EFFECTS OF TEMPORAL EXPECTATION ON COMPLEX DECISION MAKING

David Greatrex
Christ's College
Centre for Music and Science
University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy.

August 2017

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. This dissertation does not exceed 80,000 words including tables and footnotes, excluding tables of content, appendices, bibliography, and diagrams.

Abstract

Many complex decisions require integrating and assessing multiple streams of dynamic information whilst determining how to act. This dynamic information often contains rhythmic structures which our sensory systems can adapt to and use to anticipate future events. Despite the close relationship between rhythmic temporal expectations and complex decision making being self evident, no experiments explicitly attempt to understand this interdependence. If the theories that have emerged from both domains are to generalise to complex interactive behaviour, the effects of dynamic context on complex decisions must be considered.

I argue that timing research must move beyond simple decisions and develop a new experimental framework for addressing the problem. This includes increasing the complexity of experimental tasks, testing the effects of timing on perceptual averaging and subjective value decisions, incorporating timing as an inherent dimension of targets, testing degrees of aperiodicity and exploring the effects that prior knowledge about the temporal structure of a stimulus has on choice. Seven behavioural experiments are reported that implement the new experimental framework. Five use a complex auditory-spatial averaging task to examine effects of periodicity, expectation, prior knowledge and related parameters such as IOI variance. One tests the effects of rhythmic variability and stimulus duration on auditory detection to determine specificity to complex decision making, and one investigates the effects of timing on audio-visual subjective value decisions.

The results show that existing theories of temporal expectation do not necessarily generalise to complex decision making. Periodicity reduces the amount of information that is needed to form complex decisions. However, the effects of periodicity (or degree of aperiodicity) on choice are dependent on a number of factors associated with prior knowledge, stimulus rate, variance, decision type and task complexity. Using these findings I develop an explanatory framework called "dynamic inhibition and boosting" that better accounts for behavioural data in the literature compared with existing theories. This explanation is supported by the novel proposal that temporal expectations influence confidence and perceived risk.

Acknowledgements

I wish to express my sincere appreciation to my supervisors Professors Sarah Hawkins and Ian Cross for their unwavering support, encouragement and inspiration throughout this project and more generally during my time in Cambridge. Their continued patience, openness and strength has been an invaluable gift that has allowed me to follow my interests and investigate topics that I am passionate about. I am extremely grateful to them for setting examples of scholarship and scientific thinking that will stay with me as I take on new challenges. I would also like to thank them for accepting me to study at the Centre for Music and Science (CMS) and I hope that this document is evidence of that being a good decision.

I would also like to thank my many colleagues associated with the CMS who have provided a constant source of inspiration. Every year there is such a diverse range of ideas being worked on that it makes it a fascinating community to be a part of. Notably, I would like to thank Dr. Andrew Goldman, Guo Yu, Dr. Dina Kronhaus, Juan Pablo Robledo del Canto and Dr. Mark Gotham for the many conversations that I had about the ideas contained in this thesis.

The support of the wider academic community should also be noted. I owe gratitude to Professors Bradley Love, Kia Nobre and Jan Schnupp for giving me the confidence to follow my interests in decision making and timing research at a time when I was transitioning into the PhD. I would also like to thank Dr. Marcus Pearce and Dr. Benedetto De Martino for the effort they gave to review my first year report and to Professors Brian Moore and Stephen Grossberg for their insightful suggestions. I am grateful to the Cambridge Computer Laboratory for the many lectures that I attended whilst developing skills used throughout the thesis and I would like to thank Dr. Liz Garnett and Mike Williams for the many teachings that I called on. This work was kindly supported by grants from the Arts and Humanities Research Council and Christ's College Cambridge.

Finally, I would like to thank my family for all of their support. To my sister Helen for being a fantastic academic role model and friend. To my parents Jan and Tony Greatrex for their constant support of my pursuits. To my wife Anna, for being brilliant in so many ways and for sharing this journey with me.

Contents

List of Figures

List of Tables

List of Links

1. Experimental software:

   Experiment 1: https://github.com/dcgreatrex-phd/experiment_1

   Experiment 2: https://github.com/dcgreatrex-phd/experiment_2

   Experiment 3: https://github.com/dcgreatrex-phd/experiment_3

   Experiment 4: https://github.com/dcgreatrex-phd/experiment_4

   Experiment 5: https://github.com/dcgreatrex-phd/experiment_5

   Experiment 6: https://github.com/dcgreatrex-phd/experiment_6

   Experiment 7: https://github.com/dcgreatrex-phd/experiment_7

2. Experiment 7: Pilot study website

   http://dcgreatrex-phd-experiment-7.s3-website-us-east-1.amazonaws.com

Chapter 1

Introduction

1.1 Aims of the thesis

This thesis is founded on an original argument concerning the unsuitability of widely used experimental approaches in the timing literature. It argues that to understand the effects of temporal expectation on interactive behaviour, stimulus timing must be contextualised and modelled as part of, and not separate from, goal relevant information. This has required expanding the contexts in which temporal expectation is typically investigated and exploring how the timing of a stimulus biases complex decisions. Complex decision making is defined throughout this thesis as choices that are based on more than one piece of perceptual information that require integrating memorised content into decision options and/or weighting options in terms of subjective value. The benefits of this approach are two-fold. Firstly, it challenges the use of temporal expectation paradigms in which participants respond to and categorise isolated stimuli preceded by an unrelated rhythmic pulse. Whilst such experiments are effective in showing how the structure of incoming sensory information biases simple classification decisions and sensory processing (Large and Jones, 1999; Barnes and Jones, 2000; Doherty et al., 2005; Schroeder and Lakatos, 2009; Henry and Obleser, 2012; Cravo et al., 2013; Hickok et al., 2015), they provide little evidence for whether these findings generalise to more complex decisions or markably influence everyday interactive behaviour. Secondly, it ensures that timing is investigated in a way that has relevance for complex decision making and the decision making literature in general. Whilst decision scientists have been successful in mapping the behavioural and, more recently, neural correlates of a wide range of decisions (Smith and Vickers, 1988; Gold and Shadlen, 2001, 2007; Smith and Ratcliff, 2004; Glimcher, 2004;

Carpenter et al., 2009; Plassmann et al., 2007; Behrens et al., 2007; Rushworth et al., 2009; Alvarez, 2011), stimulus timing has rarely been the explicit focus of attention. In cases where it has, as in temporal discounting research (Frederick et al., 2002; Kable and Glimcher, 2007), the research has focused on the effects of long time intervals between a stimulus and reward and not on fast rhythmic event timing which is characteristic of interactive behaviours and temporal expectation paradigms. This thesis therefore brings both the temporal expectation and decision making literatures together to explore interdependence between rhythmic variability, temporal expectations and complex decision making.

The thesis contains a series of seven auditory experiments which were designed to achieve two primary aims. The first aim is to determine whether and how the timing of decision information systematically biases complex decisions. This is motivated by a desire to test the generalisability of current timing and decision theories under experimental scenarios more aligned with goal-directed everyday behaviour. The second aim is to facilitate change in the experimental psychology and neuroscientific literatures by adopting a new experimental approach that reshapes widely used experimental methods. This contributes to knowledge in three ways: 1. It develops theoretical and methodological connections between the temporal expectation and decision making literatures that have until now remained relatively separate from one another. 2. It generates data showing how onset variance within auditory sequences impacts complex averaging and value-based choice. 3. It determines whether stimulus timing should be incorporated into models of complex decision making.

1.2 Setting the scene

1.2.1 Movement and time

Neurons fire, hearts beat, eyelids blink, and bodies move. Without the ability to move it would be extremely hard for us to seek what we need in order to survive and procreate. Movement is what gives purpose to the cognitive system and leads to us experiencing a life in motion (Marsh et al., 2009). It is also what enables us to detect agency and is the medium through which we interact with others (Frith and Frith, 2010).

Movement is the reason why perception, cognition and behaviour cannot be reduced to observational exchanges in which static information is processed and responded to unidirectionally. Take the example of a busy town square. Some people walk in opposite directions, others stroll side-by-side and many stand engaged in communicative turn-taking. From afar the scene appears chaotic and disorderly; up-close it comprises a complex interaction of localised, shared and intended actions. Attention is sequentially caught and directed towards different aspects of the scene depending on your goals. This selectivity relies on active engagement with the event and on making anticipated and synchronised movements. The temporality of the event ensures that the brain must be proactive (Nobre et al., 2007).

1.2.2 Goals

Many of the movements in the square can be understood from the perspective of goals. If your goal is to arrive at your next appointment on time, then running across the square will decrease the likelihood that you are late. If you are hungry, then standing in-line to buy a sandwich is a socially accepted way to acquire food. A goal is the object to which effort or ambition is directed (OED, 2016). It can apply to an individual or constitute joint intended effort aimed at facilitating shared beneficial outcomes (Gold and Shadlen, 2007).

In addition to movement, perception and the attentional system must also be understood from the perspective of goals. Goals exist on multiple levels and are characterised by the anticipated benefit that they afford an organism. If a goal is to communicate a concept from your last meeting to a colleague then it will be unlikely that you will pay attention to the colour of a market stall tent; not to mention a person dressed as a gorilla (Simons and Chabris, 1999). However, if the gorilla makes threatening noises and suddenly starts to run across the square at you, your attention would be captured by this unlikely event. The goal is now to survive. You therefore put all of your effort into trying to avoid harm. Goal availability, anticipated reward and associated action costs can change unexpectedly with time and previous actions (Christopoulos et al., 2015). The cognitive system must therefore proactively sample the environment for information that facilitates dynamically competing goals during movement.

1.2.3 Prediction

Instead of running away from the gorilla, you smile and walk towards it. This is because you know that today your eccentric friend was planning to wear a costume whilst fundraising in the square. This is an example of prediction based on prior knowledge and expectation. As the gorilla's height, build and gait is similar to that of your friend, you predict that it does not intend to hurt you. Predicting the nature and outcome of others' actions and knowing what they are going to do next is crucial for survival (Sebanz et al., 2006; Frith, 2007). Without this information one is forced to rely on sensory patterns and ostensive social signals that are predictive of goals, intent and time to impact whilst preparing to move. This is an example of context-sensitive prediction. Examples of predictive cues are the direction of eye gaze, gestures, rate of movement, trajectory of movement and the attention and facial expressions of others in the room (Kilner et al., 2007; Frith and Frith, 2007, 2010). Context-sensitive prediction requires the fast classification of sensory information and its integration with prior knowledge for the purpose of prediction.

Pattern completion and the extrapolation of structure is an important characteristic of the predictive mind. Humans are drawn towards repeating statistical patterns and actively construct cognitive templates on which to guide their behaviour (Jones and Boltz, 1989; Wolpert et al., 2003). Dynamically extrapolating the rate and rhythmic patterning of another's footsteps provides a predictable forward-looking template on which to focus attention and synchronise action. Sensing a common pulse enables friends to coordinate their actions whilst dancing in time to music. Although a pulse may be perceived, it does not necessarily mean that there is one in the physical signal (Snyder and Large, 2005; Large and Snyder, 2009). Rather, it may constitute an emergent cognitive property that focuses attention on specific aspects of the sensory stream and facilitates the coordination of multilevel cognitive processes in a way that allows one to "make sense" of dynamic and often incomplete perceptual information (Velasco and Large, 2011). Averaging, "filling in" and detecting structure in noise are all fundamental to aspects of speech perception, intelligibility, object/event recognition and successfully timed coordination (for examples, see: Warren, 1984; Kanizsa, 1985; Repp, 2005; Large, 2008; Hawkins, 2010).

1.2.4 Context

Cognitive structures, concepts and classes emerge through one's experiences of what is functionally significant and meaningful in the immediate and recalled environment (Gallagher, 2009). The sensations and movements within the square are entirely dependent on factors such as how bodies are positioned, prior experience, attentional allocation and the types of emotions that are induced during interaction. Implicit understanding of this context strongly influences actions so that they are functionally appropriate and it is context that makes our actions meaningful or not (Beer, 2014).

1.3 The problem

Our ability to track temporal regularities in the environment and generate expectations about when events will occur is essential for successful human interaction and survival. It is therefore not surprising that our perceptual and attentional systems are tuned toward temporally structured moments in time (Jones, 1976; Nobre et al., 2007; Nobre and Rohenkohl, 2014). Passive entrainment towards predictable temporal cues and rhythmic patterns increase the accuracy of time interval judgements (Large and Jones, 1999; Jones and Boltz, 1989; Jones, 2004, 2009, 2010), enhance early sensory responses, entrain neural oscillations and increase perceptual contrast sensitivity and gain (Correa et al., 2005; Doherty et al., 2005; Rohenkohl and Nobre, 2011; Rohenkohl et al., 2012; Cravo et al., 2013; Mathewson et al., 2010, 2012). Temporal expectations, it seems, are closely related to sensory encoding and the memorisation of information.

In combination with predictive attending, humans use learned representations of goals, anticipated outcomes and perceived value as means of calculating the most beneficial behaviours to pursue (Behrens et al., 2007; Kable and Glimcher, 2009; Rangel et al., 2008; Sebanz et al., 2006; Vesper et al., 2011). This ability is commonly investigated from the perspective of animal foraging or complex decision making and requires dynamically integrating multiple sources of perceptual information with learned beliefs about the world. Knowing that colliding with a moving car can cause serious injury will increase the time taken to gather relevant perceptual evidence used to predict the car's movement. The time allocated for this prediction will depend on the importance that is placed on crossing the road at that moment in time. A greater expectation that the considered action will be beneficial will decrease the willingness to wait and increase the likelihood of engaging in risky action. Successful coordinated interaction therefore relies on interdependence between predictive attending and complex decision making.

Whilst there is a large body of literature on rhythmic temporal expectations and complex decision making, there have been no explicit attempts to understand their interdependence. The closest examples in the temporal expectation literature, albeit not specific to complex decisions, are Rohenkohl et al. (2012), Cravo et al. (2013) and Jepma et al. (2012) who used decision theoretic models to investigate what effect temporal expectations had on simple classification decisions. The closest examples made by complex decision researchers are Armel et al. (2008) and Lim et al. (2011) who varied eye-gaze duration during an economic valuation paradigm. Neither author, however, systematically varied temporal structures nor referenced any temporal expectation research. Typically, nearly all psychophysical tasks used to investigate temporal expectations require participants to classify simple isolated targets that are preceded by an unrelated rhythmic pulse (Sameiro-Barbosa and Geiser, 2016). This approach has two limitations: Firstly, it is insufficient because the meaning and nature of goal-relevant decision information (e.g. the target stimulus) is artificially separated from that of timing information (e.g. the preceding rhythmic pulse). In everyday life, however, timing is inherent within goal relevant information (such as the rhythm and rate of speech). It functions as a dimension through which relevant information is experienced and responded to. For this reason, timing must remain integrated with, and not separated from, goal relevant decision-information. Secondly, the restriction of the standard psychophysical approach to simple classification judgements has very limited application, since most interactive decisions require integrating dynamic sensory information from multiple sources and time points with that of preference and anticipated reward information. The types of judgements under investigation must be expanded if current timing theories are to generalise to complex goal-directed behaviour.

A similar critique can also be applied to the complex decision making literature. Economic decision making tasks often require participants to make preference-led decisions between pairs of visually presented items. It is rare, however, for the contextual presentation of the stimuli to be varied. Instead, decision options are viewed via static images on a computer monitor and are usually preceded and followed by a blank screen (see Plassmann et al. (2007), Lebreton et al. (2009) and Hare et al. (2010) for examples). For these experimental findings to generalise to that of complex goal-orientated behaviour, the effects of dynamic context on complex decisions must be considered. Thus, to advance an idea proposed by Summerfield and Tsetsos (2012), new experimental paradigms should be designed in which value-based decisions are not treated as abstracted economic calculations devoid of context, nor perceptual responses devoid of reward.

1.4 Contents of the thesis

Chapter 2 reviews relevant concepts, methods and definitions used throughout the thesis and provides a theoretical introduction to the investigation. Temporal expectation, rhythm, periodicity, synchronisation, entrainment, and perceptual and value-based decision making are each discussed in turn to highlight how they have been defined and studied by others. The chapter then reviews existing cognitive models of sensory entrainment and decision making to understand how key models differ from one another and to highlight areas in which they may be compatible and linked. Here, attentional entrainment and sequential sampling models are focused on due to the each being largely uncontested in its respective literature. The chapter concludes by discussing the historic lack of communication between the timing and decision making literatures, highlighting why periodicity is important and the proposed benefits of collaboration.

Chapter 3 contains the main theoretical argument of the thesis suggesting that temporal expectation research must move beyond simple decisions in order to create experimental conditions that more closely reflect the demands of everyday decisions. It begins by discussing current limitations of temporal expectation research by focusing on five areas of temporal expectation paradigms that are in need of development. These areas are: decision complexity, decision type, time as an inherent dimension of targets, rhythmic variance and prior knowledge. A new experimental approach is then proposed as a way to fix these limitations. This includes incorporating complex averaging and subjective value decision making into temporal expectation paradigms, using sound lateralization techniques to combine rhythmic sequences with goal relevant information and using analytical methods to investigate how degrees of stimulus aperiodicity bias complex decisions. This new experimental approach is used to design each of the seven behavioural experiments that are reported in later chapters.

Chapter 4 reports two sound lateralization experiments (experiments 1 and 2) aimed at testing what effect rhythmic temporal expectations and prior knowledge about stimulus timing have on complex averaging decisions. The experiments represent the first attempts at implementing the new experimental approach described in chapter 3 and were adapted from three conceptually-similar behavioural studies in the decision making literature. Effects of rhythmic variability are investigated and a decision weighting analysis run to understand how rhythmic temporal expectations weight decision information at different moments within a stimulus stream.

Chapter 5 reports two additional sound lateralization experiments (experiments 3 and 4). The purpose of the experiments is to determine the generalisability of chapter 4 findings, whilst refining the implementation of the experimental approach outlined in chapter 3. This includes reducing the complexity of the experimental design compared with experiments 1 and 2 and varying the rate of rhythmic sequences used to induce temporal expectations. As experiments 1 and 2 both used the same rate of periodic sequence it is important to determine whether the rate of the stimulus stream interacts with the effects of rhythmic temporal expectation on choice. Whilst the experimental task required complex averaging, as in earlier experiments, participants needed to make relative rather than absolute localisation decisions. The complexity of the task was also better controlled for compared with experiments 1 and 2 as a way to enhance the interpretability of the findings.

Chapter 6 reports a speeded response time paradigm that focuses on the effects of rhythmic variability on motor preparation and response. As participants' choices were previously found to be sensitive to the degree of aperiodicity in the rhythmic stimulus (experiments 1 - 4), the experiment tested whether this sensitivity was specific to complex decisions or applies to simpler perceptual tasks. This will help to identify the underlying cognitive processes that were responsible for this earlier finding and to determine whether or not it should be incorporated into cognitive models of complex decision making. A related topic that is also investigated is whether the duration of the rhythmic sequence interacts with participants' sensitivity towards IOI variance. This was not tested in experiments 1 - 4 and therefore it is unknown whether the amount of rhythmic information one is exposed to contributes to the effect. Due to the aims of the experiment, the experimental design was much simpler than the previous complex averaging tasks and it did not attempt to fulfil all of the features of the new experimental approach described in chapter 3.

Chapter 7 returns to complex averaging decisions by reporting an experiment that investigates how rhythmic temporal expectations bias decisions when participants choose how much of a stimulus to listen to. The main difference with the previous complex averaging experiments is that participants did not have to wait until the end of a stimulus stream before responding. This tailors the investigation towards behaviours in which deliberation time can be costly, whilst making it possible to dissociate the time it takes to reach a decision from responses that are restricted to a response period. Drift diffusion models (a form of computational model widely used throughout the perceptual decision making literature) were fitted to the data to estimate how different components of the decision vary under different levels of periodicity. This use of cross-disciplinary methods acts as an example of the additional insight that can be gained if the timing and decision making literatures collaborate more closely.

Chapter 8 directs the investigation towards the type of decision that is being made by testing whether rhythmic temporal expectations bias subjective value representations. This is the last unexplored feature of the new experimental approach described in chapter 3 but is key to expanding the generalisability of the thesis to everyday decision making. As most previous research into subjective value is associated with preference decisions made between items of food, the experiment was designed to include non-appetitive audio-visual stimuli. This was done in an attempt to increase the generalisability and ecological validity of value-based decision making research.

The concluding chapter ties the experimental findings together to form an overview of the effects of rhythmic temporal expectation on complex decision making. This is the first step towards building a predictive framework on which to quantitively map the underlying cognitive processes and should be used as a starting point for further experimentation and computational modelling. Real-world applications for the research are also discussed with an emphasis given to human-machine interface design and suggestions made for future experimentation.

Chapter 2

Background

2.1 Key concepts, methods and definitions

2.1.1 Temporal expectation

The term "temporal expectation" is used interchangeably with several others, including "(rhythmic) temporal predictions" (Arnal and Giraud, 2012; Nobre and Rohenkohl, 2014; Cravo et al., 2013), "predictive timing" (Arnal and Giraud, 2012), "temporal orienting" (Correa, 2010), "rhythmic expectations" (Zanto et al., 2006), "temporally adaptive predictions" (Arnal et al., 2014), "anticipatory biases" (Rohenkohl et al., 2012) and "sensory entrainment" (Sameiro-Barbosa and Geiser, 2016). They all refer to the ability to expect the time occurrence of an upcoming sensory event based on prior temporal regularity and cues in the sensory input. Temporal expectations have been widely shown to improve action preparation and execution and afford a range of behavioural benefits associated with perceptual processing in both the auditory and visual domains. For example, temporal expectations decrease response times (Woodrow, 1914; Niemi and Näätänen, 1981; Doherty et al., 2005; Stefanics et al., 2010; Rohenkohl and Nobre, 2011), improve event onset detection (Hickok et al., 2015; Mathewson et al., 2010, 2012), aid time-interval, pitch and timbre judgements (Klein and Jones, 1996; Large and Jones, 1999; Barnes and Jones, 2000; Jones et al., 2002, 2006), increase visual contrast sensitivity and gain (Rohenkohl et al., 2012; Cravo et al., 2013) and decrease threshold detection in noise (Lawrance et al., 2014). Processing gains associated with temporal expectations have also been shown to exert influence across modalities (Lange and Röder, 2006; Escoffier et al., 2010) and can arise during the multimodal perception and production of complex rhythms (Greatrex, 2011).

