R. B. R., Spencer, R. M. R., Zelaznik, H. N. H., & Diedrichsen, J. J. (2002). The

cerebellum and event timing. *Annals of the New York Academy of Sciences*, *978*, 302–317. doi:10.1111/j.1749-6632.2002.tb07576.x

Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience*, *1*(2), 136–152. doi:10.1162/jocn.1989.1.2.136

Ivry, R. B., & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences*, *12*(7), 273–280. doi:10.1016/j.tics.2008.04.002

James, W. (1890). *The Principles of Psychology*. New York: H. Holt and Company.

Janssen, P., & Shadlen, M. N. (2005). A representation of the hazard rate of elapsed time in macaque area LIP. *Nature Neuroscience*, *8*(2), 234–241. doi:10.1038/nn1386

Jaśkowski, P. (1992). Temporal-order judgment and reaction time for short and long stimuli. *Psychological Research*, *54*(3), 141–145.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, *13*(8), 1020–1026. doi:10.1038/nn.2590

Jones, M. R., & Boltz, M. (1989). Dynamic Attending and Responses to Time. *Psychological Review*, *96*(3), 459–491.

Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the Absence of Clocks: Encoding Time in Neural Network States. *Neuron*, *53*(3), 427–438. doi:10.1016/j.neuron.2007.01.006

Kersten, D., & Yuille, A. (2003). Bayesian models of object perception. *Current Opinion in Neurobiology*, *13*(2), 150–158. doi:10.1016/S0959-4388(03)00042-4

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271–304. doi:10.1146/annurev.psych.55.090902.142005

King, A. J. (2005). Multisensory integration: strategies for synchronization. *Current Biology*, *15*(9), R339–41. doi:10.1016/j.cub.2005.04.022

Knill, D. C. (2007). Learning Bayesian priors for depth perception. *Journal of Vision*, *7*(8), 13.

Knill, D. C., & Richards, W. (1996). *Perception as Bayesian Inference*. Cambridge, UK: Cambridge University Press.

Kohfeld, D. L. (1971). Simple reaction time as a function of stimulus intensity in decibels of light and sound. *Journal of Experimental Psychology*, *88*(2), 251–257. doi:10.1037/h0030891

Kok, P., Brouwer, G. J., van Gerven, M. A. J., & de Lange, F. P. (2013). Prior Expectations Bias Sensory Representations in Visual Cortex. *Journal of Neuroscience*, *33*(41), 16275–16284. doi:10.1523/JNEUROSCI.0742-13.2013

Kording, K. P., & Wolpert, D. M. (2004). The loss function of sensorimotor learning. *Pnas*, *101*(26), 9839–9842. doi:10.1073/pnas.0308394101

Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, *427*(6971), 244–247.

Kösem, A., & van Wassenhove, V. (2012). Temporal Structure in Audiovisual Sensory Selection. *PLoS ONE*, *7*(7), e40936. doi:10.1371/journal.pone.0040936.s006

Kristofferson, A. B. (1967). Successiveness discrimination as a two-state, quantal process. *Science*, *158*, 1337–1339.

Kusnir, F., Chica, A. B., Mitsumasu, M. A., & Bartolomeo, P. (2011). Consciousness and Cognition. *Consciousness and Cognition*, *20*(4), 1201–1210. doi:10.1016/j.concog.2011.01.012

Laje, R., & Buonomano, D. V. (2013). Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nature Publishing Group*, *16*(7), 925–933. doi:10.1038/nn.3405

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment

of Neuronal Oscillations as a Mechanism of Attentional Selection. *Science*, *320*(5872), 110–113. doi:10.1126/science.1154735

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology*, *94*(3), 1904–1911. doi:10.1152/jn.00263.2005

Lapid, E., Ulrich, R. R., & Rammsayer, T. (2008). On estimating the difference limen in duration discrimination tasks: A comparison of the 2AFC and the reminder task. *Perception & Psychophysics*, *70*(2), 291–305. doi:10.3758/PP.70.2.291

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159.