Although most experimental research into temporal expectations has been conducted over the last thirty years, the topic featured in early psychology texts dating back to the end of the nineteenth century (Bolton, 1894; Wundt, 1904; Woodrow, 1914; Nobre and Rohenkohl, 2014). For example, Woodrow (1914) manipulated the time between a warning signal and target (termed the foreperiod) to be either fixed to a specific time interval or allowed to vary unpredictably. Participants' reaction times to the onset of the target were fastest when the foreperiod was fixed and hence predictable and longest when it varied and hence unpredictable. Similarly, Newhall (1923) showed that the perceived brightness of a visual stimulus increased when it was presented in-time versus out-of-time with a preceding isochronous click. Niemi and Näätänen (1981) summarised these and later observations by writing that the accuracy of the timing process is inversely related to the participant's uncertainty about the time of occurrence of a response stimulus (henceforth time uncertainty). Since these early studies, researchers have focused on two conceptually distinct methods for manipulating time uncertainty. The first has been to use explicit temporal cues and learned knowledge to increase anticipation of a forthcoming event (Coull and Nobre, 1998). This approach is analogous to the spatial orientating of attention task proposed by Posner et al. (1980) and requires participants to learn explicit associations between events so that detecting one increases the anticipation that second will occur. The second has been to use rhythmic stimuli as a means of generating implicit temporal expectations that are not reliant on explicit cues. This thesis focuses on the latter.

While there are a few everyday instances of explicit temporal cues to an event which may not affect the individual directly, such as watching (but not otherwise responding to) traffic lights, the most common source of temporal expectation arises via engagement with dynamic and meaningful regularities in the environment. For example, in the social world, dancing in time with a partner requires tracking salient regularities in the sensory stream and using these to extrapolate a predictive pulse on which to coordinate future actions. These expectations have been referred to as "rhythmic expectations", "dynamic attending" and "predictive timing" and are mostly described as being defined by a sensory pattern's rate and or rhythm (Jones, 2010). This has led to the general consensus that rhythmic temporal expectations arise implicitly from passive engagement with dynamic events and are not reliant on cues or knowledge of fixed time intervals. The evidence supporting this claim is moderate but inconclusive. For example, Winkler et al. (2009) showed that rhythmic stimulation administered to sleeping newborn infants (37 - 40 weeks old) resulted in anticipatory brain responses occurring at metrically aligned moments in time. As the babies were asleep it is unlikely that they were conscious of the rhythmic stimulation and their age reduced the likelihood that expectations were caused by explicitly learned associations. More recently, Arnal et al. (2014) have claimed that rhythmic temporal expectations and learned predictions belong to biologically separate processes, each differentiable via the frequency of associated neural oscillations. Whilst theoretically attractive, evidence demonstrating that both cognitive processes do not interact, or even that a binary distinction between implicit and explicit expectations (or endogenous and exogenous processes) is possible, is lacking. Understanding the autonomy of rhythmic expectations and the degree to which they are separable from learned cues and temporal orienting is the topic of later chapters. The following paragraphs describe current approaches for investigating rhythmic temporal expectations.

Work conducted by Jones and colleagues since the mid-1970s built on observations of Helmhotlz and Bolton (Bolton, 1894) to provide a strong foundation underlying our current understanding of how rhythms modulate attention and perceptual excitability (Jones, 1976; Jones and Boltz, 1989; Klein and Jones, 1996; Large and Jones, 1999; Barnes and Jones, 2000; Jones et al., 2002; Jones, 2009). Jones focused on auditory pattern research and began by forming a dichotomy between two forms of attending. The first was termed "future-orientated" or "anticipatory" attending. This refers to the use of higher-order timing patterns, extrapolated from rhythmic events, to direct attentional energy towards future points in time. The second was termed "analytic" or "reactive" attending. This refers to the focusing of attention towards local, often adjacent, events and the constant reorienting of attention due to a lack of predictability in the environment. This distinction implies that the structural regularities of events are essential in regulating the degree, and type, of temporal attending deployed at each moment in time. For example, the reason why tracking a moving fly around a room is so difficult is that its movement is highly unpredictable and lacks rhythmic predictability. For this reason attention functions reactively in an attempt to organise unstructured information. This results in the fly always remaining one step ahead of tracking movements. Conversely, clapping in time with a beat, or even clapping on every fourth beat, is a very easy task for most people. This is because temporal regularity of the musical context provides a framework from which to direct future orientated attention for the purpose of motor preparation and synchronisation with external events. Temporal expectations are thus described by Jones as preceding perception in that they function to create a heightened receptive region for forthcoming events (Jones, 1976).

The experimental methods used by Jones and colleagues to investigate rhythmic temporal expectations share similarities with that of Newhall (1923). In a widely copied method, the temporal onset of the last tone in a stream of isochronous rhythmic precursors was systematically varied in relation to the preceding sequence (either early, on-time, or late) (Large and Jones, 1999; Barnes and Jones, 2000). The task was to determine whether a comparison time interval (heard after the stimulus) was the same, shorter or longer than the gap between the last two tones in the precursor sequence (the standard time interval). Participants' response accuracy degraded with increasing leading or lagging of the final tone and was most accurate when the final tone aligned with the isochronous sequence. Variants of this design were used to test the effect that the rate of the precursor, irregular rhythmic precursors and metrical relationships have on choice accuracy (Barnes and Jones, 2000). In addition to time interval judgements, pitch, timbre and metre perception was also investigated (Klein and Jones, 1996; Jones et al., 2002, 2006; Ellis and Jones, 2009). In a pitch comparison task, listeners judged whether the the pitches of two sine tones were the same or different from one another when separated by an isochronous acoustic sequence (Jones et al., 2002). The onset of the second tone was manipulated so that it either aligned with the isochronous sequence or was slightly early or late. The results showed that pitch discrimination was best when the second tone aligned with the isochronous sequence and degraded the further it was from the expected point in time. Common findings across a range of discrimination, detection and judgement tasks highlight the ubiquitous effects that temporal expectations have on boosting sensory processing and facilitating memory recall.

More recent neuroscientific studies have adopted similar experimental approaches to Newhall (1923) and Jones (2010) for investigating the cognitive and neural mechanisms responsible for temporal expectations. For example, Henry and Obleser (2012) used a gap detection task in which participants detected brief gaps of silence in a stream of frequency-modulated (FM) complex tones. The gaps were presented at different phases of the FM cycle and task performance was compared with the phase of entrained low-frequency neural oscillations in participants' electroencephalography data. Similarly, Hickok et al. (2015) tested whether an acoustic rhythmic sequence induced periodic fluctuations (more than one cycle) in perception that matched the period of a rhythmic precursor sequence. The stimulus comprised an amplitude-modulated noise sequence followed by a period of unmodulated noise in which a response tone, embedded in the noise, was presented at one of nine intervals in relation to the rhythmic sequence. Participants were required to press a button if they detected a tone in the noise and perception was measured using task accuracy. Although these and similar experiments (e.g. Mathewson et al., 2010, 2012; Lawrance et al., 2014; Escoffier et al., 2010) often improved the experimental designs used by Jones and colleagues (such as the use of continuous rhythmic stimulation which is more characteristic of natural rhythmic stimuli), they are all essentially asking the same question: What effect does a preceding isochronous pulse have on the detection and processing of an isolated and unrelated response stimulus?

One exception to this approach is a naturalistic behavioural paradigm used by Doherty et al. (2005), Correa and Nobre (2008) and Rohenkohl and Nobre (2011). The task involved tracking a moving ball across a computer screen before it disappeared behind an occluding barrier. When the ball reemerged from behind the barrier, participants were to decide whether or not it contained a black dot (it contained one 50% of the time). By manipulating the spatial trajectory, rate and rhythm of the ball before the occlusion, as well as the time and location in which it reemerged, the response target remained fully integrated with the temporal context of the scene. In other words, timing and associated expectations were not artificially separated from goal-relevant decision information.

As described earlier, the binary distinction between explicit temporal orientating and implicit rhythmic expectations, whilst useful, may prove to be limiting for future theoretical development. This is because future-orientated attending is an endogenous process that arises via exogenous exposure to regularities in the environment. Thus, what starts off as an implicit automatic anticipatory process will quickly contain explicit predictive features. This is because it will lead to the construction of an internal temporal template that generates predictive knowledge that can be used to inform goal-directed behaviour (Arnal and Giraud, 2012; Hermann et al., 2016). One way to frame this concept is by stating, as Jones does, that both guided imagination and event structure underlie metrical constructs (Jones, 2009, p. 87). A different way is to assume that, during interaction, implicit and explicit temporal expectations function in sustained bidirectional dialogue. The interdependence between both types of expectation ensures that predictable patterns are detected and extrapolated and predictive knowledge utilised so that attention and action function to create perceptual streams and maximize goals. This interpretation is consistent with arguments of Arnal and Giraud (2012) and Summerfield and de Lange (2014, p.745). That is, the perceptual system utilises information about stimulus frequency, conditional probability and temporal autocorrelation to build expectations about forthcoming stimuli whilst operating in a hierarchically dependent manner. The use of the term hierarchy refers to the relationship between forward and backward messages passing between surface and deep layers of the cortex that assess mismatch between goal-relevant prior statistically-based predictions (passed from deep to surface layers of the cortex) and information contained in the sensory input (passed from surface to deep layers of the cortex) at varying levels of abstraction. Rhythmic temporal expectations should thus be viewed as a component within a larger perceptual system which functions to periodically modulate the overall activity of the brain in a way that is minimally reliant on content, but maximally reliant on the temporal structure, of forthcoming information (Arnal and Giraud, 2012; Nobre and Rohenkohl, 2014; Jones, 2010).

2.1.2 Rhythm and periodicity

Rhythm can be used to describe, quantitatively measure and recall the temporal features of a wide range of dynamic events. It is defined here as a serial pattern of time intervals marked by sensory and/or motor events and is based on the interonset interval (IOI) of successive events (Chen et al., 2008; McAuley, 2010; London, 2012). Rhythms arise via patterned change in the body and environment, such as beating hearts and blinking eyes, and are often found within continuous streams of information (Glass, 2001). For this reason, applicable events are not restricted to those that contain discrete entities, such as a recurring electronically produced beep in which there is a clear and definite separation between each sub-event, but include any perceivable grouping that exist in continuous streams of information. For example, footsteps contain rhythmic information because the time at which each foot hits the ground creates time intervals forming a regular pattern that can be extrapolated and responded to. The act of walking is continuous and therefore rhythmic measurement relies on being able to detect and classify sub-elements (individual steps) within the larger event (walking).

Contrary to common uses of the term in music and some forms of speech that focus on describing the experience of metrical hierarchies, rhythm is not synonymous with periodicity. Footsteps can speed-up, slow down and generally deviate from regularity and still remain rhythmic in nature. Thus rhythm can be described by the variance of local time intervals, the duration of rhythmic patterns as well as other statistics describing the durations between sub-events. For example, a periodic sequence, which is a type of rhythmic sequence, contains zero IOI variance and has a constant IOI between sub-events. Viewing rhythm in these terms removes the need to distinguish periodicity from aperiodicity and assumes that all rhythmic sequences exist on a continuum of related time intervals embedded within a larger event. What is required, however, is a way of classifying which time intervals belong to the same rhythmic streams and are thus meaningfully related. For example, at a busy road crossing there are simultaneously occurring time intervals produced by different engines, different footsteps and unrelated conversations and gestures. This scene would typically be described as containing a number of separate rhythmic events rather than simply as being rhythmic. That is, the rhythmic relationship between the ticking sound of an engine and the gestures of someone waving are rarely perceived as meaningfully related, even though they are sub-events within a larger event (the road crossing at a specific point in time). Therefore, the definition of rhythm should be changed from that in the previous paragraph: Rhythm is a serial pattern of time intervals marked by sensory and/or motor events and is based on the regularity of the IOI of successive meaningfully-related events.

Although all rhythmic sequences can be described and experienced in terms of IOI variance, averages and absolute duration, periodic rhythms are traditionally characterised in terms of period and phase. Period (T) constitutes the base time unit of a periodic rhythm and is the time it takes for one complete cycle of an oscillation without repetition (Pikovsky et al., 2001). This unit is required when computing the frequency of the periodic rhythm: $f = 1/T$. Phase is a quantity that increases by 2π within each oscillatory cycle, proportional to the fraction of the period (Pikovsky et al., 2001). It can therefore be thought of as a quantity that moves in uniform circular motion during each periodic cycle (Strogatz et al., 1993). This means that the phase of an oscillation starts at 0°and increases linearly until it reaches 360°at which point it returns to zero for the start of the next oscillation. Both period and phase are useful measurements for quantitatively describing the relationship between two or more periodic events and can be used to determine the degree of synchrony that exists between them. For example, two separate events may share the same period but be out of phase with one another whilst still remaining in tight coupled synchrony. This concept is especially relevant for understanding models of neural oscillatory entrainment that are discussed in section 2.2.

Apart from controlled examples, such as pendulums, experimenter-generated beeps and metronome clicks, rhythm and periodicity in the natural world is usually a lot more variable. For example, a person's footsteps may be perceived as regular, yet the measurement of these time intervals would indicate that they are not strictly isochronous. Understanding what constitutes regularity and how this varies from isochronous time intervals, both in terms of the physical signal and perception of that signal, is important for the study of temporal expectations. This is because rhythmic stimuli used to investigate temporal expectation are often isochronous and the effects of rhythmic variability on perception are under-investigated. For this reason the discussion must turn away from definitions and methods for describing isolated serial patterns in time and focus on rhythmic interaction and perception. Specifically, how do humans and other species use rhythmic information to engage and coordinate successfully with dynamic events?

2.1.3 Synchronisation and entrainment

Synchronisation and entrainment are important concepts for understanding theories of rhythmic temporal expectation. This is due to the required alignment and coupling of anticipatory cognitive processes with externally generated rhythms and because the terms are widely used throughout the timing literature. This section provides definitions and examples which can be used as a foundation for understanding the cognitive models of expectation and choice that are described in section 2.2.

When two or more rhythmic processes co-occur and form a sustained non-random relationship with one another they are said to be in synchrony (Bernieri et al., 1988). Synchrony can arise within and between biological organisms, inanimate physical processes as well as between living organisms and inanimate processes. For example, two autonomous metronomes located in separate corners of a room would be described as being in synchrony if they are set to oscillate with the same period. This is because both have the same period and exist in the same phase space, regardless of the fact that they may be out of phase with one another. If a person were then to clap in time with one of the metronomes, their clapping would be described as synchronised with the metronome, as both the hand movements and the cognitive processes responsible are set to oscillate in time with a fixed rhythmic process (McGrath and Kelly, 1986). Synchronisation, therefore, is the act of adaptively coupling of one rhythm to another without the need for both sources to adapt to each other.

When mutual adaptation between multiple self-sustaining rhythms forms a shared synchronous state it is called entrainment. To cite Clayton et al. (2005), "entrainment describes a process whereby two rhythmic processes interact in such a way that they adjust towards and eventually "lock-in" to a common phase and/or periodicity." Clayton et al. (2005) and Pikovsky et al. (2001) point out that there are a number of criteria that must be met for entrainment to occur. Firstly, there needs to be more than one oscillatory process. Secondly, each oscillatory process must be autonomous. Autonomy in this context means that each rhythm must have its own energy source and be able to continue oscillating when not interacting. For example, the air molecules surrounding a vibrating tuning fork would not be entrained because they are fully dependent on the vibrations of the fork. Instead, their movements arise via resonance whereby they get all energy from an external source (van Noorden and Moelants, 1999; Large, 2008). Lastly, the oscillators must to some degree be coupled and be able to influence one another through interaction. This means that both oscillations must share a sustained connection which can be defined in terms of a common period and "lock-in" phase relationships. Coupling requires that each process is able to adapt its period and phase dynamically in response to systematic changes in the others behaviour. If this coupling is too strong then both processes lose their individuality and hence their autonomy. If coupling is too weak then a change in one would not affect the other and thus they not influence each other.

As entrainment does not constitute exact rhythmic imitation, knowing how to quantify "weak interaction" between coupled oscillators is problematic (Clayton et al., 2005). For example, as two people walk side-by-side it is common for the period and sometimes phase of their footsteps to entrain towards sustained synchrony. This is an automatic and largely unconscious phenomenon that is robust to random and non-random perturbations (Large and Jones, 1999). Therefore, if the difference in frequency $\Delta f = f_1 - f_2$ of two separate rhythmic processes decreases over time to be some small value, even after being perturbed, then the processes are entrained and the starting frequencies said to have existed within an "entrainment region" (Pikovsky et al., 2001; Jones, 2010). The entrainment region is thus the range between zero Δf and maximal Δf in which two oscillators will entrain. Perturbations can occur via small random deviations in the duration of an oscillatory cycle or via isolated and temporally deviant event onsets. This is called "frequency detuning" and can be tested by perturbing either of the two processes and observing whether Δf minimises. Additional characteristics of entrained systems to consider are: i) Rhythmic dominance whereby one oscillation acts as a stronger attractor than the other (Jones, 2010), ii) Phase and period correction in which out-of-time oscillations are reset to fall within an entrainment region (Large and Jones, 1999; McAuley, 2010), and iii) Phase locking whereby both processes entrain towards opposite positions of phase space (i.e. they exist 180°out of phase) (Clayton et al., 2005).

Entrainment is important to the investigations reported in this thesis because it provides mechanistic clues as to why rhythmic temporal expectations are induced by quasi-periodic regularities in the environment. Weak oscillatory coupling ensures that strict isochrony is not a prerequisite for entrainment and that percepts of regularity could emerge via the interaction of cognitive oscillations with sensory streams. I will return to these ideas in section 2.2 when reviewing current cognitive models of temporal expectations as well as when discussing the experimental findings in later chapters. Before doing so, however, thought must be given to how rhythmic expectations bias higher-order cognitive function and how this information shapes our decisions.

2.1.4 Decision making

Nearly all psychophysical tasks designed to investigate timing can be understood from the perspective of decision making. For example, Large and Jones (see section 2.2) built their dynamic attending theory on the observation that rhythmic expectation increased the accuracy of categorical judgments and the speed in which these could be made (Large and Jones, 1999; Jones, 2010). Likewise, Cravo et al. (2013) concluded that neural entrainment enhances contrast sensitivity due to the speed and accuracy with which participants could correctly decide on the angle of a visual grating. As the same behavioural metrics are used to measure both temporal expectations and the integration of noisy decision evidence during choice, temporal expectations should be conceptualised as facilitating decision making in some way (Smith and Ratcliff, 2004).

A broad but encompassing definition of a decision is that it is a "commitment to a mental state that prescribes a course of future action and decision making is the name given to the neural, cognitive and computational mechanisms by which such commitments are made" (Summerfield and Egner, 2014, p. 837). Others have described decisions as the "cognitive process of choosing an option or an action among a set of two or more alternatives" (Wang, 2008, p. 215), "the behaviour observed when individuals select one among many available options" (Padoa-Schioppa and Assad, 2006) and "a deliberate process that results in the commitment to a categorical proposition" (Gold and Shadlen, 2007, p. 536). Decision making is a widely discussed process and it is commonplace within many everyday utterances concerning healthcare, money and consumption to name just a few.

It is unclear whether the cognitive processes underlying perceptual judgment (such as the choices made in timing paradigms) are the same as those underlying complex goal-directed choice. Is the process of deciding whether a musical note is out of time any different from deciding in more economic terms whether a musical passage is any good and how one should respond? Whilst both processes are labeled as decisions and can often be interpreted as building on one another, these types of processes are mostly handled separately in the literature (Gold and Shadlen, 2007; Smith and Ratcliff, 2004; Summerfield and Tsetsos, 2012). If the same underlying mechanisms govern both, however, then the known effects of temporal expectation associated with perceptual decision making may have relevance to complex goal-directed choice.

Although perceptual and economic decisions are conceptually distinct, it seems increasingly unlikely that the brain processes involved in each will completely exclude the other. All decisions are claimed to rely on an accumulation of evidence relevant to the identification or evaluation of a sensory stimulus - termed a decision variable (DecV) - and all experiments in the literature require participants to identify a stimulus and then to select a response that will lead to reward (Summerfield and Tsetsos, 2012; Gold and Shadlen, 2007; Gold and Heekeren, 2014). This implies that there should be common theories capable of describing the neural circuits underlying both perceptual decision making (PDM) and economic decision making (EDM) (Krajbich et al., 2010; Krajbich and Rangel, 2011; Tsetsos et al., 2012). As argued by Summerfield and Tsetsos (2012), however, the conceptual distinction has strongly defined the field and there is a tendency not only to study one type of decision experimentally whilst disregarding the other, but to separate the fields of enquiry into distinct research programmes, and for the two fields to work within different theories and to use different experimental methods. Before discussing how common theories might interact with those of temporal expectation, both types of decision and their related experimental methods must be further defined.

PDM is the process by which humans quantitatively categorise sensory information into discrete classes and use these to perform appropriately in a given task or context (Gold and Shadlen, 2001; Summerfield and Egner, 2014). It is usually investigated by having participants classify whether a noisy stimulus belongs to one of two categories. Different task conditions are then achieved by varying i) the level of noise, ii) the probability that the stimulus belongs to a specific category and iii) the task instructions. For example, when measured in terms of reaction time and accuracy, choice behaviour varies depending on whether the task instructions emphasise speed or accuracy. Seen from this perspective it is clear that most perceptual experiments investigating temporal expectations are in fact investigating the effects of periodicity and stimulus timing on PDM. Whilst this statement may appear obvious it is important. This is because historically there has been very little dialogue between the temporal expectation and decision making literatures resulting in the entirely separate cognitive models being developed that should be relatable to one another.

PDM has been traditionally described via statistical concepts such as signal detection theory (SDT). SDT defines a method for measuring the criterion an observer uses when deciding on a signal's existence by calculating the sensitivity of the observer towards a signal and the log likelihood value at which the observer discriminates between two choices, normally referred to as observer bias (Green and Swets, 1966; Gold and Shadlen, 2007; Summerfield and Egner, 2014). A key limitation of SDT is that it neglects the role of time and prior context within the decision making process. As discussed in section 2.2, choice requires a dynamic accumulation of evidence over time and therefore current cognitive models have since advanced to account for temporality in choice.

To understand how temporal expectations affect complex choice, value, reinforcement and learned preferences must be considered. EDM addresses these topics and is concerned with understanding the processes that enable humans to choose between actions or goods on the basis of learned value (Kable and Glimcher, 2009; Chib et al., 2009; Hare et al., 2010). The literature is characterised by the general framework that human decision making comprises three distinct attributes: a multicomponent mechanism of valuation, which encodes and retrieves subjective value associated with decision alternatives (Rushworth et al., 2009; Lim et al., 2011; Hare et al., 2010; Plassmann et al., 2007); competitive choice, which takes value representation as input and then selects the option with the largest value (Glimcher, 2004; Kable and Glimcher, 2009; Wang, 2008); and reinforcement, which updates value associations depending on whether the expected outcome of a decision was violated or not (Rangel et al., 2008; Rushworth et al., 2009). This framework closely relates to the field of reinforcement learning, which aims to computationally automate goal-directed learning - see Sutton and Barto (1998) for an introduction. For need of a definition, subjective value is the anticipated future benefit of a considered outcome. A key aspect of the valuation component is that subjective value is encoded as a form of "common currency" or scale as a means of facilitating comparisons between incompatible options (Montague and Berns, 2002; Kable and Glimcher, 2009; Rushworth et al., 2009). Subjective value therefore adds an additional layer of complexity to the choice comparison because any DecV that is influenced by subjective value is not just reliant on sensory information about a fixed and present event, but must also represent predictions about the benefits of a yet-to-be-experienced outcome. One important question for this thesis is whether temporal expectations and the timing of how decision-relevant information is presented weight the construction of subjective value. In other words, are events more attractive and subjectively valuable when they occur at a predictable moment in time (or possibly, when they occur as an integral part of a wider stimulus context that is regularly structured, as opposed to irregularly structured), and if so, why?

Recent work has begun to show that subjective value is context-dependent and reliant to some degree on how decision information is presented to an observer. Factors such as the number of available options within a given location, the perceptual attributes of surrounding options, the time between making a decision and receiving a reward and the recency of past information are all known to modulate value and influence option selection (Tversky and Simonson, 1993; Green et al., 1994; Iyengar and Lepper, 2000; Tsujimoto and Sawaguchi, 2005; Kable and Glimcher, 2007; Louie and De Martino, 2013). This has been shown via visual experiments in which static and easily observable items (i.e. free from noise) are presented on a computer screen and participants are to decide which option they think will lead to the most beneficial outcome.

Visual selective attention has also been shown to weight subjective value judgements, in that the longer an item is looked at, the more likely it is to be chosen from an array of equally preferred options (Krajbich et al., 2010; Lim et al., 2011). Simply paying more attention to one decision alternative can therefore influence both the decision time and the item that is selected. The authors of these studies suggest that the findings are caused by the value of the unattended option being discounted relative to that of the fixated option and that activity in the prefrontal cortex, thought to be representative of a DecV, is proportional to the weighted difference between the values of the attended and unattended options. Although this example is restricted to the narrow context of a two alternative forced choice task, in which the decision options were pictures of food, it highlights that the investigation of goal-directed decision making needs to be expanded in order to account for selective attention. One way that this could be achieved is by presenting participants with decision options that contain within them dynamic properties.

This would make the decisions representative of the huge number of everyday decisions that involve tracking dynamically-changing information whilst making value judgments and decisions consequent upon them.

The limitations of using static visual images to test and map out the cognitive processes underlying EDM are obvious. Imagine a predator chasing prey. As soon as the chase begins the value of considered outcomes varies constantly depending on contextual features such as how much energy the predator has left, whether there are new additional dangers in the environment, how fast the prey is and whether there are now easier options for catching food. If the predator is hunting in a pack, the value of a considered action will also vary depending on the actions of other hunters and the degree to which the predator believes that they will co-operate fairly throughout the hunt (Skyrms, 2001; Bech and Garratt, 2003). The computations involved are therefore far more complex than those responsible for selecting between static images because they must account for the uncertain and changing nature of the environment and the future actions of co-actors. Selective attention and anticipatory attending are therefore essential processes responsible for keeping track of what options are available at any moment in time and what the likely outcomes might be. Goals, value and risk must be understood as having some form of dependence on the flow of sensory information and dynamic anticipatory expectations.

2.2 Cognitive models of temporal expectation and decision making

A central aim of this thesis is to determine whether and how the timing of decision information systematically biases complex choice. To achieve this it is important to review up-to-date cognitive models that are influential in the temporal expectation and decision making literatures. This section therefore aims to identify similarities between influential models and to establish a theoretical foundation on which new, more holistic (i.e. complete), models can be built. This will inform the interpretation of experimental findings and underlie the predictive framework described in chapter 9. The section describes two classes of models: 1. Attentional entrainment models (AEMs). 2. Sequential sampling models (SSMs). These have been chosen because each class of model has substantial empirical support and remains largely uncontested in its respective literature. AEMs include the theoretical and biological realisation of Jones and colleagues' dynamic attending theory

(DAT). They provide a formal account of how driving environmental rhythms dynamically capture and entrain attentional processes in the human brain. SSMs describe a class of decision-theoretic models whose basic premise is that decision evidence accumulates dynamically over time towards a response boundary at which point a choice is made. As both AEMs and SSMs have relevance to a number of subfields, several research groups have contributed to the development of each type of model.

2.2.1 Attentional entrainment models

DAT provides a theoretical backbone to AEMs. It was first proposed by Boltz and Jones (Jones and Boltz, 1989) and developed by Large, McAuley and Jones in two highly original doctoral theses and a follow-up publication (Large, 1994; McAuley, 1995; Large and Jones, 1999). DAT is a mathematical model of attentional dynamics. It proposes that temporal expectations arise due to biological limit-cycle oscillations synchronising to regular timing patterns in the environment. This synchronisation is unidirectional in that it refers to the adaptive alignment of one rhythm (an attending biological oscillator) to the other (an environmental rhythm), but the end result is bidirectional entrainment between multiple internal biological oscillators.

DAT states that temporal expectations are self-sustaining periodic processes carried out by a biological oscillation. When coupled to an external rhythm they automatically synchronise resulting in a stable connection that is robust to small perturbations. As discussed in section 2.1.3, coupling of an attending oscillator arises if its intrinsic resonant period falls within an entrainment region of the driving rhythm. Attending rhythms remain adaptive to changes in the external rhythm by adapting their period and phase in response to rate changes that may occur. The driving rhythm can therefore be thought of as causing changes in attending oscillators' behaviour that move the entrained system through various states in phase space towards the attractor state of synchrony (Jones, 2010).

Rather than attention being conceptualised as occurring at an isolated point in time within a limit-cycle (a closed trajectory reflecting traversal through states of period and phase), it is described as a concentration of energy within a fluctuating pulse or energy stream. It is modelled as a repeating periodic probability density function, with time on the x-axis and energy on the y-axis, which contributes an expectancy region about the mode where attentional energy is non-zero (Large and Jones, 1999). The pulse is defined by its phase location within the limit cycle, termed its locus ϕ, and its focus k (the concentration of attentional energy within each cycle). As k increases, due to greater synchronisation between the entrained rhythms, the pulse narrows so that all concentrated attentional energy is near the mode. This results in more narrowed probability density functions and hence focused temporal expectations at specific moments within the entrainment cycle. As k decreases, due to decreased synchronisation and/or perturbations, the pulse widens to reflect greater rhythmic uncertainty. When $k = 0$ the pulse function is flat indicating a uniform distribution of attentional energy (Large and Jones, 1999). This occurs when entrainment between the two systems is not possible. Attentional focus therefore reflects the accumulated effect of expectancy violations rather than sequence variability and is conceptualised as a modulated oscillatory pulse that is strongest at metrically aligned moments in time (Large and Jones, 1999; Jones, 2009, 2010).

In addition to single oscillatory coupling DAT assumes that people simultaneously use multiple attending oscillators to track quasiperiodic components in complex environmental rhythms (Large and Jones, 1999). This arises via the entrainment of metrically-related attending oscillators to the synchronised attentional pulse. For example, imagine you hear a ticking metronome with an IOI of 400 ms. First, an internal attending oscillator whose intrinsic frequency falls within the entrainment region will synchronise to the external rhythm. This synchronisation will cause other metrically-related internal oscillators from higher and lower time spans (e.g. with IOIs of 200 and 800 ms) to entrain (Jones, 2009). In other words, those internal oscillators whose frequencies are harmonically related to the 400 ms IOI will adapt their period and phase to the synchronised 400 ms oscillator. This results in sustained bidirectional entrainment emerging between multiple attending oscillators at different frequencies. Over time these persisting interrelationships lead to "metrical clustering" and the percept of metre which expresses structural information about the metronome. A metric cluster "comprises sets of co-occurring oscillations with interrelationships that persist due to acquired internal bindings" (Jones, 2009). Internal bindings strengthen as a function of resonance among oscillator periods, phase coincidences and the durations of co-occurring oscillatory activity. Jones calls this aspect of the model the "metrical binding hypothesis" (Jones, 2009). Once acquired, metrical clusters enable observers to direct attentional energy towards metrical levels within this persisting metrical form that do not exist in the sensory stream. This is what allows one to hear a periodic rhythm but "feel" a faster or slower metrically-related pulse.

Whilst DAT elegantly describes the shaping of attentional energy it does not adequately define attention nor describe how it might fit within a larger decision framework. This ambiguity is problematic for further development as it leads to a number of important questions. For example: When oscillators are entrained to an external stimulus, is the perceptual salience of the incoming information boosted when it is temporally expected, or suppressed when it is not expected, or both? If boosted, does this apply to all information within the perceptual stream, learned associations, or just goal-relevant features? If boosting does not apply to goal specific information, what is the evolutionary purpose for this system and what benefit does it afford? Do attentional oscillators function at multiple stages of information processing? How does prior knowledge interact with driven oscillations, and can attentional entrainment be voluntarily suppressed? Many of these questions do not yet have clear answers because only a small set of paradigms has been used to test DAT. Knowing what aspects of a signal become enhanced or suppressed and how this relates to action representation and selection is key to both this investigation and the wider investigation of DAT. One way to start understanding some of these questions is to consider AEMs that have been developed within the neuroscience community.

The idea that oscillatory brain activity can entrain or be directed towards external rhythms is supported by a large number of neurophysiological studies that show neuronal oscillatory activity as a possible mechanism for timing in the brain. Engel et al. (2001) and Buzsáki and Draguhn (2004) provide early reviews. Regularities such as the rate of microsaccades (Bosman et al., 2009; Nobre and Rohenkohl, 2014), periodic fluctuations in selective attention (Busch and VanRullen, 2010) and phenomena such as the preferred tempo of a musical pulse (van Noorden and Moelants, 1999) have all been shown to occur at rates similar to rates of rhythmic brain activity. This indicates that neuronal oscillations may function to regulate both perceptual processing and timed behaviour.

An early hypothesis for describing the role of neuronal oscillatory networks was that they act as a partially independent context that affects the processing of the content of sensory information (Buzsáki and Chrobak, 1995). This led to the idea that rhythmic fluctuations in large neuronal populations create "windows of depolarisation" in which arriving inputs are more influential than during hyperpolarised states (Engel et al., 2001). The phase-specific gating of information and temporal alignment of neuronal discharges were thus thought to facilitate the transmission of sensory information from lower to higher levels of a cognitive hierarchy (Engel et al., 2001; Jones, 2010; Calderone et al., 2014; Ng et al., 2012;

Henry and Obleser, 2012; Rohenkohl et al., 2012; Cravo et al., 2013). Schroeder and colleagues provide a range of biological evidence for this activity in the primary sensory cortex. They showed that the phase of entrained low-frequency neural oscillations in the delta range (1 - 4 Hz) not only systematically amplified response gain and reduced reaction times, but determined momentary power in higher frequency oscillatory activity (Lakatos et al., 2007, 2008; Schroeder and Lakatos, 2009; Stefanics et al., 2010). This finding suggests that phase entrainment of large neuronal populations occurs during early sensory processing and acts to filter all perceptual information depending on its temporal predictability.

AEMs are also thought to be influenced by conscious "top-down" commands which are transmitted via the back propagation of high-frequency oscillatory activity. For example, Iversen et al. (2009) investigated whether individuals can alter their oscillatory brain responses at will. Participants listened to short repeating auditory rhythm and were instructed to imagine a beat to occur either on the first or second tone of the sequence. By omitting either tone the authors could determine whether there was a beat in the brain activity that was not present in the stimulus. On omitted tones, they found beat-like activity in beta-band (defined as 15 - 30 Hz) oscillations, but not in faster gamma-band (defined as 30 - 100 Hz) activity. This contrasted with a control condition in which neither tone was omitted and beat like activity was found in both beta and gamma-band activity. These findings highlight that rhythmic perception and oscillatory activity can be driven by endogenous processes and that beta-band oscillations play an important role in this process. Arnal and Giraud (2012) have recently developed this idea by suggesting that different bands of oscillatory activity have different filtering and message-passing functions. The authors claim that temporal expectations arise as the result of entrained slow cortical oscillations to the period and phase of driving environmental rhythms. This entrainment modulates the overall activity of sensory cortices and facilitates sensory processing regardless of the informational content of forthcoming information. This process is complemented by "top-down" expectations about informational content which are communicated via back-propagated beta-band activity and the validity of these expectations forward propagated through the cognitive hierarchy via faster gamma-band oscillations (Arnal and Giraud, 2012).

2.2.2 Sequential sampling models

SSMs are generally referred to as describing a process of either "sequential sampling" (Smith and Ratcliff, 2004; Gold and Heekeren, 2014), "serial sampling" (Summerfield and Tsetsos, 2012; Summerfield and Egner, 2014) or "sequential analysis" (Gold and Shadlen, 2007). These terms are used interchangeably and refer to processes that enable humans and other primates to make simple classification decisions based on streams of dynamically presented information. SSMs first originated over fifty years ago within mathematical statistics and are an extension of the sequential probability ratio test (SPRT). Given a fixed error rate, the SPRT uses the minimal amount of evidence required to form a categorical decision (Wald, 1947; Gold and Shadlen, 2007). This is achieved by updating, at each sampled moment in time, the log-likelihood ratio of the evidence given a two-alternative hypothesis until the decision signal exceeds a predefined criterion at which point the option with the greater likelihood is selected (Gold and Shadlen, 2007; Summerfield and de Lange, 2014). All SSMs follow this framework and are assumed to consist of three components: sensory encoding, evidence accumulation and a stopping criterion (Smith and Vickers, 1988). Sensory encoding acts as an intermediary between the sensory receptors and the decision stage. This encoding produces a DecV whose statistical features are a function of the stimulus being presented, which is then fed into the accumulator stage as evidence. Evidence accumulation determines how successive samples of evidence are accumulated and contains information about decision options. The stopping criterion determines at which point in time the accumulation is terminated and which response to make. This is thought to vary depending on the conditions and goals of the decision and to be represented biologically as option relevant neurons reaching a critical firing rate (Gold and Shadlen, 2007).

There are two broad classes of SSMs proposed throughout the decision making literature. Both agree that decisions are the result of an accumulation of noisy information over time but disagree on how evidence accumulation occurs (Smith and Ratcliff, 2004). The first are called "random-walk" or "diffusion" models which were first proposed by Laming (1968), Link and Heath (1975) and Ratcliff (1978). These assume that decision evidence during two-alternative choice tasks accumulates as a single total so that evidence in favour of one alternative is evidence against another. This is conceptualised as a random walk or continuous diffusion of a particle with Brownian motion moving between two stopping criteria (boundaries representing each option), from a starting location somewhere between either boundary. The drift of this motion is determined by momentary inclination towards one or the other option. The second class of models are "accumulation" or "counter" models, first proposed by Vickers (1970, 1979) and Smith and Vickers (1988). These assume that evidence favouring each option accumulates independently and that choice is a race between competing evidence totals. The selected response is determined by the first accumulator to reach criterion (Townsend and Ashby, 1983; Smith and Ratcliff, 2004). Finally, complex hybrids of both models have been proposed such as the leaky competing accumulator (LCA) (Usher and McClelland, 2001). The LCA assumes the accumulator framework but incorporates both a gradual decay in the accumulation process and lateral inhibition similar to the diffusion models between competing accumulators. The key area of contention between SSMs is therefore whether the input to the decision process contains absolute or relative decision information (Summerfield and Tsetsos, 2012).

This project concerns the intersection between SSMs and AEMs, and it is this intersection that is the focus of the rest of this section. Neural evidence suggests the intersection probably occurs between sensory encoding (AEMs) and the creation of a DecV during later central and motor stages (SSMs) (Summerfield and Tsetsos, 2012; Wyart et al., 2012; Cravo et al., 2013). Nevertheless, a clear understanding of a standard SSM is required so that a holistic framework can be proposed to account for the effects of temporal expectation on complex choice. For this reason, the following paragraph describes the Ratcliff Drift Diffusion model (DDM) (Ratcliff, 1978; Ratcliff and McKoon, 2008). The reason for this selection is that the DDM is supported by an extremely wide range of behavioural and neurobiological evidence spanning psychology, neuroscience and physiology. It has also been used to analyse data in chapter 7, so a clear understanding of its components will aid the comprehension of later chapters.

Figure 2.1 shows a schematic illustration of the DDM. The DDM distinguishes the quality of decision evidence entering the accumulation process from endogenously controlled decision criteria, as well as other non-decision processes, such as sensory encoding and response execution (Ratcliff and McKoon, 2008). The model is characterised by: 1. The distance between each response boundary and the starting point of the accumulation process (decision threshold a). 2. The rate at which decision evidence accumulates over time (drift rate v). 3. The starting point of the diffusion process relative to the two response boundaries (bias z as a proportion of boundary separation ($0 < z < 1$)). 4. The total duration of memory encoding and response output processes that occur before and after the deci-

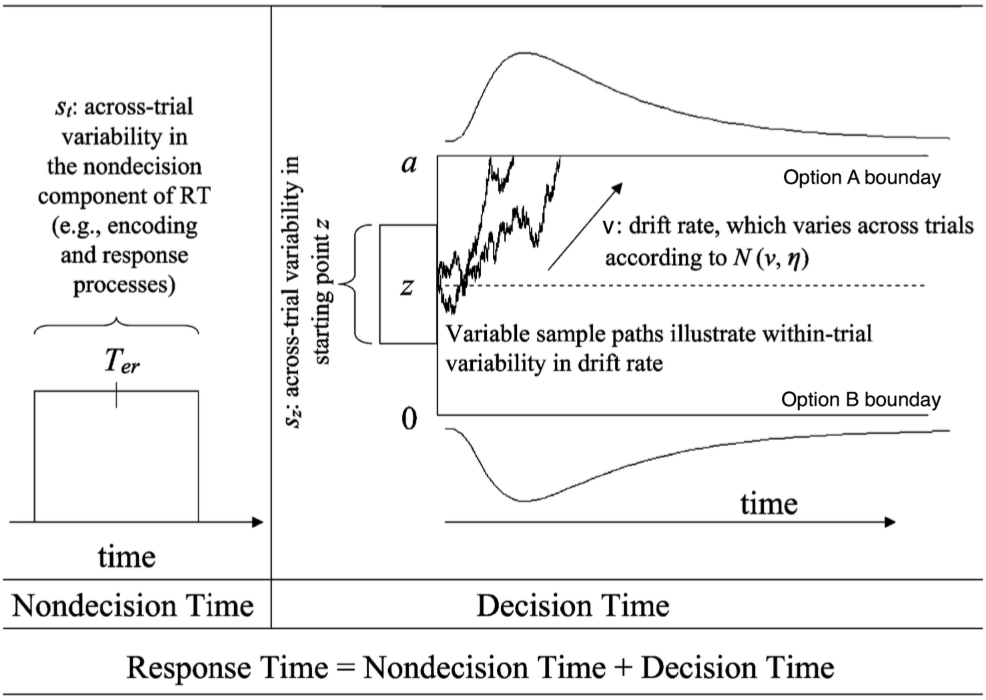


Fig. 2.1 A schematic illustration of the Drift Diffusion Model adapted from Wagenmakers (2009).

sion (non-decision time T_{er}). Variables a and z are assumed to be endogenously controlled and v to be a representation of the quality of the signal being accumulated. The model allows for trial-by-trial variability in some or all of the estimated parameters, in effect allowing them to vary across trials for each participant (represented by s_t, s_z and η in figure 2.1). This means that a process with the same drift rate or boundary condition will not necessarily produce the same decision outcome or latency on each trial. The DDM infers the behaviour of decision components by accounting for the shape and variance of response distributions and is able to describe common response biases. For example, the instruction to favour speed over accuracy is modelled as a decrease in a. This reduces the amount of accumulated information required to make a response. Because the accumulation process is modelled as a random walk and is thus prone to errors caused by randomness, participants make faster decisions but with more errors. Similarly, an increase in the signal-to-noise ratio enhances v whilst keeping the other decision variables constant: this describes situations when participants make both faster and more accurate decisions. To summarise, the DDM describes separate components of the decision process by accounting for the shape and variance of response distributions that cannot be computed using averaged response data alone.