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, *26*(1), 1–37. doi:10.1207/s15516709cog2601_1

Large, E. W., & Snyder, J. S. (2009). Pulse and Meter as Neural Resonance. *Annals of the New York Academy of Sciences*, *1169*(1), 46–57. doi:10.1111/j.1749-6632.2009.04550.x

Lejeune, H., & Wearden, J. H. (2009). Vierordt's The Experimental Study of the Time Sense (1868) and its legacy. *Eur. J. Cogn. Psychol.*, *21*(6), 941–960. doi:10.1080/09541440802453006

Lewicki, M. S. (2002). Efficient coding of natural sounds. *Nature Neuroscience*, *5*(4), 356–363. doi:10.1038/nn831

Lewis, P. A., & Miall, R. C. (2003). Distinct systems for automatic and cognitively controlled time measurement: evidence from neuroimaging. *Current Opinion in Neurobiology*, *13*(2), 250–255. doi:10.1016/S0959-4388(03)00036-9

Lucas, C. G., & Griffiths, T. L. (2009). Learning the Form of Causal Relationships Using Hierarchical Bayesian Models. *Cognitive Science*, *34*(1), 113–147. doi:10.1111/j.1551-6709.2009.01058.x

Lunney, H. W. (1974). Time as heard in speech and music. *Nature*, *249*(457), 592–592. doi:10.1038/249592a0

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, *9*(11), 1432–1438. doi:10.1038/nn1790

Maloney, L. T., & Mamassian, P. (2009). Bayesian decision theory as a model of human visual perception: Testing Bayesian transfer. *Visual Neuroscience*, *26*(01), 147–155. doi:10.1017/S0952523808080905

Mamassian, P., Landy, M. S., & Maloney, L. T. (2002). Bayesian modelling of visual perception. In *Probabilistic models of the brain: Perception and neural function* (pp. 13–36). Cambridge, MA: MIT Press.

Mangels, J. A., Ivry, R. B., & Shimizu, N. (1998). Dissociable contributions of the prefrontal and neocerebellar cortex to time perception. *Brain Research. Cognitive Brain Research*, *7*(1), 15–39.

Manning, F., & Schutz, M. (2013). "Moving to the beat" improves timing perception. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-013-0439-7

Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Cognitive Brain Research*, *21*(2), 139–170. doi:10.1016/j.cogbrainres.2004.06.012

Mauk, M. D., & Buonomano, D. V. (2004). The neural basis of temporal processing. *Annual Review of Neuroscience*, *27*, 307–340. doi:10.1146/annurev.neuro.27.070203.144247

McAuley, J. D. (1995, July 1). *Perception of time as phase: Toward an adaptive-oscillator*

model of rhythmic pattern processing. Indiana University.

McAuley, J. D., & Jones, M. R. (2003). Modeling Effects of Rhythmic Context on Perceived Duration: A Comparison of Interval and Entrainment Approaches to Short-Interval Timing. *Journal of Experimental Psychology: Human Perception and Performance, 29*(6), 1102–1125. doi:10.1037/0096-1523.29.6.1102

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

McDonald, J. J. J., Teder-Sälejärvi, W. A. W., Di Russo, F. F., & Hillyard, S. A. S. (2005). Neural basis of auditory-induced shifts in visual time-order perception. *Nature Neuroscience, 8*(9), 1197–1202. doi:10.1038/nn1512

McNeill, W. H. (1995). *Keeping together in time: Dance and drill in human history*. Cambridge, MA: Harvard Univ. Press.

Meck, W. H. (2003). *Functional and Neural Mechanisms of Interval Timing*. (W. H. Meck). Boca Ration, FL: CRC.