As well as describing perceptual decisions, there is growing evidence to suggest that the DDM and other SSMs can be used to describe more complex decisions such as economic choice and value-based decision bias (Basten et al., 2010; Krajbich and Rangel, 2011; Gold and Heekeren, 2014). For example, Tsetsos et al. (2012) accurately predicted economic choice and biases using a leaky competing accumulator. Mimicking computational processes described above, the authors proposed that the process of value-based decision making consisted of an accumulation of subjective value (rather than perceptual evidence) associated with each choice alternative until a response threshold was reached (Tsetsos et al., 2012). Similar approaches have also been used to successfully model the effects of selective attention and eye gaze on value-based decision making (Armel et al., 2008; Krajbich et al., 2010; Krajbich and Rangel, 2011; Krajbich et al., 2012).

Like AEMs, SSMs have limitations. One area of concern is the assumption that evidence accumulation is a continuous process that occurs at a constant rate. As discussed in section 2.2.1, forward and backward message-passing between different levels of cognitive hierarchy is thought to arise via the rhythmic sampling of sensory information and neural oscillatory activity. If information sampling, processing and message-passing is rhythmic, then why is the accumulation of decision evidence continuous? How are rhythmic signals transformed into a constant without the loss of important information? Wyart and colleagues asked a similar question and argued that the human brain should exhibit slow rhythmic fluctuations in the rate of evidence accumulation during decision making (Wyart et al., 2012). To test this, the authors had participants classify whether, on average, the tilt of 8 sequentially presented Gabor patterns fell closer to either the cardinal or diagonal axes whilst recording electroencephalography activity. The data showed that perceptual information was first weighted according to the phase of on-going delta-band (defined as 1 - 4 Hz) oscillations overlying the parietal cortex and that lateralized beta-band (defined as 10 - 30 Hz) activity integrated the weighted information in an additive fashion (Wyart et al., 2012). In other words, neural activity that represented an accumulation of decision evidence was not constant and fluctuated rhythmically resulting in decision refractory periods that followed the processing of salient decision information. Importantly, both the rate of recorded delta and beta-band activity was separate from the intrinsic rate of the stimulus. This suggests that the neural encoding of perceptual and categorical decision information is dependent on endogenous neural oscillations and not just the entrainment of neural oscillations towards a rhythmic stimulus. Recent studies have found similar findings which together call into question the fundamental assumption of SSMs that decision evidence accumulates at a constant rate (Cheadle et al., 2014; Wyart et al., 2015; Spitzer et al., 2016). Cognitive mechanisms traditionally associated with AEMs may, therefore, have explanatory power for decision making.

2.3 A lack of communication between disciplines

To conclude this review it is important to ask why there has been a lack of communication between the temporal expectation and decision making literatures and to address arguments for bringing them together. This is required because the assumption that rhythmically generated temporal expectations bias complex decision making is not obviously correct. Relatively few events in nature are strictly isochronous, so why should periodicity and quasi-periodic structures have significance for perception and behaviour? Under what situations does periodicity afford behavioural benefit and what does this have to do with decision making? To begin addressing these questions this section reviews ways in which periodicity appears to be functionally important. It focuses on interpersonal coordination, communication and computational efficiency in an attempt to highlight relevance for complex decision making. Arguments are then made in support of future collaboration between the temporal expectation and decision making literatures and studies discussed that have already tried to bridge this gap.

2.3.1 Why is periodicity important?

Periodicity has relevance for complex decision making because it should affect how observed and anticipated information is valued. This is because there exist associations between periodicity and cognitive/social functions (described below) that can be used strategically to facilitate communication and survival. Understanding these strategies and how they relate to the goals and context of the decision maker will be key to determining how laboratory findings generalise to real world behaviours. A focus on timing will also ensure that decision making is not separated from context, but rather described as a process that emerges as a result of interacting with the world.

The most obvious benefit of periodicity is that, because it decreases temporal uncertainty, it can be used to facilitate interpersonal coordination between people or groups. Regular actions such as walking at a steady pace, making exaggerated movements in time to music, or rhythmically saying "three", "two", "one" out loud, allow others to predict with a high degree of certainty when and where you will be in the future. This reduction in time-space uncertainty forms structured and predictive information that others can use when planning how to move. It is especially helpful in the context of organised, goal-directed, group action, where multiple individuals must coordinate their behaviours to achieve a goal (such as lifting a heavy weight). The benefits of this ability are context dependent and only apply when the goal is to coordinate. Making exaggerated quasi-periodic movements whilst being tracked by a sniper will increase the likelihood that you are shot. If your goal is to engage in successful conversational turning taking, however, the use of predictive rhythmic speech patterns will make it easier to coordinate responses and increase the likelihood that the temporal flow of the conversation is maintained (Wilson and Wilson, 2005; Sebanz et al., 2006).

A related benefit is that periodicity can be used to aid the transfer of information between people. As described in section 2.2.1, periodic signals can induce in observers a persisting metrical form on which they can direct their attention towards upcoming, structurally salient, moments in time. Using this structure, one can strategically place information at predictable moments within this metrical hierarchy as a way of increasing the likelihood that it will be communicated efficiently. This is similar to what Clark (2005) describes as "material signalling", where actions are used to capture attention and strategically communicate commands and feedback during joint activity. A different way to view this material signalling is as an ostensive educational tool for prioritising the importance of information during communication (Csibra and Gergely, 2006). For example, a teacher may use predictable rhythmic movement to emphasise the importance of a piece of information to a class of students. This explicitly reduces temporal uncertainty in the interaction and as a result students should find it easier to detect, process and memorise important information that has been deliberately timed. The use of periodicity as a knowledge transfer tool may provide one reason for why humans who wish to successfully interact appear predisposed to unconsciously imitate each other's movements, gestures and timing (Clayton et al., 2005).

A third benefit is that periodicity can be used to increase cognitive computational efficiency by both reducing the amount of time the sensory system is required to sample the environment and aiding pattern recognition and completion. For example, if one predicts that goal relevant information is likely to occur in a periodic stream, perhaps via the detection of structure in the environment or an explicit cue, the sensory system need not sample the environment continuously.

Rather, it can use this knowledge to only sample at rhythmically-determined moments of anticipated interest. This strategy would help to maximise the amount of goal-relevant information that can be detected whilst minimising the amount of sampling that is required (Large and Jones, 1999; Henry and Herrmann, 2014). Similarly, entrained neural oscillators should help to actively "fill in" missing or incomplete information within the sensory stream (Velasco and Large, 2011). This is because entrained oscillators are self-sustaining and robust to small perturbations. As a result, higher dimensionality representations and groupings can be formed without the need of representing large amounts of low-level variance.

2.3.2 The benefits of collaboration

The most attractive aspect of increased connection between the temporal expectation and decision making literatures is that it will be relatively easy to achieve. This is because both fields use common psychophysical methods, build models based on similar behavioural measurements, and use common cognitive concepts. It is therefore likely that the potential rewards of collaboration will outweigh the risk of slower progress within each field. Rohenkohl et al. (2012) and Cravo et al. (2013) provide a good example of this. After running a rhythmic temporal expectation experiment they fit a DDM (see section 2.2.2) to the behavioural data to determine what effect rhythmic temporal expectations had on decision model parameters. They showed that the drift rate was higher in the periodic versus aperiodic conditions and that timing did not affect other components of the model. This suggests that rhythmic predictability enhanced the quality of perceptual information entering the decision stage, but did not change the decision criterion, nor other non-decision related processes. Although there were limitations to their experimental design which I address in chapter 3, this cross-topic analysis increases the accessibility of the topic for researchers in both fields. The next step is to design a battery of timing experiments that move beyond simple perceptual decisions and address complex decision making.

Chapter 3

Beyond simple decisions

3.1 Overview

Chapter 3 moves beyond simple decisions by describing a new experimental approach for investigating the effects of stimulus timing and temporal expectation on decisions that are more complex than those typically studied in the timing literature. This approach focuses on ways to test the generalisability of current timing theories under experimental conditions that more closely reflect the demands of everyday decisions. The chapter begins by addressing areas of existing temporal expectation paradigms that are in need of development. This includes the complexity of the decision, the types of decision being tested, the separation of goal-relevant decision information from properties of rhythmic precursors, rhythmic variability and prior knowledge about the timing of a stimulus and the purpose of the experiment. The experimental approach is then outlined and split into three components. The first expands the complexity of the decision via the use of complex averaging and valuation tasks. The second incorporates stimulus timing within goal-relevant decision information on each trial. This is described from the perspective of a sound localisation paradigm in which the timing of spatially lateralized auditory sequences are systematically varied. The third tests multiple degrees of rhythmic variability to understand what effect IOI variance has on complex decision making. The chapter concludes by discussing the benefits of the approach.

3.2 Areas in need of development

As discussed in sections 1.3 and 2.1.1, most perceptual studies investigating temporal expectation are restricted to a relatively small number of experimental contexts and tasks. As a result they nearly all investigate, to some degree, the effect that a preceding isochronous pulse has on the detection and processing of an isolated and unrelated response stimulus. This temporal invariance somewhat undermines the generalisability of the results of a large number of studies and hence the generality and even validity of the complex cognitive models that have been proposed to explain the data coming from these experiments. To determine whether these theories generalise to everyday behaviour and more complex goal-directed decisions, new experiments must be designed that explicitly deal with temporal variability and varying task demands. To design such studies requires understanding current limitations and identifying aspects of temporal expectation paradigms in need of development. The following sections describe five of these areas.

3.2.1 Decision complexity

Compared with the complexity of decisions one makes whilst walking across a busy square, the decisions used to investigate the effects of rhythmic temporal expectations on perception are highly simplistic. The most basic of these requires a single speeded response near the onset of an isolated target that is preceded by, or embedded within, a rhythmic sequence. Here, participants are to press a response button if they detect a target which takes the form of brief gaps of silence within a continuous auditory stream (Lange, 2009; Henry and Obleser, 2012), auditory tones and visual dots (Doherty et al., 2005; Stefanics et al., 2010; Mathewson et al., 2010, 2012; Miller et al., 2012; Breska and Deouell, 2014).

The complexity of the decision has been increased by some researchers by using two-alternative forced-choice tasks that require participants to make a binary classification decision on each trial. This requires either classifying the existence or absence of a target stimulus (Hickok et al., 2015), or making decisions based on stimulus attributes such as whether a target is a complex or pure tone (Rimmele et al., 2011), brighter or louder than the preceding rhythmic context (Marchant and Driver, 2012; Geiser et al., 2012; Herrmann et al., 2016), or has a different pitch or timbre in relation to its context (Klein and Jones, 1996; Morillon et al., 2016). The fact that these decisions are based on a single event and not multiple pieces of information, such as the trajectory, rate and actions of people walking in a square, represents an abstraction—indeed, a retreat—from real-world complexity that needs to be addressed.

Two ways to address this limitation is to impose memory constraints by having participants decide between two or more pieces of memorised information or to have them multitask. As decisions are often based on recalled rather than immediately available perceptual information this should increase both the complexity and ecological validity of the experimental task. This idea has already featured in some timing paradigms in which participants either compared a target stimulus with a previously heard standard (Large and Jones, 1999; Jones et al., 2002), engaged in a separate working memory task (Curtanda et al., 2015; de la Rosa et al., 2012) or decided which of two previously heard auditory sequences contained a deviant tone (Large and Jones, 1999; Lawrance et al., 2014). Whilst these studies are beginning to move towards the domain of real-world decisions by increasing the complexity of the decision, they are still limited by their reliance on isolated stimuli or the comparison of isolated events.

A different way in which they can be improved is by having participants make decisions based on sequences of information. This would not only demand the use of memory, but also require participants to form ensemble representations (i.e. averages) on which to base their decisions. Morillon et al. (2014) provides the only published example of this approach being used in the temporal expectation literature. That study was published midway through the research leading to this thesis, after the method had been adopted for the present research. It comprised a pitch comparison task in which participants decided whether the average pitch of a rhythmic sequence of frequency modulated pure tones was higher or lower than a reference pitch. Whilst the details of the task contained limitations that are discussed later in this chapter, the experimenters use of complex averaging successfully increased the complexity of the decision under investigation.

Averaging is a useful method for refocusing the research question away from how stimulus timing affects single event classification to how it affects the construction and representation of complex decision variables (DecVs). This is especially important for this investigation because averaging requires using structure and redundancy to form compressed and efficient representations of information that can be used to describe objects, assign value and compare complex options (Alvarez, 2011) - see section 3.3.1 for further discussion. Whilst there has been relatively little research into which neural mechanisms are responsible for differing types of averages (such as spatial or temporal), there is evidence that the parietal cortex plays an important role during numerical averaging and that value representations associated with complex decision options arise in the prefrontal cortex prior to choice (Piazza and Veronique, 2009; Kable and Glimcher, 2009). As the latter applies to scenarios such as charitable decision making in which decision makers must incorporate information associated with a range of different factors such as ethics and wealth (Hare et al., 2010), averaging appears to be a high order cognitive function that is closely related to memory and appears fundamental to the construction of most complex DecVs. Whilst averaging itself does not capture the complexity of decisions such as who to hire, where to invest money or who to marry, its strong relationship to complex decision making makes it an ecologically valid and scalable method for increasing the complexity of tasks in temporal expectation paradigms.

3.2.2 Decision type

All experimental paradigms used to test rhythmic temporal expectation require participants to either detect, classify or compare perceptual features of existing objects. This means that there is always an objectively correct or incorrect answer to the decision as it is referring to something that is already present in the world. For many everyday decisions, however, there is no objectively correct answer. Imagine that you must decide whether to walk to the left or right side of someone walking towards you, or choose which type of coffee to order in a coffee shop. These decisions are based on expectations about the future state of the world and are therefore probabilistic inferences that must account for preferences, goals and prior knowledge. This does not mean that they do not rely on existing perceptual features within the scene, but simply that they require the integration of attributions or representations of value and uncertainty predictions.

As there is increasing evidence to suggest that subjective value representations are weighted by selective attention (Armel et al., 2008; Krajbich et al., 2010; Krajbich and Rangel, 2011; Krajbich et al., 2012), it is important to test whether rhythmic temporal expectations bias other types of decisions associated with subjective value and risk. This is because within many research programmes temporal attention is claimed to function as a dimension of selective attending (Jones, 1976, 2010; Nobre et al., 2007; Nobre and Rohenkohl, 2014; Henry and Herrmann, 2014). These tests would require a new experimental approach in which participants make value-based decisions based on rhythmically varying sensory information. Any positive effect of temporal expectation would force theorists to expand existing cognitive models to describe not only how rhythmic temporal expectations bias low-level sensory processing, but what interaction this has with higher-order cognitive and motor functions that underlie DecVs.

3.2.3 Time as an inherent dimension of targets

Temporal expectation arises during implicit exposure towards rhythmically predictable regularities in the environment. For this reason it is common for experiments studying its effects on perception to consist of trials in which an entrainment period precedes the presentation of a target stimulus. An entrainment period consists of a sequence of at least three rhythmic precursors presented in either the auditory or visual domain. These precursors can be presented isochronously, when stable temporal expectations are desired, or irregularly, as a comparison condition. A review of twenty experiments in the literature shows that the mean number of rhythmic precursors was 7 with a standard deviation of 4.04. The only function of rhythmic precursors in the literature so far is to induce rhythmically fluctuating temporal expectations by the time the target stimulus is presented.

A major problem with the above approach is that it allows for rhythmic context to be artificially separated from goal relevant information on which decisions are made. This separation is sometimes made explicit for participants by telling them to ignore the rhythmic precursors (Jones et al., 2002; Escoffier et al., 2010; de la Rosa et al., 2012; Bolger et al., 2013). This approach is problematic because in everyday life timing is inherent within and between meaningful events and therefore it is not a dimension that is easily separable from context or the task at hand (Jones, 1990). Therefore, in addition to asking whether a concurrent working memory task influences the effects of rhythmic precursors on response times (Curtanda et al., 2015; de la Rosa et al., 2012), a useful secondary question is whether the temporal presentation of the working memory task stimuli impacts working memory performance. It would be an alien world in which short sequences of isochronous clicks accounted for timing and preceded each spoken utterance or change in facial expression. Similarly, goal-relevant information is unlikely to be devoid of timing and stored in a format that is entirely separate from the changing structure of the environment. A small number of studies have started to address the issue of rhythmic contexts being separated from response targets by trying to make experimental tasks more representative of natural rhythmic stimuli. For example, Henry and Obleser (2012), Rohenkohl et al. (2012) and Cravo et al. (2013) presented multiple targets at random intervals within a long stream of rhythmic precursors and Doherty et al. (2005), Correa and Nobre (2008) and Rohenkohl and Nobre (2011) overlaid targets with spatial information (see section 2.1.1). As these attempts still separate rhythmic information from goal-directed decision information, much more needs to be done to fix the problem.

A different way to consider timing is as a dimension of goal relevant events that functions to weight the relevance of feature information in some way. From this perspective it is clear that timing should not be separated from goal relevant decision information and that a solution is for experimental targets to contain dynamic properties. Were this the case, rhythmic precursors and separated entrainment periods would not be required.

3.2.4 Rhythmic variance

It is common for experiments in the timing literature to contain both periodic and aperiodic rhythmic conditions. Aperiodic trials typically include a temporally jittered precursor sequence (i.e. with IOIs that are probably different from one another) that functions as a comparison for periodic trials in which precursor stimuli occur at identical temporal intervals. The common finding is that, compared with periodic trials, responses on aperiodic trials are slower, less accurate and are not accompanied by the entrainment of ongoing neural brain activity. Whilst this categorical distinction between periodic and aperiodic is a good starting point, more needs to be done to understand why this is the case, particularly as natural stimuli such as speech and music, to which much of the previous literature on timing claims relevance, are only very rarely strictly periodic. This treatment of timing either periodic or not leaves unexplored questions such as: do varying degrees of aperiodicity and IOI variability bias choice in different ways? Are decision latencies, accuracy and momentary power in oscillatory brain activity linearly dependent on IOI variance between precursors? Is the influence of increasing aperiodicity linearly related to decision quality, or is the relationship quantal? These questions can be answered by paying more attention to the generation and analysis of aperiodic precursors and their impact on choice, which will be key to mapping out the computational processes that underlie rhythmic sampling and its impact on complex choice.

Mathewson et al. (2012) and Herrmann et al. (2016) are, to the author's knowledge, the only studies to have investigated how degrees of rhythmic variability, and not predictability (as in Morillon et al. (2016)), affect perceptual decision making. During data processing, Mathewson et al. (2012) binned all aperiodic trials into low and high variability groups based on the IOI variance in each precursor sequence. By analysing each variability group separately, they showed that high variability trials had a lower target detection rate than both the low variability and strictly periodic trials, thus demonstrating that perceptual decisions can be sensitive to varying degrees of IOI variability in a precursor sequence, i.e. to degree of aperiodicity. Herrmann et al. (2016) provides a more in-depth investigation. They measured whether temporal expectations arise in temporally variable tone sequences and to what degree these affect perceptual processing. Each trial consisted of a rhythmic sequence of 25 irregularly presented tones and contained between 2 to 4 randomly positioned target tones. Target tones were identical to other tones in the sequence except that they were slightly louder and the task was to detect the targets. Repeated sampling ensured that sequence IOIs had an average frequency of 2 Hz (500 ms) and were randomly sampled from a range ±55 ms either side of the average. This meant that some parts of each sequence were more irregular than others and that overall degree of irregularity varied trial-by-trial. After the experiment, a mathematical oscillator model developed by Large and Jones (1999) was fitted to the data to determine what the instantaneous phase of attending oscillators should have been at the time of each target tone. The analysis showed that, regardless of the degree of aperiodicity of the context, participants' hit rate improved as a function of the strength of the entrained mathematical oscillator. In other words, the more predictable the aperiodic context became, even though it was not completely periodic, the more likely participants were to detect a target tone.

As humans are sensitive to rhythmic variability, and, as noted above, both speech and music rhythm is rarely strictly periodic, it makes little sense to continue addressing periodicity and aperiodicity as simply categorically distinct. Rather, effort must be made to understand how humans respond to varying degrees of aperiodicity and how this is accounted for in models of temporal expectation and decision making. A simple way to do this, as Mathewson et al. (2012) did, is to bin randomly generated aperiodic sequences into subgroups according to IOI variance and to then analyse dependent variables by each subgroup of variability.

3.2.5 Prior knowledge

Imagine that your goal is to catch a tennis ball. As it moves through the air you focus on what you need to do to catch the ball. This involves tracking its trajectory and preparing your body to catch the ball based on predictions concerning where it will land. Although timing is key to the task, it is no more (nor less) important than other sources of information such as visuo-spatial information and motoric control. It therefore remains an integrated dimension of a dynamic visuo-motor spatial-coordination task and not the explicit focus of attention. This contrasts with many temporal expectation paradigms in which it is obvious to the participant that the primary purpose of the experiment is to investigate timing. For example, if each trial begins with a rhythmic sequence of repeating stimuli that have no apparently meaningful relationship to the task, participants seem likely to infer that they are structurally important to the experiment, and therefore probably warrant explicit attention. Rhythmic precursors that fall into this category are streams of repeating flashes and tones (Large and Jones, 1999; Barnes and Jones, 2000; Mathewson et al., 2012; Marchant and Driver, 2012; Miller et al., 2012; Herrmann et al., 2016), frequency modulated complex tones (Henry and Obleser, 2012) and amplitude modulated noise (Hickok et al., 2015). The problem with this approach is that participants will likely perform differently on the task if they assume the rhythmic precursors to be functionally important to the experiment. As a result, their behaviour may have limited relevance to real-world scenarios in which timing is an implicit dimension of a task.

The more abstract the relationship between rhythmic precursors and the task, the greater the above problem may be. Escoffier et al. (2010) provides perhaps the best example of this. Participants were shown sequences of images of different faces and houses and told to respond as fast and accurately as possible whether each image was presented upright or inverted. To test whether acoustically induced temporal expectations enhanced task accuracy, participants heard a rock drum beat over headphones whilst they did the task that either coincided with the onset of the images or did not. For participants' doing the task, it is a possibility that due to the rock beat not being related to the images, or accompanied by other instruments, they may have inferred that its function was to induce temporal expectations. A small number of studies *have* successfully avoided this problem by using tasks in which different types of feature information are overlaid onto the rhythmic precursors. For example, studies have combined different colours (Breska and Deouell, 2014), spatial locations (Doherty et al., 2005), and pitches (Morillon et al., 2014) with (visual or auditory) precursor sequences as a way of reducing dissociation between targets and task timing.

As rhythmic temporal expectations are thought to be largely implicit and driven by the structure of rhythmically predictable events, there have been few direct attempts to understand what effect prior knowledge about the timing of a rhythmic sequence has on task performance. This is important for the current project because complex decision making relies on predictions about future states of the world and the integration of learned knowledge with sensory information. One way that the effect of prior knowledge can be investigated is to compare responses to trials in which subjects know whether stimuli will be periodic or not, with those where they do not know beforehand. Jones et al. (2006) used a similar method to test the differences between voluntary and automatic attending. In their third experiment they explicitly told participants about the timing of the target tone and ways in which it could be manipulated prior to them starting the experiment. Task performance was then compared with a previous study (experiment 2) in which participants were not told about the stimulus timing. The analysis showed that prior knowledge about the timing of rhythmic sequences selectively impacted upon choice by increasing response accuracy.

In summary, section 3.2 has addressed five areas of temporal expectation paradigms that are in need of development. Firstly, the complexity of decisions required in experimental tasks is overly simplistic compared with many everyday decisions. Secondly, tasks have been restricted to perceptual decisions about features of existing objects and other types of decisions have been overlooked. Thirdly, rhythmic precursors are often dissociated from response targets resulting in timing being artificially separated from goal-relevant information. Fourthly, a binary distinction is often made between periodic and aperiodic rather than investigating degrees of aperiodicity. Lastly, prior knowledge about the timing of the stimulus has not been well controlled for making it difficult to determine what effect this has on experimental responses.

3.3 A new experimental approach

The experimental approach outlined in this section proposes methods for investigating the effects of rhythmic temporal expectation on complex decision making. As stated in chapter 1, complex decision making is defined in this thesis as choices that are based on more than one piece of perceptual information that require integrating memorised content into decision options and/or weighting options in terms of subjective value. It addresses the five limitations described in section 3.2 and focuses on making psychophysical testing more representative of everyday decisions. This is achieved by increasing the complexity of the decision required on each trial whilst at the same time retaining controlled conditions. The approach combines auditory psychophysics with features of existing temporal expectation and decision making paradigms with the aim of finding common ground between disparate disciplines. The proposed approach is not a plea to abandon traditional psychophysical methods, but rather an attempt to develop and enhance them from within. If successfully implemented, the experimental approach should facilitate change in the experimental psychology and neuroscientific literatures by encouraging cross-topic collaboration. The following sections discuss key aspects of the approach, explaining how it works and what benefit it affords the project.

3.3.1 Complex averaging and subjective value

The first feature of the proposed experimental approach involves increasing the complexity of decisions required on each trial during temporal expectation paradigms. This will help to determine the degree to which current theories of temporal expectation generalise to everyday decision making and expand the investigation beyond that of simple perceptual decisions. There are at least two ways in which this can be done: complex averaging and subjective value decision making.

A major challenge for the perceptual system is that the complexity and amount of information in the surrounding environment usually exceeds sensory processing capabilities (Albrecht and Scholl, 2010). For this reason, humans use at least two strategies to detect and process information. One is to direct selective attention towards important goal-relevant features in the environment. The other is to construct low-resolution probabilistic representations that rapidly provide a broad overview of the scene. These statistical summaries have been termed ensemble representations and arise when it is important to process and respond to large amounts of spatially and temporally dispersed perceptual information (Alvarez, 2011). This explains why you might be able to accurately recall the approximate size and direction of a group of people walking through a square, but cannot remember specific features about them. Research investigating ensemble representations has traditionally been limited to the visual domain but there is recent evidence to suggest that the ability to average is also strong in audition (Albrecht et al., 2012; Piazza et al., 2013; Nelken and de Cheveigné, 2013). One benefit of incorporating ensemble representations into a temporal expectation task is that it removes the need to use rhythmic precursors. This is because timing relationships are inherent within and between the different perceptual elements used to construct ensemble representations, meaning that the temporal presentation of goal relevant information can be directly manipulated. A second benefit is that it will determine the extent to which rhythmically presented decision information contributes to the response. This is useful because it will allow stimulus timing and feature information to be investigated and modelled in terms of decision weights.

A further benefit of using complex averaging is that it will force theoretical models of temporal expectation and decision making to jointly describe how dynamic sensory information is integrated within complex DecVs. This is currently lacking within the temporal expectation literature. In a well-cited study, Ariely (2001) demonstrated that when shown a display of discs of varying sizes, humans can accurately report the average size of the discs, but are unable to recall if particular sizes were present. This phenomenon has been widely replicated and is robust to high frequency stimulus presentation, memory delays, wide variation in the number and density of discs and different types of statistical distribution from which stimulus features are drawn (Chong and Treisman, 2003, 2005b; Albrecht and Scholl, 2010; Alvarez, 2011). One explanation for these findings is that individual stimulus properties are computed, combined and then disregarded leading to a form of compression in which only ensemble representations are available as input for memory and decision making (Ariely, 2001; Alvarez, 2011). An alternative suggestion is that individual representations are simply so noisy and inaccurate that observers cannot reliably recall their feature information (Alvarez and Oliva, 2008). Whichever explanation is true, the finding raises important theoretical questions for this project concerning whether effects of temporal expectation on sensory processing are transferred to ensemble representations or disregarded along with individual stimulus representations.

A simple way to implement complex averaging in a temporal expectation paradigm is by having participants compare an average feature of a dynamically presented stimulus array with a referent of some kind. This would allow the onset timing of the stimulus array to be manipulated and response data used to determine the degree to which individual onsets weight the ensemble representation. This is similar to methods used by Morillon et al. (2014) - see section 3.2.1. The array could also be designed to have the same duration as standard entrainment periods and would only need to contain a small number of varying but related stimuli. This is because the accuracy of mean estimations is relatively constant as the number of items to be averaged passes four (Chong and Treisman, 2005a; Haberman et al., 2009; Alvarez, 2011; Piazza et al., 2013).

Subjective value decision making affords many of the same benefits as that of complex averaging. This is because it requires not only investigating how rhythmic temporal expectations bias low-level sensory processing, but also what impact stimulus timing has on the integration of prior knowledge and the formation of complex DecVs. One way that subjective value decision making can be used is by having participants make preference decisions between pairs of easily recognisable visual or auditory options (such as pictures of food items, or the same word spoken by different people). Rather than presenting the stimuli statically, or in a way that timing is largely irrelevant, the information should be presented dynamically within a rhythmic context. By recording participants' preferences for each option separately before the experiment and then measuring the effect of timing on choice, it should be possible to determine whether rhythmic temporal expectations impact upon subjective value and can lead to preference reversals between similarly valued options.

3.3.2 Sound lateralization

The second feature of the experimental approach combats issues associated with isolated response targets (section 3.2.3) and prior knowledge (section 3.2.5). It does this by using sound lateralization techniques and conforming to the following rules: 1. Rhythmic sequences must be combined with goal relevant decision information, and 2. experimental tasks must be representative of everyday decision making. The first rule has already been addressed by a small number of experiments in the timing literature (see section 3.2.5). For example, Morillon et al. (2014) combined goal relevant pitch information by having participants estimate the average pitch of rhythmic tone sequences. Whilst their task nicely integrates timing with goal relevant information, it does not satisfy the second rule. This is because determining the average pitch of a randomly generated tone sequence is neither representative of everyday decisions nor musical judgements. Sound lateralisation, on the other hand, provides an excellent means of satisfying both conditions whilst allowing for the integration of additional features in the experimental approach.

The ability to locate a sound source is essential in many everyday situations and is of considerable importance to both humans and animals. It relies on the use of various cues generated by the interaction of sound waves with the head, ears and torso and enables us to both track the location of goal relevant sounds and segregate different sound sources in acoustically complex environments (Moore, 2012; Ahveninen et al., 2014; Keating and King, 2015). Spatial hearing is highly relevant to complex averaging (Kumpik et al., 2010). This is because every salient head or torso movement changes the spatial location of a sound source in relationship to the body. Therefore, perceiving the spatial location of a fixed sound source requires continuously extracting spatial information and integrating this into a model of the environment that accounts for body movement. This is why it is possible to detect the position of a singing bird hidden amongst a large number of rustling leaves whilst walking down a path. The relevance of complex averaging to spatial hearing is especially clear when goal relevant sounds do not have a fixed location. Imagine that you hear a very quiet buzzing sound and infer that it is a mosquito that you cannot yet see. To avoid being bitten you want to locate it. This task is especially hard because the buzzing sound only occurs when the mosquito is flying and therefore the sound source is always moving. Success relies not only on the ability to rapidly extract and average dynamic spatial information from a moving sound source, but also being able to cross-reference this with changes in the acoustic signal caused by one's own body movements.

To attain ecological validity, the use of sound localization in temporal expectation paradigms need not be as complex as the above mosquito example. This is because it is common for humans and animals to intentionally stop moving when trying to locate a sound source, as a strategy for reducing task complexity. Therefore, a simpler task might be for participants to determine the average location of a sequence of spatially localised sounds without moving their head or torso.

3.3.3 Rhythmic variability

As described in section 3.2.4, human hearing and perceptual judgements are sensitive to rhythmic variability. Temporal expectation paradigms must therefore extend beyond the binary classification of periodic versus aperiodic and be able to investigate what effect varying degrees of aperiodicity has on choice. The third feature of the new experimental approach proposes an analytical solution to this problem. Rather than grouping all aperiodic trials together when analysing the experimental data, aperiodic trials should be divided into subgroups of variability according to the degree of IOI variance contained within the stimulus sequence. Dependent variables can then be analysed by each subgroup of rhythmic variability. The number of subgroups used in the analysis can vary by experiment and be tailored to the experimental question being asked. This approach is compatible with traditional temporal expectation paradigms in that each trial of the experiment will either contain a periodic rhythm or an aperiodic rhythm. The only difference is that a sufficiently large number of aperiodic trials are required to maintain statistical power when analysing variability subgroups.

3.4 Conclusion

This chapter has addressed five limitations of existing temporal expectation paradigms and proposed a new experimental approach aimed at correcting, or at least improving on, them. This is intended to move the investigation beyond that of simple decisions in order to determine whether and how temporal expectations systematically bias complex choice. The approach requires increasing the complexity of the experimental task by having participants make either complex averaging or subjective value decisions. It exploits sound lateralization as an ecologically valid way of combining rhythmic sequences with goal relevant information whilst avoiding the use of irrelevant rhythmic precursors. Rhythmic variability should also be investigated by binning aperiodic trials into subgroups of IOI variance and adopting analytical methods that distinguish degrees of stimulus aperiodicity. Together these suggestions add more types of decision to the investigation, increase the complexity of these decisions, integrate timing with goal relevant information, avoid the use of isolated response targets, allow for the investigation of rhythmic variability and avoid making timing the explicit focus of attention.

The following chapters describe seven behavioural experiments that use and develop the new experimental approach. Each experiment required participants to make dichotomic responses and avoided testing decision scenarios with more than two options. Whilst dichotomic responses restrained the complexity of decisions, they were deemed necessary so as to facilitate the comparison of experimental findings to those of published timing studies. They also allowed for common sequential sampling models that are widely used in the decision making literature to be applied to the data. Ultimately, however, researchers should be aiming to investigate the effects of timing on a range of different and unbounded decision scenarios. This thesis should therefore be seen as being associated with common testing methods, yet to be taking small steps towards greater decision complexity and ecological validity.

The first experiment tests the effects of temporal expectation on complex averaging and addresses what effect prior knowledge about stimulus timing has on decision making (chapter 4). Related questions are then investigated in later chapters with experiments that focus on stimulus rate and complexity (chapter 5), rhythmic variability (chapter 6) and evidence accumulation (chapter 7). The final experiment (chapter 8) varies the type of decision under investigation by testing the effects of temporal expectation on subjective value. Sample sizes for all experiments were not estimated but instead based on studies in the rhythmic temporal expectation and decision making literature (Jones et al., 2002; Correa and Nobre, 2008; Lim et al., 2011; Mathewson et al., 2012; Henry and Obleser, 2012; Wyart et al., 2012; Jepma et al., 2012; Cravo et al., 2013; Morillon et al., 2014). This decision was partly governed by project budget but mainly followed norms in the literature at the time of running each experiment. In order to reduce variance and increase statistical power each study used a within-subject design and required participants to complete at least 30 repetitions per experimental condition (mean: 54, SD: 34.1). This is similar to a number of timing and averaging studies cited in the thesis (Barnes and Jones, 2000; Jones et al., 2002; Chong and Treisman, 2005b; Piazza et al., 2013; Cravo et al., 2013; Curtanda et al., 2015). Chapter 9 ties the experimental findings together to propose a predictive theoretical framework that describes key cognitive processes responsible for the interdependence between temporal expectation and complex decision making.

Chapter 4

Complex averaging and prior knowledge (Experiments 1 and 2)

4.1 Experiment 1 and 2: Introduction

Chapter 4 describes the first attempts at implementing the new experimental approach described in chapter 3 and reports the findings of two new psychophysical experiments. The work has two primary aims: 1. To gather preliminary evidence as to whether rhythmic temporal expectations systematically bias complex decision making. 2. To determine what effect prior knowledge about the relative timing of decision-relevant information has on complex decision making and bottom-up anticipatory processing. Both aims should help to establish whether findings in the temporal expectation literature generalise to decision scenarios that are more representative of everyday decision making. The paradigm used for both experiments is novel and avoids the experimental limitations described in section 3.2. This is achieved by using a complex averaging task in which participants make decisions about the average location of an acoustic sequence of spatially lateralized tones. The method has the benefit of both increasing the complexity of the decision under investigation, and ensuring that timing is not artificially separated from goal-relevant decision information. As well as standard response time and accuracy analysis, classification models were used to estimate decision "weights". Decision weights quantify the degree to which participants prioritise stimulus information in their decision. Here, they were used to assess whether information conveyed by temporally-expected tones influenced choice more than that of temporally-unexpected tones. Rhythmic variability was also investigated by dividing aperiodic trials into subgroups of particular IOI variance and then analysing each subgroup separately.

The experiments were designed to answer two specific research questions. The first question asks whether complex auditory-spatial averaging decisions are sensitive to temporal variability in the stimulus. As both experiments used the same task, this question applies to both studies. The second question asks whether prior knowledge about the rhythmic variability of a stimulus enhances or inhibits the ability to make complex auditory-spatial averaging decisions. This was tested by randomising the order of trials in experiment 1, so that participants did not know whether a trial would contain a periodic or aperiodic sequence, and by providing explicit instructions about the timing of the sequence whilst blocking trials in experiment 2. As discussed in section 3.2.5, determining what influence prior knowledge has on the processing of rhythmic sensory information is key to determining the generalisability of temporal expectation theories. It is also needed to explore the idea presented by Arnal et al. (2014) who claim that sensory entrainment is computationally and biologically distinct from learned top-down predictions (see section 2.2.1). As the findings of Jones et al. (2006) and Iversen et al. (2009) suggest otherwise, experiments 1 and 2 should help to confirm whether or not bottom-up and top-down anticipatory processes are autonomous.

Sound lateralization was used in both experiments because it provides an ecologically valid way of combining rhythmic information with goal-relevant decision information (see section 3.3.2). It was administered by varying interaural cues, so that a sound perceived to ones right would be made slightly louder in the right compared with left ear (interaural level differences) and onset fractionally later in the left compared with the right ear (interaural time differences). The benefits of using this method is that it allowed for the precise manipulation and measurement of the acoustic signal on each trial. Since the stimulus was delivered over headphones, the method also removed the effects of room reverberation that would normally cause acoustic interactions during free field sound localization tasks. If standard sensory entrainment models apply to complex averaging, participants should make faster and more accurate decisions on periodic versus aperiodic trials. Similarly, if top-down anticipatory processes are autonomous from bottom-up processes, participant performance should not differ significantly between the two experiments. If it does differ significantly, it will provide evidence of a more complex anticipatory system in which learned knowledge tunes temporal attending and decision strategies during complex decision making. Due to the similarities between experiments 1 and 2, the method and results of both are reported together to facilitate comparisons between them.

4.2 Experiments 1 and 2: Method

4.2.1 Participants

20 participants performed in Expt. 1 (10 females, aged between 18 - 28, mean = 21.95, SD = 2.54) and 21 performed in experiment 2 (11 females, aged between 21 - 35 in Expt. 2, mean = 26.62, SD = 4.20). Participants only ever performed in one of the two experiments and were paid £10 an hour. An additional two participants were excluded from Expt. 1 because they did not finish the experiment. Neither performed well enough to successfully complete the initial configuration procedure. All participants were right handed students and reported having normal hearing, normal or corrected-to-normal vision and no colour blindness. In Expt. 1, none were practicing musicians. In Expt. 2, two were practicing amateur musicians. Both experiments received ethical approval from the Cambridge Faculty of Music Research Ethics Review Committee.

4.2.2 Task

Both the task and the experimental design were adapted from three conceptually-similar studies which investigated perceptual decision making in the visual domain: Wyart et al. (2012, 2015) and Cheadle et al. (2014).

Figure 4.1 provides a task schematic. On each trial participants listened to a tone sequence comprising 8 presentations of tone k, presented via headphone lateralisation at 8 different spatial locations bounded within a range from far left (-90° azimuth) to far right (+90° azimuth). Tone lateralisation was manipulated via both interaural intensity differences and interaural time differences. This was paired with an on-screen image marking the cardinal (-90°, 0°, +90°) and diagonal (-45°, +45°) spatial axes relative to the participant's midline. The sequence was preceded and followed by a click and presented either periodically, with fixed interonset intervals (IOI) of 333 ms (3 Hz), or aperiodically, with pseudo randomised IOIs between 200 ms and 500 ms (see section 4.2.3.1 for details). The latter resulted in a jittering unpredictable rhythm. Upon hearing the final click, participants judged whether the average orientation of the eight tones was closer to the cardinal or diagonal spatial axes. Participants made a forced binary re-

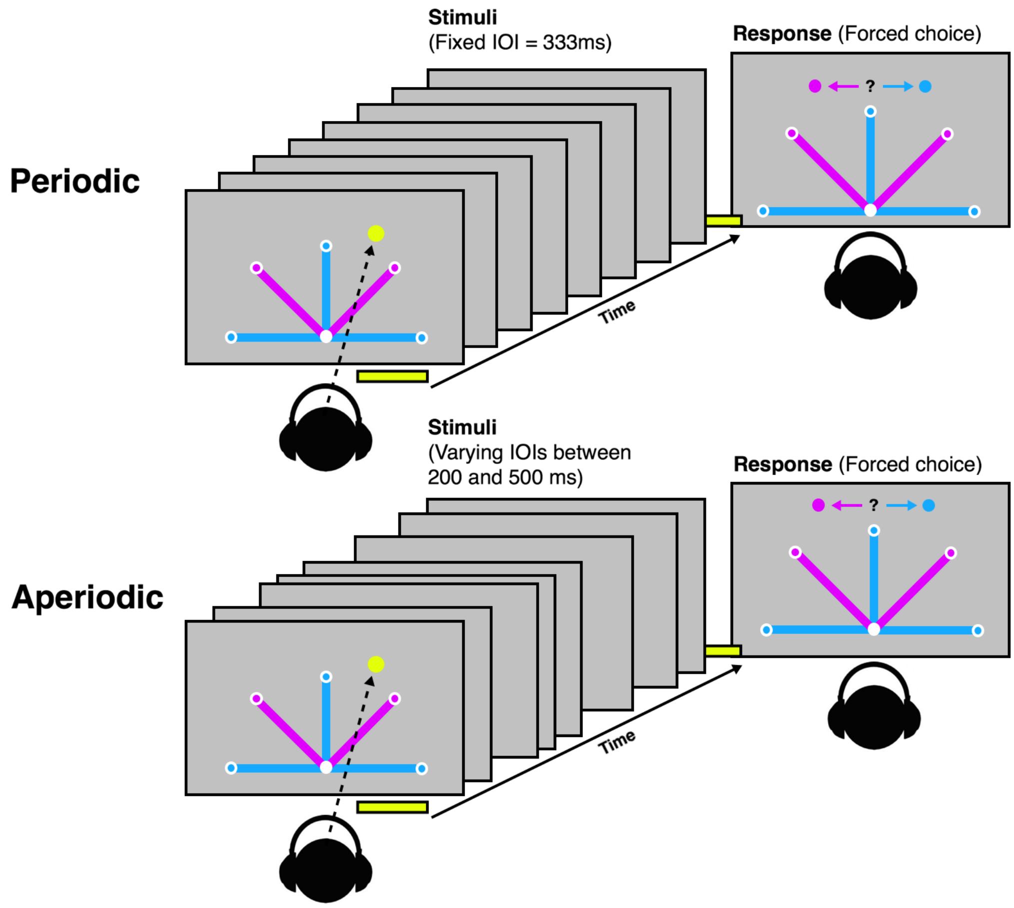


Fig. 4.1 Expt. 1 and 2. Schematic illustration of task structure: On each trial an auditory sequence containing eight tones was presented with either a periodic or aperiodic rhythm. This was preceded and followed by a click that signalled the start and end of each sequence (yellow horizontal bars). Top panel: periodic. Bottom panel: aperiodic. Each tone was lateralized to sound at different spatial locations ranging from -90° to 90° azimuth at 0° elevation (see the yellow dots in the figure). Tone lateralisation was manipulated via both interaural intensity differences and interaural time differences. Participants were to decide whether on average the spatial location of the tones in the sequence was closer to the cardinal or diagonal axes. Cyan lines represent the cardinal spatial axes and the purple lines represent the diagonal spatial axes.

sponse and received corrective feedback on each trial. The experiment contained no trials in which the average was exactly mid-way between both categories.