Meck, W. H. (2005). Neuropsychology of timing and time perception. *Brain and Cognition, 58*(1), 1–8. doi:10.1016/j.bandc.2004.09.004

Merchant, H., & de Lafuente, V. (2014). Introduction to the neurobiology of interval timing. *Advances in Experimental Medicine and Biology, 829*, 1–13. doi:10.1007/978-1-4939-1782-2_1

Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat: a neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*(1664), 20140093–20140093. doi:10.1016/j.cub.2009.03.038

Merker, B. H., Madison, G. S., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex, 45*(1), 4–17.

Miall, C. (1989). The storage of time intervals using oscillating neurons. *Neural Computation, 1*(3), 359–371.

Miller, J. J., & Ulrich, R. R. (2001). On the analysis of psychometric functions: the Spearman-Kärber method. *Perception & Psychophysics, 63*(8), 1399–1420.

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

Miller, J., & Schwarz, W. (2006). Dissociations between reaction times and temporal order judgments: a diffusion model approach. *J. Exp. Psychol. Human.* 32, 394–412.

Miyazaki, M., Nozaki, D., & Nakajima, Y. (2005). Testing Bayesian models of human coincidence timing. *Journal of Neurophysiology, 94*, 395–399.

Miyazaki, M., Yamamoto, S., Uchida, S., & Kitazawa, S. (2006). Bayesian calibration of simultaneity in tactile temporal order judgment. *Nature Neuroscience, 9*(7), 875–877. doi:10.1038/nn1712

Nobre, A. C., Correa, A., & Coull, J. T. (2007). The hazards of time. *Current Opinion in Neurobiology, 17*(4), 465–470. doi:10.1016/j.conb.2007.07.006

Ohyama, T., Gibbon, J., Deich, J. D., & Balsam, P. D. (1999). Temporal control during maintenance and extinction of conditioned keypecking in ring doves. *Animal Learning & Behavior, 27*(1), 89–98.

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

Petzschner, F. H., & Glasauer, S. (2011). Iterative Bayesian Estimation as an Explanation for Range and Regression Effects: A Study on Human Path Integration. *Journal of Neuroscience, 31*(47), 17220–17229. doi:10.1523/JNEUROSCI.2028-11.2011

Petzschner, F. H., Glasauer, S., & Stephan, K. E. (2015). A Bayesian perspective on magnitude estimation. *Trends in Cognitive Sciences, 19*(5), 285–293. doi:10.1016/j.tics.2015.03.002

Piéron, H. (1913). II. Recherches sur les lois de variation des temps de latence sensorielle en fonction des intensités excitatrices. *L'année Psychologique, 20*(1), 17–96. doi:10.3406/psy.1913.4294

Piéron, H. (1952). *The sensations, their functions, processes, and mechanisms*. New Haven, CT: Yale University Press.

Pins, D., & Bonnet, C. (1996). On the relation between stimulus intensity and processing time: Piéron's law and choice reaction time. *Perception & Psychophysics, 58*(3), 390–400.

Pouget, A., Dayan, P., & Zemel, R. (2000). Information processing with population codes. *Nature Reviews: Neuroscience, 1*(2), 125–132. doi:10.1038/35039062

Povel, D. J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance, 7*(1), 3–18.

Povel, D. J. (1984). A theoretical framework for rhythm perception. *Psychological Research, 45*(4), 315–337.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59–108.

Repp, B. H. (1987). The sound of two hands clapping: an exploratory study. *The Journal of the Acoustical Society of America, 81*(4), 1100–1109. doi:10.1121/1.394630

Repp, B. H. (1999). Detecting deviations from metronomic timing in music: effects of perceptual structure on the mental timekeeper. *Perception & Psychophysics, 61*(3), 529–548.

Repp, B. H. B. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic Bulletin & Review, 12*(6), 969–992.