The main experimental factors being tested for were periodicity [periodic, aperiodic] and spatial category [diagonal, cardinal]. Spatial category featured as a control factor in the analysis to identify whether percepts of lateralised acoustic space remained balanced throughout the experiment or were subject to perceptual biases. This is important to test for due to the novel aspects of the experimental design and because the only published examples of similar experimental tasks occur in the visual and not auditory domain.

4.2.3 Stimuli

4.2.3.1 Tone sequence

Each tone in the sequence was the same 200-ms audio clip of a real bass clarinet note on Bb2 (f_0 = 116.53 Hz, sample rate: 44 kHz, recorded at the Electronic Music Studio, University of Iowa (Fritts, 2012)). It was fit with 10-ms onset and offset ramps as part of the 200-ms duration. This sound was chosen because it contained partials across a broad frequency range (116.53 Hz - 22 kHz) and was reported as being pleasant to listen to. Importantly, piloting revealed that it was the easiest sound to locate in space compared with an oboe producing Bb3 (f_0 = 233.08 Hz) and a wood block click.

The tone sequence was preceded and followed by a click presented in mid-line (a shortened audio clip of a wood block being hit with a drum stick - duration: 200 ms including onset and offset ramps: each 10 ms). The clicks signalled the start and end of each sequence and were included to disrupt priming effects between trials. This is similar to masking methods used by Wyart et al. (2012, 2015) and Cheadle et al. (2014). The last click also controlled for response foreperiod by ensuring that the IOI following the last tone and the response window was the same across both the periodic and aperiodic conditions.

4.2.3.2 Sound lateralization

The perceived location of the tone sequence was manipulated using both interaural time differences (ITD) and interaural level differences (ILD). ITDs were calculated using:

$$ITD = \frac{r(\theta + \sin(\theta))}{c} \tag{4.1}$$

where r = half the radius of the head, constant at 0.09 metres, θ = azimuth in radians and c = the speed of sound in air constant at 343 m/s (Howard and Angus, 2009, p. 98). ILDs varied such that the two headphone channels delivered equal intensities when the sound was deemed to be in mid-line (0°). As the lateralised sound source moved towards azimuths of ±90° relative to the midline, the sound in the opposite channel attenuated linearly to a minimum value of -20 dB relative to that of the signal channel when the latter reached ±90°.

4.2.3.3 Timing

Temporal expectation was manipulated by varying the duration of the IOI between each tone k in the sequence. In periodic trials the IOI between tones was constant. It was set at 333 ms (3 Hz) in accordance with similar studies investigating attentional and neuronal oscillatory entrainment (Barnes and Jones, 2000; Henry and Obleser, 2012; Rohenkohl et al., 2012; Cravo et al., 2013; Hickok et al., 2015). This frequency is thought to be slow enough to avoid the effects of a decision refractory period (Wyart et al., 2012) and fast enough to ensure that each trial did not last too long (average time per trial: 4.5 s).

On aperiodic trials the IOI between each tone was randomly selected from a range spanning 200 to 500 ms. A constraint was applied to this selection in that IOIs lying within ±30 ms of the two preceding IOIs were resampled until a more suitable IOI was found. This value was used because it was considerably larger than the just noticeable tempo difference associated with a base rate of 3 Hz reported by Drake and Botte (1993). This resampling ensured that aperiodic sequences never contained adjacent IOIs that were perceptually similar to one another, which reduced the likelihood of metrical relationships forming throughout the sequence. Piloting revealed that this method of IOI generation produced rhythms that were perceived as being more aperiodic than when randomly selecting IOIs from a small array of non-metrically related values.

4.2.3.4 Ideal decision values

Each tone k in the sequence was coded with an ideal decision value ($\boldsymbol{ID_k}$) which was selected from a continuous array ranging between -1 and 1. $\boldsymbol{ID_k}$ represented the type and amount of categorical information that each tone carried and described each tones location within the decision space. $\boldsymbol{ID_k}$ was negative if the location of the tone was closer to the diagonal axes and positive if it was closer to the cardinal axes. $\boldsymbol{ID_k}$ was zero if the tone was located exactly half way between both axes. $ID_k = -1$ thus represented fully diagonal locations (-45°, +45°), $ID_k = +1$ represented fully cardinal locations (-90°, 0°, +90°) and $ID_k = 0$ represented the locations of the category boundaries (-67.5°, -22.5°, 22.5°, 67.5°).

By using an auditory space that covered multiple axis points ***ID*** values and azimuth information remained orthogonal to one another. For example, different azimuth values (e.g. 20° and -70°) could contain the same categorical information and thus the same ***ID*** value. This ensured that the decision process could be dissociated from any perceptual biases attaching to specific azimuth locations. Figure 4.2 illustrates an adapted version of Wyart et al. (2012) saw-tooth decision-mapping rule that was used to assign ID and azimuth values.

4.2.3.5 Azimuth selection

On each trial, ***ID*** and azimuth values were generated dynamically by sampling 8 values from a probability density function spanning the ***ID*** space [-1 : +1] and then by assigning a corresponding azimuth to each value using the decision-mapping rule (see figures 4.2).

Specifically, this required first sampling an eight digit array from one of two probability density functions (PDF) that had been shifted either positively or negatively by a fixed value (the reference value) within the ***ID*** space [-1 : +1]. The PDF selection depended on whether the trial was to be more diagonal (PDF with mean < 0) or more cardinal (PDF with mean > 0). The reference values, marking the means of both PDFs, were participant-specific and described the distribution for which participants could correctly classify the sequence 75% of the time (see section 4.2.4). The standard deviation (SD) of both PDFs were set to 0.25 to ensure that ***ID*** values could be selected from the complete range of decision space. To standardise the array across trials, resampling was done until the following criteria were met:

1. The resulting mean and SD of the ***ID*** array differed by no more than ±0.01 from the mean and SD of the PDF from which it was sampled.

2. An equal number of ***ID*** array values fell either side of the array mean.

Both PDFs were sampled evenly throughout the experiment to ensure that the number of trials containing predominantly diagonal versus predominantly cardinal tones was equal.

Once an ***ID*** array was found whose mean and SD was similar enough to the PDF, each of its eight elements was randomly assigned to one of the four corre-

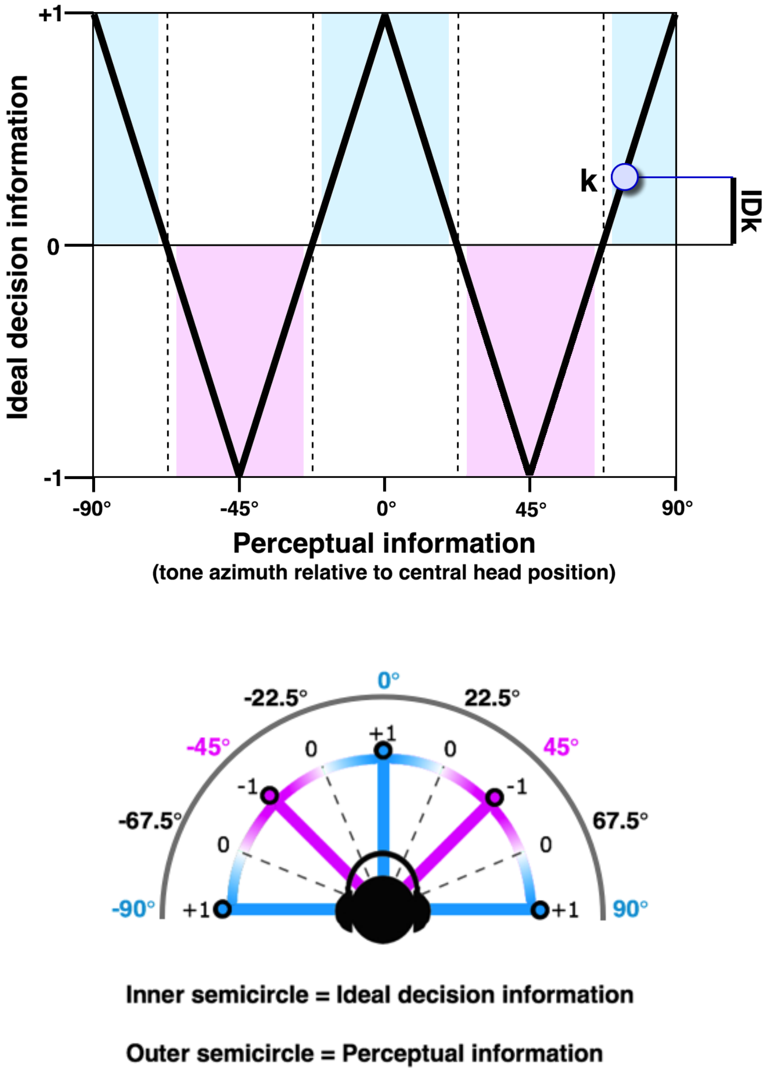


Fig. 4.2 Top panel: a schematic decision-mapping rule, adapted from Wyart et al. (2012), and used to map the location of tone k (perceptual information, x-axis) to the diagonal/cardinal decision axis (ideal decision information, y-axis). All stimuli located between the diagonal axes [±45° - purple areas] and the points of ambiguity [±67.5°, ±22.5° - dotted lines] were assigned a negative ideal decision value (*ID*) between [-1,0]. This represented the relative position within the diagonal category. All stimuli located between the cardinal axes [±90°, 0° - blue areas] and the points of ambiguity [±67.5°, ±22.5° - dotted lines] were assigned a positive ID value between [0,1]. This represented the relative position within the cardinal category. ID_k corresponds to the absolute amount of categorical evidence provided by onset k in isolation and is a measure of ideal decision information. Bottom panel: *ID* values mapped onto the auditory space. Purple spectrum represents *ID* values associated with the diagonal category. Cyan spectrum represents *ID* values associated with the cardinal category. The inner semicircle marks ideal decision information and the outer semicircle marks perceptual information.

sponding azimuth values with the decision space. To understand this last step, note that on the top panel of figure 4.2, every value on the y-axis corresponds to four points on the thick black diagonal line. This increased the likelihood that the azimuths for each trial would be assigned uniformly and ensured that perceptual information remained orthogonal to ID_k values.

4.2.3.6 Auditory feedback

Following Cheadle et al. (2014), auditory feedback consisted of two 100-ms tones, one of 400 Hz and the other of 800 Hz. An ascending sequence (400 Hz then 800 Hz) signalled a correct response, a descending sequence signalled an incorrect response. These feedback sequences began 250 ms after each response.

4.2.4 Procedures

4.2.4.1 Experiment 1

Participants were tested individually and completed all stages of the experiment in sound attenuated recording studio located in the Centre for Music and Science at the University of Cambridge. Each participant underwent an adaptive practice and a calibration session directly before the experiment. The practice consisted of 80 trials (40 periodic and 40 aperiodic, randomised) in which task difficulty was adapted based on the participant's performance. The calibration session comprised an adaptive three-up one-down psychophysical staircase procedure to estimate the threshold for correctly classifying sequences 75% of the time (Kaernbach, 1991). This procedure adaptively varied the mean of the PDF on each trial and selected the threshold after 21 reversals. The resulting mean of the PDF associated with this threshold was then fixed throughout all subsequent trials of the experiment for that participant.

After the calibration session participants performed 8 blocks of 52 trials, 26 periodic and 26 aperiodic, presented in random order. They were told to decide whether the sequence belonged to either the diagonal or cardinal category by pressing the appropriate one of two response keys on a computer keyboard using their index fingers. Response keys were counterbalanced across participants. The response period was cued in two ways: by the appearance of two colour-coded circles in the top left and right segments of the screen (Cyan for cardinal; Purple for diagonal), aligned appropriately for the particular counterbalanced key condition; and by the simultaneous onset of the final woodblock click of the sequence, which sounded 333 ms after the final tone. Response times were collected from the onset of the final woodblock click. Trials timed out if no response had been made after 3 seconds. Feedback was given 250 ms after a response or else at the timeout duration in the case of no response. There was always a 666 ms pause before the next trial, which began from the offset of the auditory feedback.

After each block of 52 trials, participants were presented with a virtual roulette wheel that randomly selected one trial from the previous block. Participants won a bonus of £0.50 if their response on the selected trial was correct. The calibration procedure ensured that participants typically won £3 in additional bonuses throughout the experiment on top of the base payment of £10. This measure controlled the incentive and decision strategy used by all participants on each trial. The experiment lasted about 1.25 hours.

4.2.4.2 Experiment 2

Expt. 2 differed from Expt. 1 in three ways: 1. both the practice and experimental trials were blocked by periodicity. Periodic and aperiodic blocks alternated throughout the experiment, with their orders counterbalanced across participants. 2. A phrase, either "Predicable beat trials" or else "Unpredictable beat trials", was presented on-screen prior to the start of each block. This indicated whether a block would contain periodic or aperiodic sequences. On reading this phrase participants needed to click the computer mouse to start the block. 3. Auditory markers were included in the practice session. The purpose of the markers was to help participants develop and maintain a clear sense of spatial awareness for the task before starting the main experimental session. The markers comprised two tones presented sequentially at each axis location. The presentation order started at 0° , followed by each of the ±45° positions in turn and then the ±90° positions (with left/right order counterbalanced). The markers were presented 5 times throughout the practice session at evenly-spaced intervals, i.e. every 16 trials.

4.2.5 Apparatus

All stimuli were partially or fully generated and behavioural responses recorded using Psychophysics-3 Toolbox (Brainard, 1997) in addition to custom scripts written for MATLAB (MathWorks). These scripts were written by the author of this thesis and can be downloaded from an online repository:

`https://github.com/dcgreatrex-phd/experiment_1`
`https://github.com/dcgreatrex-phd/experiment_2`

All audio files were edited using Audacity v.2.0.5. Images were presented on a 22-inch Iiyama Prolite E2202WS screen with a vertical refresh rate of 60 Hz, which was positioned 100 cm in front of the participant. Responses were collected via an Apple keyboard with numeric keypad and sound heard through a Beyerdynamic DT 990 Pro headset.

4.3 Experiments 1 and 2: Results

The results of Experiments 1 and 2 were analysed separately, but subjected to the same types of analyses. For ease of comparison, each type of analysis reports results for both experiments.

4.3.1 Threshold values

There was a wide range of ability in both experiments. This is seen in the distribution of fixed threshold values estimated during the initial calibration stage. Threshold values were defined by the absolute mean of the PDFs that were used for each participant and can therefore be viewed as a measure of each Ps ability at the task. The closer the mean was to zero (the point of ambiguity), the harder the experimental conditions were due to the underlying PDF being less biased towards one of the two categories. In Expt. 1, the threshold values (i.e. the absolute means of the PDFs) ranged from 0.20 to 0.60 (mean = 0.37; SD = 0.11) within the ***ID*** space [-1 to + 1]. In Expt. 2, they ranged from 0.27 to 0.60 (mean = 0.48; SD = 0.09). The average threshold in Expt. 1 was therefore 0.11 smaller than that of Expt. 2, indicating that a greater number of participants in Expt.1 could do the task under harder conditions compared with those in Expt. 2.

4.3.2 Response time and categorisation accuracy

Figure 4.3 shows the mean proportion of errors and response times (RTs) for Expt. 1 and 2. The behavioural data contains obvious differences between responses in Expt. 1 and Expt. 2, with both periodicity and spatial category having a strong effect on RTs and choice accuracy. Starting with the similarities, RTs were faster during the cardinal and periodic trials of both experiments. This was confirmed by submitting log-transformed RTs to two separate two-by-two repeated measures Analysis of Variance. The analyses revealed strong main effects of periodicity (Periodic: mean = 0.475, SD = 0.144; Aperiodic: mean = 0.569, SD = 0.136; $F_{(1,19)}$ = 92.056; p = < 0.001, partial $n^2 = 0.83$) and spatial category (Cardinal:

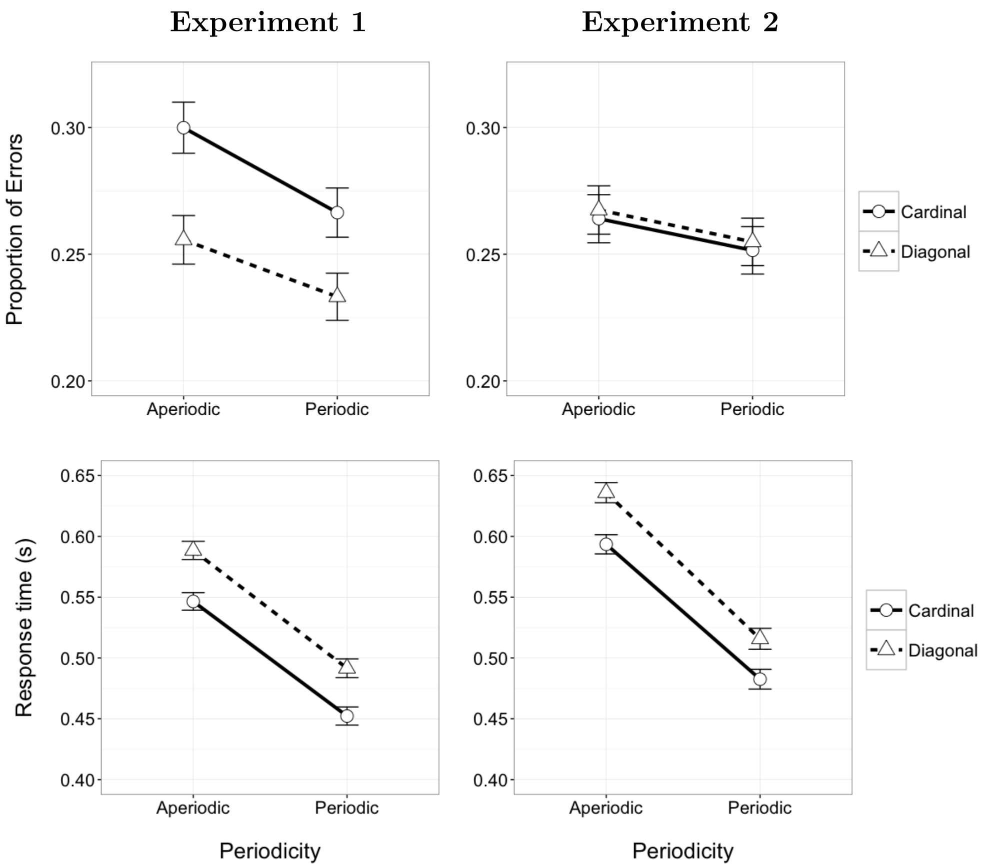


Fig. 4.3 Expt. 1 and 2. Left panel column: data from Expt. 1. Right panel column: data from Expt. 2. Top panel row: mean proportion of errors associated with the periodicity [Periodic, Aperiodic] and spatial category [Diagonal - dashed lines, Cardinal - solid lines] conditions. Bottom panel row: mean response time (s) associated with the periodicity [Periodic, Aperiodic] and spatial category [Diagonal - dashed lines, Cardinal - solid lines] conditions. Error bars = standard error of the mean.

mean = 0.505, SD = 0.131; Diagonal: mean = 0.540, SD = 0.161; $F_{(1,19)} = 5.908$; p = 0.025, partial $n^2 = 0.24$) in Expt. 1. Participants responded faster on periodic compared with the aperiodic trials as well as on cardinal compared with diagonal trials. Similar findings were found in Expt. 2, with main effects of periodicity (Periodic: mean = 0.506, SD = 0.170; Aperiodic: mean = 0.621, SD = 0.156, $F_{(1,20)} = 64.827$; p = < 0.001, partial $n^2 = 0.76$) and spatial category (Cardinal: mean = 0.544, SD = 0.150; Diagonal: mean = 0.584, SD = 0.192; $F_{(1,20)} = 5.950$; p = 0.024, partial $n^2 = 0.23$) on RTs. There were no interactions: Expt. 1 ($F_{(1,19)}$ = 0.366; p = 0.553, partial $n^2 < 0.01$), Expt. 2 ($F_{(1,20)}$ = 0.018; p = 0.895, partial $n^2 < 0.01$).

Differences were found in proportion of errors. Participants made fewer errors on diagonal and periodic trials in Expt. 1, but not in Expt. 2. This suggests that knowing about the timing of the stimulus in advance of hearing it helped participants to suppress the effects of the experimental conditions on decision accuracy. This was supported by subjecting the response accuracy data to two separate two-by-two repeated measures Analyses of Variance: In Expt. 1, there was a main effect of periodicity (Periodic: mean = 0.250, SD = 0.084; Aperiodic: mean = 0.278, SD = 0.084; $F_{(1,19)} = 8.268$; p = < 0.01, partial $n^2 = 0.30$) and spatial category (Cardinal: mean = 0.283, SD = 0.086; Diagonal: mean = 0.244, SD = 0.079; $F_{(1,19)} = 4.989$; p = 0.04, partial $n^2 = 0.21$) on decision accuracy. Participants made fewer errors on periodic compared with aperiodic trials as well as on diagonal compared with cardinal trials. The interaction between the two factors was not significant, $F_{(1,19)} = 0.082$; p = 0.778; partial $n^2 < 0.01$. These findings were not found in Expt. 2 with no effects of periodicity (Periodic: mean = 0.254, SD = 0.090; Aperiodic: mean = 0.267, SD = 0.100, $F_{(1,20)} = 2.281$; p = 0.147, partial $n^2 = 0.10$) nor spatial category (Cardinal: mean = 0.259, SD = 0.100; Diagonal: mean = 0.262, SD = 0.094; $F_{(1,20)} = 0.042$; p = 0.840, partial $n^2 < 0.01$).