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review, 20*(3), 403–452. doi:10.3758/s13423-012-0371-2

Roach, N. W., Heron, J., & 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. doi:10.1037/0033-2909.88.3.638

Rohenkohl, G., & Nobre, A. C. (2011). Alpha Oscillations Related to Anticipatory Attention Follow Temporal Expectations. *Journal of Neuroscience, 31*(40), 14076–14084. doi:10.1523/JNEUROSCI.3387-11.2011

Rohenkohl, G., Coull, J. T., & Nobre, A. C. (2011). Behavioural Dissociation between Exogenous and Endogenous Temporal Orienting of Attention. *PLoS ONE, 6*(1), e14620. doi:10.1371/journal.pone.0014620.g002

Rohenkohl, G., Gould, I. C., Pessoa, J., & Nobre, A. C. (2014). Combining spatial and temporal expectations to improve visual perception. *Journal of Vision, 14*(4), 8–8. doi:10.1167/14.4.8

Sato, Y., & Aihara, K. (2009). Integrative Bayesian model on two opposite types of sensory adaptation. *Artificial Life and Robotics, 14*(2), 289–292. doi:10.1007/s10015-009-0675-0

Schiffman, H. R., & Bobko, D. J. (1977). The role of number and familiarity of stimuli in the perception of brief temporal intervals. *The American Journal of Psychology*, 85–93.

Schirmer, A. (2004). Timing speech: a review of lesion and neuroimaging findings. *Cognitive Brain Research, 21*(2), 269–287. doi:10.1016/j.cogbrainres.2004.04.003

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as

instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. doi:10.1016/j.tins.2008.09.012

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of Active Sensing and perceptual selection. *Current Opinion in Neurobiology*, *20*(2), 172–176. doi:10.1016/j.conb.2010.02.010

Schulze, H. H. (1978). The detectability of local and global displacements in regular rhythmic patterns. *Psychological Research*, *40*(2), 173–181.

Schulze, H. H. (1989). The perception of temporal deviations in isochronic patterns. *Perception & Psychophysics*, *45*(4), 291–296.

Schwartze, M., & Kotz, S. A. (2013). A dual-pathway neural architecture for specific temporal prediction. *Neuroscience & Biobehavioral Reviews*, *37*(10), 2587–2596. doi:10.1016/j.neubiorev.2013.08.005

Sciutti, A., Burr, D., Saracco, A., Sandini, G., & Gori, M. (2014). Development of context dependency in human space perception. *Experimental Brain Research*, *232*(12), 3965–3976. doi:10.1007/s00221-014-4021-y

Shi, Z., Church, R. M., & Meck, W. H. (2013). Bayesian optimization of time perception. *Trends in Cognitive Sciences*, *17*(11), 556–564. doi:10.1016/j.tics.2013.09.009

Simoncelli, E. P. (2003). Vision and the statistics of the visual environment. *Current Opinion in Neurobiology*, *13*(2), 144–149. doi:10.1016/S0959-4388(03)00047-3

Spence, C., & Parise, C. (2010). Prior-entry: A review. *Consciousness and Cognition*, *19*(1), 364–379. doi:10.1016/j.concog.2009.12.001

Spence, C., & Squire, S. (2003). Multisensory Integration: Maintaining the Perception of Synchrony. *Current Biology*, *13*(13), R519–R521. doi:10.1016/S0960-9822(03)00445-7

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, *130*(4), 799–832.

Srinivasan, M. V., Laughlin, S. B., & Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proceedings of the Royal Society B: Biological Sciences*, *216*(1205), 427–459.

Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: Fundamental issues and a general model. *Attention and Performance IV*, 629–685.

Sternberg, S., Knoll, R. L., & Gates, B. A. (1971). Prior entry reexamined: Effect of attentional bias on order perception.

Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*, *9*(4), 578–585. doi:10.1038/nn1669

Su, Y.-H. Y., & Pöppel, E. E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological Research*, *76*(3), 373–382. doi:10.1007/s00426-011-0346-3

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403–409. doi:10.1016/j.tics.2009.06.003

Taatgen, N., & van Rijn, H. (2011). Traces of times past: representations of temporal intervals in memory. *Memory & Cognition*, *39*(8), 1546–1560. doi:10.3758/s13421-011-0113-0

Tanaka, S., Tsuzaki, M., Aiba, E., & Kato, H. (2008). Auditory sensitivity to temporal deviations from perceptual isochrony: Comparison of the starting point and ending point of acoustic change1. *Japanese Psychological Research*, *50*(4), 9.

Teki, S., Grube, M., & Griffiths, T. D. (2011). A unified model of time perception accounts for duration-based and beat-based timing mechanisms. *Frontiers in Integrative Neuroscience*, *5*. doi:10.3389/fnint.2011.00090/abstract

ten Hoopen, G. G., Hartsuiker, R. R., Sasaki, T. T., Nakajima, Y. Y., Tanaka, M. M., & Tsumura, T. T. (1995). Auditory isochrony: time shrinking and temporal patterns. *Perception*, *24*(5), 577–593.

ten Hoopen, G. G., Van Den Berg, S., Memelink, J., Bocanegra, B., & Boon, R. (2011). Multiple-look effects on temporal discrimination within sound sequences. *Attention, Perception & Psychophysics*, *73*(7), 2249–2269. doi:10.3758/s13414-011-0171-1

Thomas, E. C., & Brown, I. (1974). Time perception and the filled-duration illusion. *Perception & Psychophysics*, *16*(3), 449–458.

Tobin, S., Bisson, N., & Grondin, S. (2010). An Ecological Approach to Prospective and Retrospective Timing of Long Durations: A Study Involving Gamers. *PLoS ONE*, *5*(2), e9271. doi:10.1371/journal.pone.0009271.t002

Treisman, M. (1963). Temporal discrimination and the indifference interval. Implications for a model of the "internal clock". *Psychological Monographs*, *77*(13), 1–31.

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit Perceptual Anticipation Triggered by Statistical Learning. *Journal of Neuroscience*, *30*(33), 11177–11187. doi:10.1523/JNEUROSCI.0858-10.2010

Ulrich, R. R. (1987). Threshold models of temporal-order judgments evaluated by a ternary response task. *Perception & Psychophysics*, *42*(3), 224–239.

Ulrich, R. R., Rinkenauer, G., & Miller, J. (1998). Effects of stimulus duration and intensity on simple reaction time and response force. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 915–928.

Umemoto, A., Scolari, M., Vogel, E. K., & Awh, E. (2010). Statistical learning induces discrete shifts in the allocation of working memory resources. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1419–1429. doi:10.1037/a0019324

Van der Burg, E., Alais, D., & Cass, J. (2013). Rapid Recalibration to Audiovisual Asynchrony. *Journal of Neuroscience*, *33*(37), 14633–14637. doi:10.1523/JNEUROSCI.1182-13.2013

van Rijn, H., Gu, B.-M., & Meck, W. H. (2014). Dedicated Clock/Timing-Circuit Theories of Time Perception and Timed Performance. In *Advances in Experimental Medicine and Biology* (Vol. 829, pp. 75–99). New York, NY: Springer New York. doi:10.1007/978-1-4939-1782-2_5

van Wassenhove, V., Buonomano, D. V., Shimojo, S., & Shams, L. (2008). Distortions of Subjective Time Perception Within and Across Senses. *PLoS ONE*, *3*(1), e1437. doi:10.1371/journal.pone.0001437.s003

VanRullen, R., & Dubois, J. (2011). The psychophysics of brain rhythms. *Frontiers in Psychology*, *2*, 1–10. doi:10.3389/fpsyg.2011.00203

VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Sciences*, *7*(5), 207–213. doi:10.1016/S1364-6613(03)00095-0

Vierordt, C. (1868). *Der zeitsinn nach versuchen*. Tübingen, Germany: Laupp.