To support the analyses, d-prime and criterion values were computed (Green and Swets, 1966). D-prime quantifies participants' categorisation sensitivity whereas criterion accounts for categorisation bias. In line with the previous analysis, paired-sample t-tests showed that mean d-prime values were significantly greater in the periodic compared with aperiodic condition of Expt. 1 (Periodic: mean = 1.395, SD = 0.420; Aperiodic: mean = 1.216, SD = 0.325; $t_{(19)} = 3.119$, p < 0.01), but were not significantly different in Expt. 2 (Periodic: mean = 1.37, SD = 0.520; Aperiodic: mean = 1.34, SD = 0.680; $t_{(20)} = 0.362$, p = 0.721). Criterion values were also, on the whole, significantly greater than zero across all conditions of Expt. 1 (mean = 0.06, SD = 0.121, $t_{(19)} = 2.2$, p = 0.04) but indistinguishable from zero in Expt. 2 ($t_{(20)} = 0.041$, p = 0.968) This indicates that participants' responses were biased towards the diagonal category in the first but not second experiment.

The main effects of periodicity and spatial category on RTs in both Expt. 1 and 2 suggest that both factors regulated the time it took for participants to either encode or recall sequence information from memory. This may highlight a general processing limitation that operates irrespective of prior knowledge or ability at the task. Additionally, RTs were on the whole noticeably longer in Expt. 2 (Periodic: mean = 0.506, Aperiodic: mean = 0.621) compared with Expt. 1 (Periodic: mean = 0.475, Aperiodic: mean = 0.569). As there was no effect of periodicity or spatial category on choice accuracy in Expt. 2, this may be indicative of a speed-accuracy trade off. In other words, by knowing more about the task (due to task instructions, the blocking of trials and the inclusion of auditory sweeps), and taking longer to respond, participants could reduce the degree to which the experimental conditions biased the accuracy of their decisions.

4.3.3 Variance analysis

Experiments 1 and 2 contained two sources of uncontrolled variability due to random sampling across trials. The first was variance in the seven IOIs that separated tones on aperiodic trials. As aperiodic IOIs were randomly sampled from a continuous array of values, IOIs were more variable in some sequences than in others, as the top panel of figure 4.4 illustrates. The second source of variability was the spread in azimuth values on each trial. This was due to the random assignment of ID_k values to one of four corresponding spatial positions that lay between the cardinal and diagonal axis points - see section 4.2.3.5 for details and the bottom panels of figure 4.4 for an example. Some sequences therefore contained tones that would be heard as located in a small area of space, whereas others contained tones located across a much wider area.

To test whether IOI or azimuth variance influenced the data, two additional analyses were done. The first tested whether the SD of IOI values on each trial reliably influenced responses. The second tested the same question using the SD of azimuth values on each trial. This was achieved by first computing the SD of IOI values on each trial and binning each trial into one of four equally sized IOI variance groups (52 trials each). These groups were participant-specific and defined using the 25th, 50th and 75th percentiles of each participant's total

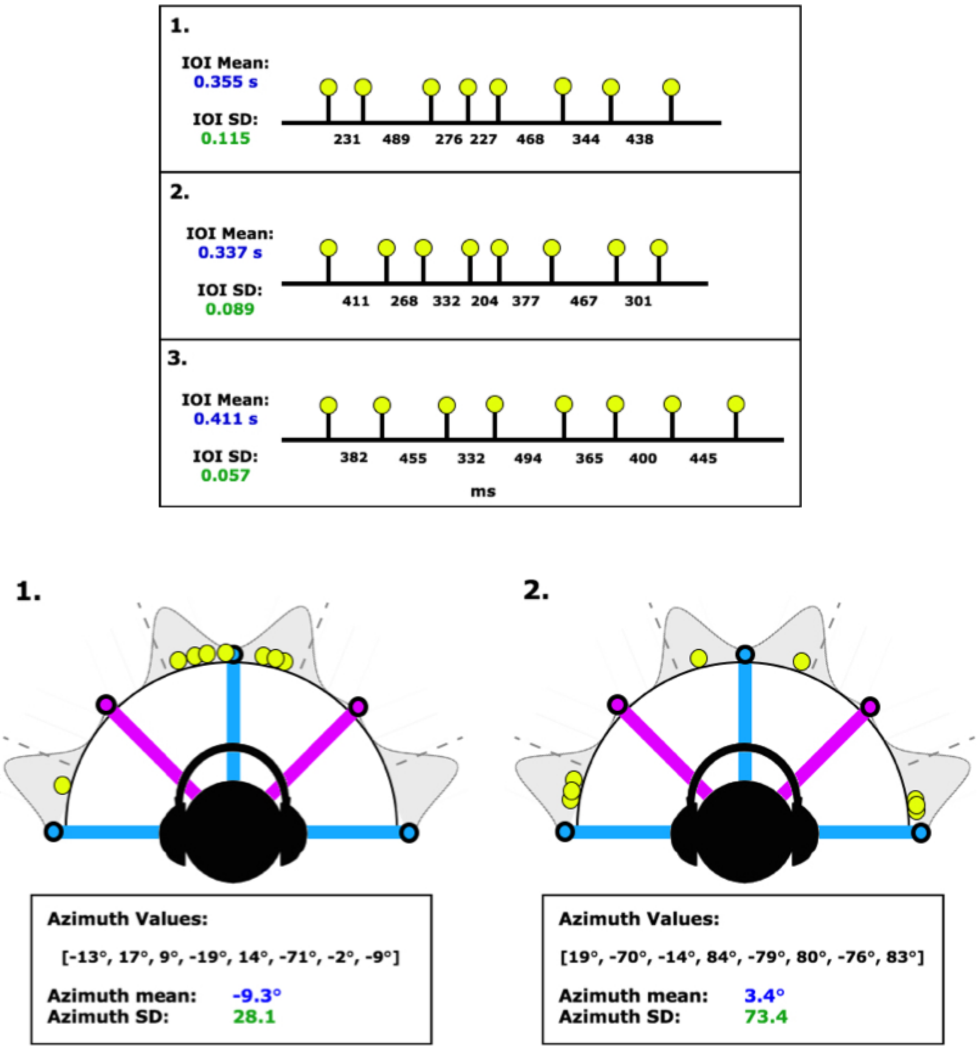


Fig. 4.4 Expt. 1 and 2. Two sources of uncontrolled variability: sequence IOIs and azimuth assignment. Top panel: three IOI sequences that conform to the sampling rule used to generate all aperiodic sequences. 1. High variance sequence. 2. Medium variance sequence. 3. Low variance sequence. Note that each sequence has a different IOI mean and standard deviation. Bottom panel: two cardinal (blue axes) azimuth sequences. Left panel: low variance azimuth distribution. Right panel: high variance azimuth distribution. Although each azimuth sequence has a different azimuth mean and standard deviation, both were generated using exactly the same *ID* array. Variance between sequences was caused by the random assignment of each ID_k value to one of four corresponding ID distributions within the decision space on each trial (see the four grey normal distributions). If each of the 8 ID_k values had been assigned to the same grey distribution, the azimuth sequences would be identical.

IOI SD distribution. This procedure was then repeated for the SD of azimuth values on each trial. Note, that because the spread of azimuths on cardinal versus diagonal trials are necessarily different from one another, the grouping procedure was applied to each spatial category condition separately. This enabled the testing of azimuth variance on the main experimental factors.

4.3.3.1 IOI variance

Figure 4.5 shows RTs and proportion of errors by IOI variance for both Expt. 1 and Expt. 2. IOI variance clearly affected RTs in both experiments but did not systematically affect error rates on aperiodic trials. The more variable the IOIs were on a trial (in relation to all of the aperiodic trials presented during one experimental session), the slower participants were to respond. This was confirmed by submitting mean log-transformed RTs from correct aperiodic trials to a one-way repeated measures ANOVA. There was a significant effect of IOI variance in both Expt. 1 ($F_{(3,57)} = 21.07$, $p < 0.001$, partial $n^2 = 0.77$) and Expt. 2 ($F_{(3,60)} = 31.806$ $p < 0.001$, partial $n^2 = 0.86$). To understand which levels of IOI variance were affected, Bonferroni corrections were used. In Expt. 1, decisions made in the 0:25th percentile group trials (mean = 0.53, sd = 0.13) were faster than both the 50:75th (mean = 0.58, sd = 0.14) and >75th (mean = 0.60, sd = 0.14) percentile groups trials (both $p < 0.001$). Similarly, decisions made in the 25:50th percentile group trials (mean = 0.56, sd = 0.15) were faster than those in the high >75th percentile group trials ($p = 0.003$). In Expt. 2, all levels of IOI variance were significantly different from one another (means = 0.57, 0.61, 0.64, 0.66, SD = 0.15, 0.15, 0.14, 0.17, from low to high degrees of IOI variance on aperiodic trials), except between the 50:75th and the >75th percentile groups ($p = 0.10$). A follow-up 4X2 repeated measures ANOVA ran on the same RT data showed that the IOI variance groupings [0:25th, 25:50th, 50:75th, >75th percentiles] did not interact with the spatial category conditions [Diagonal, Cardinal] during aperiodic trials (Expt. 1: $F_{(3,57)} = 2.28$, $p = 0.09$, partial $n^2 = 0.23$; Expt. 2: $F_{(3,60)} = 0.758$, $p = 0.522$, partial $n^2 = 0.10$). Finally, two additional one-way repeated measures ANOVAs with factor IOI variance showed that IOI variance did not systematically affect error rates during aperiodic trials (Expt. 1: $F_{(3,57)} = 0.1187$, $p = 0.949$, partial $n^2 < 0.01$; Expt. 2: $F_{(3,60)} = 0.156$, $p = 0.926$, partial $n^2 < 0.01$).

The finding that IOI variance systematically increased RTs on aperiodic trials but did not affect error rates is novel. A possible interpretation may be that humans systematically regulate the speed at which auditory information is re-

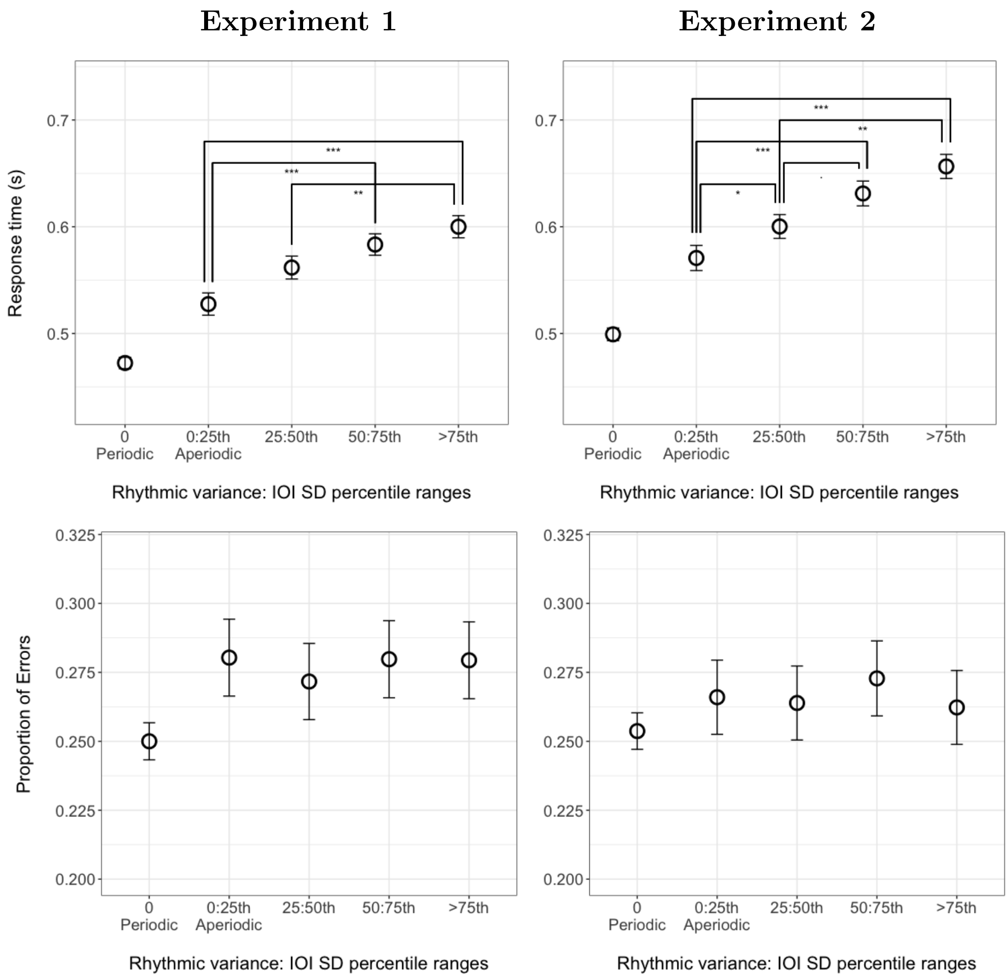


Fig. 4.5 Expt. 1 and 2. Left panel column: data from Expt. 1. Right panel column: data from Expt. 2. Top panel row: mean response times by IOI variance. The x-axis contains five levels of IOI variance: "Zero" marks all periodic trials in which IOI variance was zero. The remaining four levels mark degrees of IOI variance presented during aperiodic trials (from low to high variance). These were made by binning aperiodic trials into one of four equally sized groups depending on the 25th, 50th and 75th percentiles of each participant's total within-trial IOI SD distribution (see section 4.3.3 for more details). Bottom panel row: mean proportion of errors by IOI variance. Error bars = standard error of the mean.

sponded to depending on how temporally variable it is. This regulation would not influence the quality of the information that is encoded, and thus the accuracy of the decision, but simply the time it takes to store and recall this information during an averaging judgement.

4.3.3.2 Azimuth variance

Figure 4.6 shows the mean response times and proportion of errors by azimuth variance and spatial category for both experiments. The most obvious difference between experiments was found in error rates, with there being an interaction between azimuth variance and spatial category in Expt. 1 but not in Expt. 2 ($F_{(3,57)}$ = 6.233 p = 0.001, partial $n^2 = 0.33$, in a 2X2X4 repeated ANOVA with periodicity [Periodic, Aperiodic], spatial category [Diagonal, Cardinal] and azimuth variance as factors). This means that participants made more errors on cardinal compared with diagonal trials in Expt. 1 when azimuth locations were located close to one another, as compared with being widely spread. Follow-up comparisons showed that this was due to higher error rates on cardinal compared with diagonal trials with low azimuth variance (0:25th: $p < 0.001$, 25:50th: $p = 0.019$). For Expt. 2, although there was no interaction, there was a main effect of azimuth variance on decision accuracy ($F_{(3,60)} = 5.418$, $p = 0.002$, partial $n^2 = 0.53$). Follow-up comparisons showed this to be caused by participants making more errors in the highest (most variant) percentile group (>75th). The same analysis was run on log transformed RTs. Azimuth variance interacted with RTs on cardinal trials in Expt. 1 ($F_{(3,57)} = 3.644$, $p = 0.018$, partial $n^2 = 0.38$), but had no significant effect in Expt. 2 ($F_{(3,60)} = 2.388$, $p = 0.08$, partial $n^2 = 0.34$). In Expt. 1, cardinal trials in the two highest variance groups (50:75th and >75th percentiles) were responded to significantly faster than their diagonal-trial counterparts ($p = 0.008$ and $p = 0.01$ respectively).

Azimuth variance was not controlled in the experimental design, yet the analysis highlights that it did affect decision making. It is hard to explain from these data why interactions were found in Expt. 1 and not in Expt. 2. Future designs must therefore account for stimulus variability so that findings can be more easily interpreted.

4.3.4 Decision weight analysis

To test whether serial positioning or implicit metrical functions resulted in participants relying more on information associated with some tones in the sequence

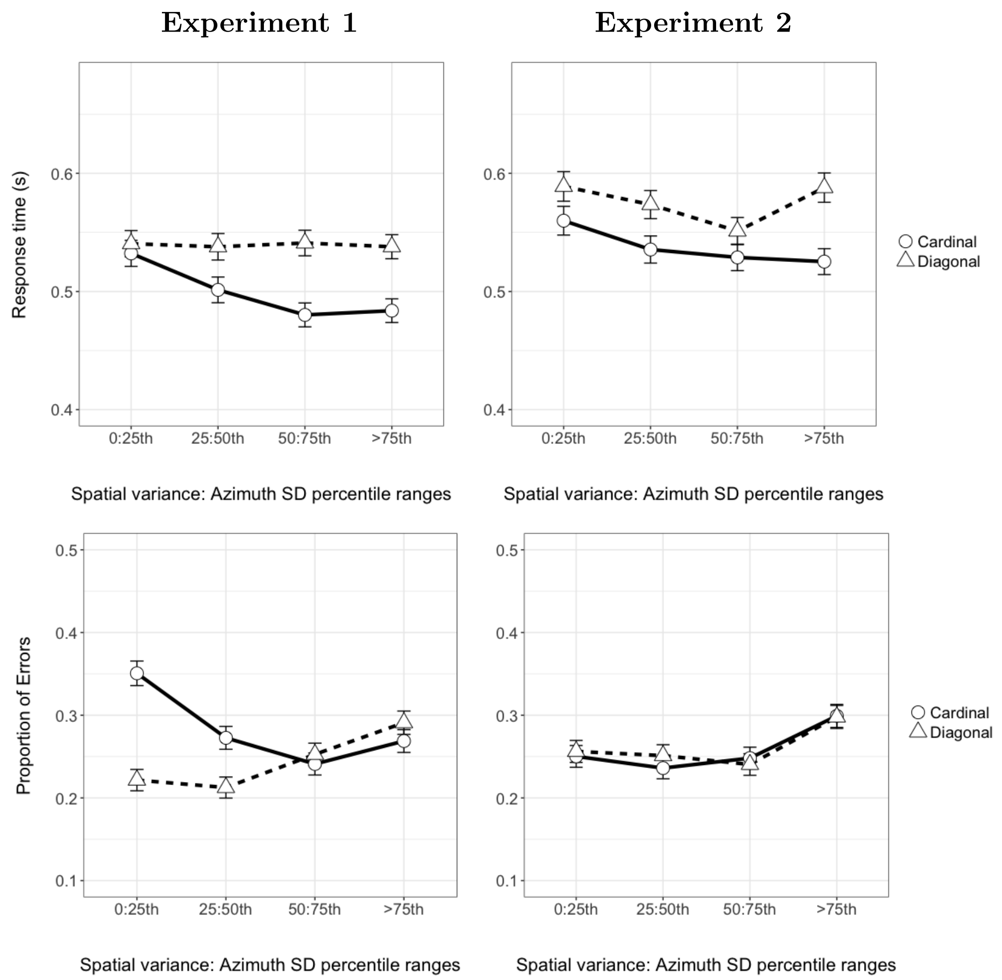


Fig. 4.6 Expt. 1 and 2. Left panel column: data from Expt. 1. Right panel column: data from Expt. 2. Top panel row: mean response times on correct trials by within-trial azimuth variance and spatial category. Within-trial azimuth variance (x-axis) is split into four levels from low to high. These groups were made by binning trials into one of four equally sized groups depending on the 25th, 50th and 75th percentiles of each participant's total within-trial azimuth SD distribution (see section 4.3.3 for more details). Bottom panel row: mean proportion of errors by within-trial azimuth variance and spatial category. Error bars = standard error of the mean.

over others, a decision-weighting analysis was made. This meant submitting the data to a multivariate logistic regression built from a weighted linear combination of ideal decision (*ID*) values. As described in section 4.2.3.4, ID_k values associated with each of the eight tones k on each trial were thus used as predictors in participant-specific regression models. The estimated coefficients associated with each predictor were then taken as a measure for decision-weight. The analysis followed Wyart et al. (2012, 2015) procedures and standard statistical methods, appropriate since the experimental methods were modelled on theirs.

Specifically, the logistic regression estimated decision weights associated with each of the eight tones in the sequence (W_k), each representing the contribution of its corresponding ID_k value to the binary [Diagonal, Cardinal] choice. The model was as follows:

$$P\,(\,Cardinal\,) = \Phi\left(\beta_0 + \sum_{k=1}^{8} W_k \cdot ID_k\right) \tag{4.2}$$

where $P(cardinal)$ corresponds to the probability of judging the sequence as being cardinal, $\phi()$ to the cumulative distribution function (the probit link), and β_0 to a participant-specific bias towards one of the two responses (the intercept).

The model contained 9 free parameters (the intercept term β_0 and eight coefficient values W_{1-8}, entered in this order) and was fitted to each participant's choice data separately for the periodic and aperiodic conditions. This resulted in two sets of decision weights (maximum likelihood coefficients) being estimated for each participant, one corresponding to periodic trials, and the other to aperiodic trials. Patterns within these decision weights were then tested across the participant group via hypothesis testing.