Vilares, I., & Körding, K. P. (2011). Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Annals of the New York Academy of Sciences*, *1224*(1), 22–39. doi:10.1111/j.1749-6632.2011.05965.x

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception & Psychophysics*, *72*(4), 871–884. doi:10.3758/APP.72.4.871

Vuust, P., & Witek, M. A. G. (2014). Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology*, *5*, 1111–1111. doi:10.3389/fpsyg.2014.01111

Wackermann, J., & Ehm, W. (2006). The dual klepsydra model of internal time

representation and time reproduction. *Journal of Theoretical Biology*, *239*(4), 482–493. doi:10.1016/j.jtbi.2005.08.024

Wainwright, M. J. (1999). Visual adaptation as optimal information transmission. *Vision Research*, *39*(23), 3960–3974.

Wallace, M., & Rabin, A. I. (1960). Temporal experience. *Psychological Bulletin*, *57*(3), 213–236.

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": application of a model of the internal clock in humans. *The Quarterly Journal of Experimental Psychology. B, Comparative and Physiological Psychology*, *51*(2), 97–120. doi:10.1080/713932672

Wearden, J. H., Norton, R., Martin, S., & Montford-Bebb, O. (2007). Internal clock processes and the filled-duration illusion. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(3), 716–729. doi:10.1037/0096-1523.33.3.716

Wearden, J. H., Todd, N. P. M., & Jones, L. A. (2006). When do auditory/visual differences in duration judgements occur? *The Quarterly Journal of Experimental Psychology*, *59*(10), 1709–1724. doi:10.1080/17470210500314729

Wei, X. X., & Stocker, A. (2012). Efficient coding provides a direct link between prior and likelihood in perceptual Bayesian inference. *Advances in Neural Information ….*

Wei, X.-X., & Stocker, A. A. (2015). A Bayesian observer model constrained by efficient coding can explain "anti-Bayesian" percepts. *Nature Neuroscience*, 1–11. doi:10.1038/nn.4105

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, *5*(6), 598–604. doi:10.1038/nn858

Wing, A. M., & Kristofferson, A. B. (1973). The timing of interresponse intervals. *Perception & Psychophysics*, *13*(3), 455–460.

Winkler, I., Haden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Pnas*, *106*(7), 2468–2471. doi:10.1073/pnas.0809035106

Wolpert, D. M. (2007). Probabilistic models in human sensorimotor control. *Human Movement Science*, *26*(4), 511–524. doi:10.1016/j.humov.2007.05.005

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, *3 Suppl*(supp), 1212–1217. doi:10.1038/81497

Yamamoto, S., & Kitazawa, S. (2001). Reversal of subjective temporal order due to arm crossing. *Nature Neuroscience*, *4*(7), 759–765. doi:10.1038/89559

Yamamoto, S., Miyazaki, M., Iwano, T., & Kitazawa, S. (2012). Bayesian Calibration of Simultaneity in Audiovisual Temporal Order Judgments. *PLoS ONE*, *7*(7), e40379. doi:10.1371/journal.pone.0040379.g005

Yarrow, K., Jahn, N., Durant, S., & Arnold, D. H. (2011). Shifts of criteria or neural timing? The assumptions underlying timing perception studies. *Consciousness and Cognition*, *20*(4), 1518–1531. doi:10.1016/j.concog.2011.07.003

Zakay, D. (1993). Time estimation methods do they influence prospective duration estimates? *Perception*, *22*, 91–101.

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

Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental Brain Research*, *152*(2), 198–210. doi:10.1007/s00221-003-1536-z

Zampini, M., Shore, D. I., & Spence, C. (2005b). Audiovisual prior entry. *Neuroscience Letters*, *381*(3), 217–222. doi:10.1016/j.neulet.2005.01.085

Zanto, T. P., Snyder, J. S., & Large, E. W. (2006). Neural correlates of rhythmic expectancy. *Advances in Cognitive Psychology*, *2*(2), 221–231. doi:10.2478/v10053-

008-0057-5

Zemel, R. S., Dayan, P., & Pouget, A. (1998). Probabilistic interpretation of population codes. *Neural Computation*, *10*(2), 403–430.