The process of fitting models to each participant's data and then testing group differences using null hypothesis significance testing on the estimated parameters is the widely-used "summary statistics approach" to computational modelling - see Daw (2011) for a review. The approach is related to hierarchical regression modelling in that it treats each parameter estimate as a random variable (random effect) by drawing a participant from the population at random and then running the entire experiment and analysis on that person's data. In contrast, were only one model fit to the entire dataset (i.e. including all participants), the estimated parameters would be treated as fixed effects and as a result, all between-participant variability would be neglected (Daw, 2011, pp. 7-8).

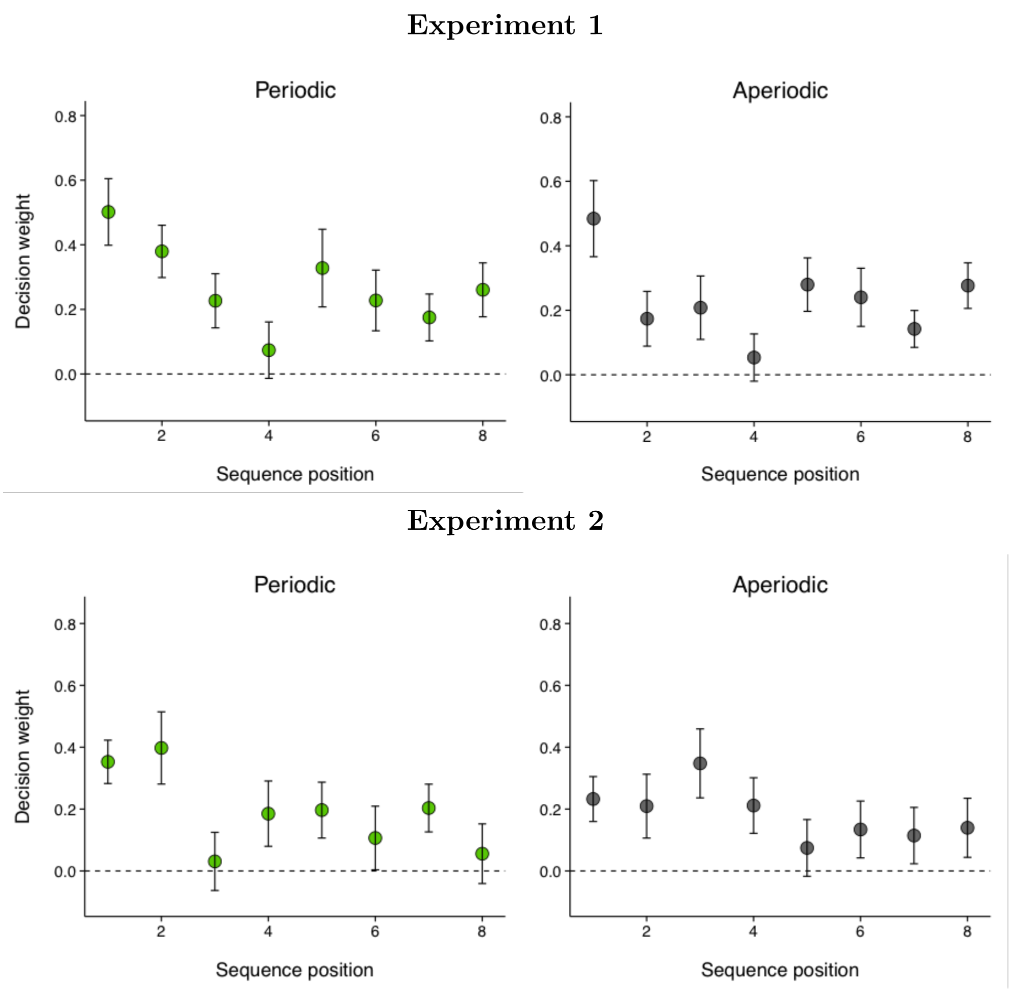


Fig. 4.7 Expt. 1 and 2. Mean logistic regression coefficients indexing the 8 decision-weights in the periodic (left panels) and aperiodic (right panels) conditions, for each position in the noise burst sequence: Expt. 1 (top panel row), Expt. 2 (bottom panel row). Error bars = standard error of the mean.

4.3.4.1 Tone by tone decision weighting

Experiment 1: The top panel row of figure 4.7 shows the group averaged regression coefficients indexing the average decision-weights associated with each of the 8 tones in both the periodic and aperiodic conditions of Expt. 1. Appendix A contains figures of the same data at a participant-specific level (figures A.1 and A.2). All group averaged coefficients were positive and were on average significantly different from zero (Periodic: mean = 0.272, SD = 0.131, $t_{(7)}$ = 5.86, p =< 0.001, Aperiodic: mean = 0.232, SD = 0.126, $t_{(7)}$ = 5.20, p = 0.001). They were not, however, equal assigned and some positions in the sequence were weighted more strongly than others. This was confirmed using one-way repeated measures ANOVAs (Periodic: $F_{(7,133)} = 57.64$, p < 0.001, partial $n^2 = 0.77$, Aperiodic: $F_{(7,133)} = 58.182$, p < 0.001, partial $n^2 = 0.77$). To determine which tones were most influential across the group, follow-up pairwise comparisons with Bonferroni corrections were made. The only signifiant difference, when accounting for across participant variability, was that the 1st tone was weighted more strongly than the 4th tone in both Periodicity conditions (Periodicity: $p = 0.034$, Aperiodic: $p = 0.015$). The lack of other significant differences shows that although decision weights were not evenly assigned throughout the sequence, the weighting pattern was mostly participant-specific. As Bonferroni comparisons are conservative measurements, it is worth noting that the largest averaged weights in Expt. 1 were on tones 1, 2 and 5, and the largest averaged weights in Expt. 2 were on tones 1 and 5.

The enhanced averaged weighting of the 1st tone and suppression of the 4th might imply that participants imposed a metrical structure whilst listening to the sequence. Attention is known to entrain towards periodically embedded time ratios in periodic but not in aperiodic contexts (Large and Jones, 1999; Barnes and Jones, 2000). As similar weighting patterns were found in both periodic and aperiodic conditions, the pattern is likely to have been caused by event statistics rather than metricality: that is, the stimulus always comprised eight tones in both periodic and aperiodic conditions and therefore the 1st and 5th tones could have been used to group the sequence into two perceptual wholes. This may have led to less attention being assigned to the 4th tone due to its proximity to the start of a new group (tone 5). This is equivalent to the phenomenon of "subjective accenting" in which identical sound events are perceived as unequal (Brochard et al., 2003). It was also unsurprising that the 2nd tone had on average lower weighting on aperiodic compared with periodic trials. This is because the IOI between the starting click and the first tone was always 333 ms and therefore it was the second tone that told participants whether the sequence was periodic or aperiodic via the confirmation or violation in the expected beat

A second 2x2 repeated-measures ANOVA tested whether the decision weights were larger in the periodic compared with aperiodic condition and whether they were subject to a primacy or recency bias. The test contained factors periodicity [Periodic, Aperiodic] and time [Early (tones 1 - 4), Late (tones 5 - 8)]. These values were calculated for each level of the periodicity condition, resulting in four averaged decision weights being entered into the analysis for each participant. Mean decision weights were larger in the periodic compared with the aperiodic condition (Periodic: mean = 0.272, SD = 0.179; Aperiodic: mean = 0.232, SD = 0.146; $F_{(1,19)} = 14.198$, p = 0.001, partial $n^2 = 0.43$). There was not, however, any evidence of primacy or recency biases: decisions weights associated with the first half of the sequence (early tones 1-4) were not statistically different from those in the second half of the sequence (late tones 5-8), Early: mean = 0.263, SD = 0.180; Late: mean = 0.241, SD = 0.147; $F_{(1,19)} = 0.483$, p = 0.496, partial $n^2 = 0.02$. This is contrary to findings in similar visual decision making studies that show a strong recency bias affecting averaging decisions (Wyart et al. 2012;2015; Cheadle et al., 2014).

Experiment 2: The bottom panel row of figure 4.7 shows the group averaged regression coefficients indexing the 8 decision weights in both the periodic and aperiodic conditions of Expt. 2. Appendix A contains figures of the same data at a participant-specific level (figures A.3 and A.4). The averaged weights were on the whole positive (Periodic: mean = 0.191, SD = 0.131, $t_{(7)} = 4.126$, $p = 0.004$, Aperiodic: mean = 0.183, SD = 0.086, $t_{(7)} = 6.01$, $p < 0.001$) but noticeably smaller than those in Expt. 1. As in Expt. 1, the averaged weights did not appear to be evenly assigned according to repeated measures ANOVAs (Periodic: $F_{(7,140)} = 60.67$, p < 0.001, partial $n^2 = 0.85$, Aperiodic: $F_{(7,140)} = 61.056$, p < 0.001, partial $n^2 = 0.76$). These differences were not, however, significant at any level when accounting for Bonferroni corrections, meaning that no effects of metricality or fourth tone suppression were found. This suggests that weighting patterns were participant-specific and that prior knowledge helped participants to deploy a more uniform attending strategy which counteracted bias between the periodicity conditions.

Summary of tone by tone decision weighting: Contrary to findings reported in the visual literature, participants' decisions were not biased by the recency of the stimulus. The size of decision-weights, although positive, were smaller and more variable than those reported by Wyart et al. (2012) and Cheadle et al. (2014) and decision weights appeared to be influenced by the statistical grouping in Expt. 1 only. This suggest that spatial resolution in the auditory domain may not be as strong as in the visual domain in this experimental design. If true, participants would have found it harder to extract specific low-resolution categorical information from each tone when perceived as part of a fast moving rhythmic sequence. Likewise, lack of prior knowledge about the stimulus structure may have caused participants to focus their attention towards tones that are important for statistical grouping in Expt. 1, rather than adopting a more optimal uniform attending strategy as evident in Expt. 2. For this reason, decision weights appeared to be more affected by prior knowledge about the timing of the stimulus rather than the periodicity of the sequence.

4.3.4.2 Inlying versus outlying evidence

The regression analysis was repeated to test whether inlying categorical evidence (ID_k values located close to the ID array mean) was weighted differently to outlying evidence (ID_k values located far from the ID array mean). Such biasing has been reported for vision (De Gardelle and Summerfield, 2011; Cheadle et al., 2014), but has yet to be examined for audition.

Following the analytical methods of De Gardelle and Summerfield (2011), the 8 ID_k values from each trial were sorted into ascending order for each participant separately before including them as predictors in the model. To recap, ID_k values lie within a decision space ranging between -1 and +1, where -1 represents fully diagonal and +1 represents fully cardinal locations - see figure 4.2. The resulting coefficients were normalised by dividing by their root mean square in order to reduce the contribution of individual participants whose weights were associated with very low or high error rates (see De Gardelle and Summerfield, 2011, for method). If a bias were to exist, ranked ID_k values 3, 4, 5, and 6 (inlying evidence - ID_k values located close to the ID mean) would have higher decision weights than ranked ID_k values 1, 2, 7 and 8 (outlying evidence - ID_k values located far from the ID mean). Therefore, plotted on a graph, the weights should form a negative parabola if coefficient size were to be measured on the y-axis and ranked ID_k values plotted along the x-axis.

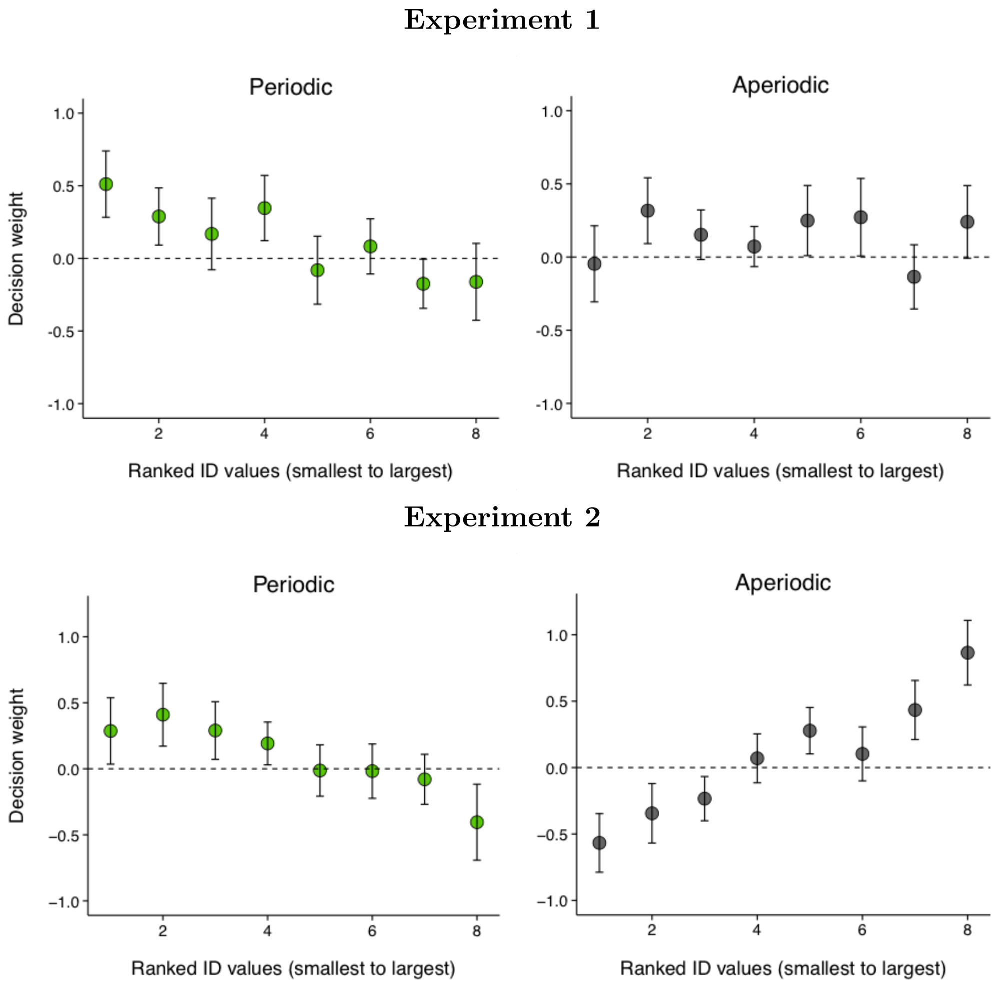


Fig. 4.8 Expt. 1 and 2. Normalised mean regression coefficients indexing the 8 ranked decision weights in the periodic (left panels) and aperiodic (right panels) conditions. Ideal decision (ID_k) values were ranked by size and not sequence position before being entered as predictors in the regression model: Expt. 1 (top panel row), Expt. 2 (bottom panel row). ID_k values represent the amount of categorical information [Diagonal, Cardinal] that each tone k carried regarding the tones spatial location. ID_k values thus lie within a decision space ranging between -1 and +1, where -1 = fully diagonal and +1 = fully cardinal locations. The smaller the ranked ID_k value on the x-axis, the closer it is located to the diagonal axes in relationship to all other tones in the sequence. Conversely, the larger the value, the closer it is located to the cardinal axes in relationship to all other tones in the sequence. Error bars = standard error of the mean.

Figure 4.8 shows the mean regression coefficients indexing the 8 ranked decision weights in both Expt. 1 and Expt. 2. The lack of negative parabolas showed that inlying evidence was not weighted more heavily than outlying evidence in either condition and therefore the predictions of De Gardelle and Summerfield (2011) and Cheadle et al. (2014) were not supported. Instead, some of the data followed a different trend. The weights in three out of the four panels of 4.8 (top left, bottom left, bottom right) appeared to be positively or negatively correlated with ranked ID_k values, depending on the periodicity of the sequence. In Expt. 1, periodic weights seemed to be larger for low ranked ID_k positions (ranks: 1, 2, 3, 4) compared with high ranked ID_k positions (ranks: 5, 6, 7, 8), suggesting that tones closer to the diagonal category were weighted more strongly on periodic trials. This trend was not statistically significant ($F_{(1,19)} = 1.514$, p = 0.223, partial $n^2 = 0.07$). This was tested using a repeated-measures ANOVA with ID_k rank size [the mean ID_k value for small rank positions - 1, 2, 3, 4, versus mean ID_k value for large rank positions - 5, 6, 7, 8] and periodicity [Periodic, Aperiodic] as factors. In Expt. 2, the trends formed opposing positive and negative correlations between rank position and decision weight in the periodic (negative correlation: rho = .-18, p = 0.02) and aperiodic (positive correlation: rho = .40, p < 0.001) conditions. Using the same ANOVA as above, this interaction was significant ($F_{(1,20)} = 11.314$, p = 0.003, partial $n^2 = 0.36$).

To better understand this interaction, the participant-specific models were refit separately to the ranked ID_k data for each of the four periodicity and spatial category conditions [Aperiodic:Diagonal, Aperiodic:Cardinal, Periodic:Diagonal, Periodic:Cardinal]. Figure 4.9 shows the mean regression coefficients across all participants indexing the 8 ranked decision weights for each of these conditions. The x-axis marks the average location within the ideal decision space that each ranked ID_k value occurred within the experiment, rather than ordinal rank position as in figure 4.8). Interestingly, the interaction was caused by a relationship between the periodicity and spatial category conditions. On diagonal trials, the weights of aperiodic rhythms increased the further away the ID_k value was from the diagonal category, but not when the rhythm was periodic. The exact opposite was seen on cardinal trials. On cardinal trials, the weights of periodic rhythms increased the further away the ID_k value was from the cardinal category, but not when the rhythm was aperiodic. This was confirmed using a 2X2X2 repeated measures ANOVA with spatial category [Diagonal, Cardinal], periodicity [Periodic, Aperiodic] and rank size [mean ID_k value for rank positions 1:4, mean ID_k value for rank positions 5:8] as factors. As expected there was a main effect of

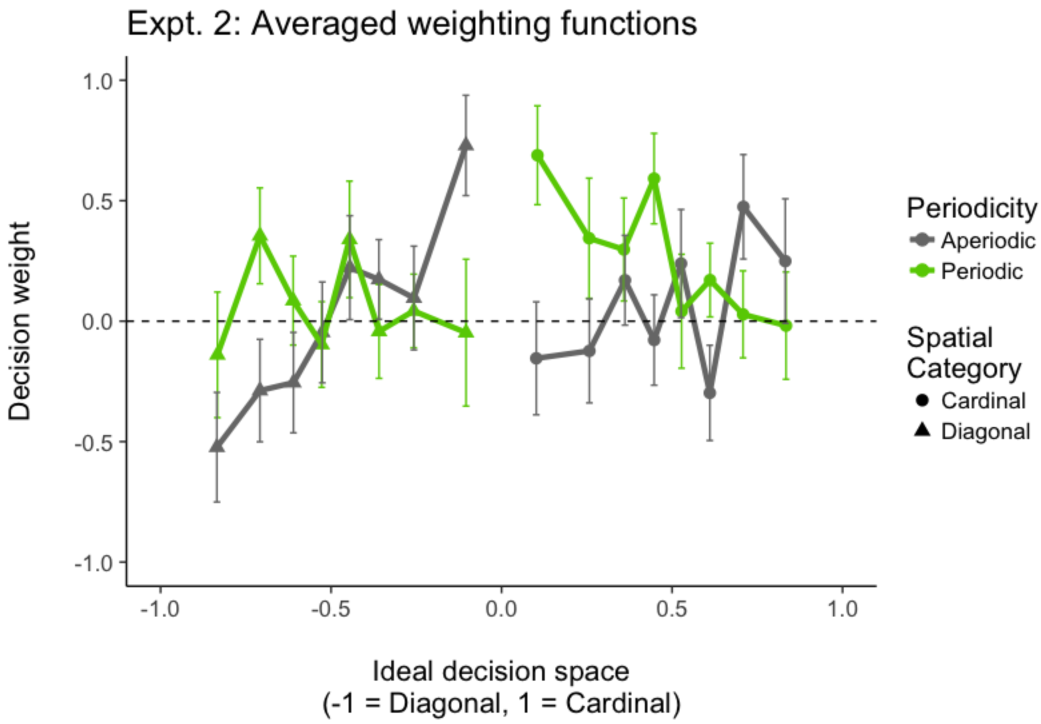


Fig. 4.9 Expt. 2. Normalised mean regression coefficients indexing the 8 ranked decision weights in each of the periodicity [Periodic, Aperiodic] and spatial category [Diagonal, Cardinal] conditions. The x-axis marks the average location within the ideal decision space that each ranked ID_k value occurred within the experiment. Error bars = standard error of the mean.

spatial category ($F_{(1,20)} = 4.966$, p = 0.037, partial $n^2 = 0.20$) and an interaction between periodicity and rank size ($F_{(1,20)} = 7.436$, p = 0.013, partial $n^2 = 0.27$).

Summary of inlying versus outlying evidence: The interaction in Expt. 2, implies that a more complex classification process was in operation than the one described in equation 4.2. Ideal decision information did not appear to be simply accumulated and weighted in a linear fashion. Rather, decision weights interacted with both the rhythmic onset of tones and their perceived location in space. This highlights the complexity of the underlying process and the importance of developing more flexible computational models that better account for stimulus features, such as timing and location, during complex decision making.

4.4 Experiments 1 and 2: Discussion

Experiments 1 and 2 used a novel auditory averaging task to investigate the effects of timing on complex averaging decisions. Their purpose was to highlight whether or not rhythmic temporal expectations bias complex decision making and to what degree prior knowledge about the timing of the stimulus influences this process. The main findings were:

1. Stimulus-induced temporal expectation facilitated auditory averaging decisions by decreasing RTs and error rates.
2. Responses were affected by IOI and azimuth variance in the stimulus sequence.
3. Prior knowledge about the timing of the stimulus reduced the degree to which the experimental conditions biased decision accuracy and led to the use of more uniformly-distributed decision-weights.

Effects of periodicity: There were noticeable differences between the results of both experiments and therefore the hypothesis that top-down predictions about the timing of the stimulus and bottom-up rhythmic temporal expectations are independent of each other, each operating autonomously, is rejected. In experiment 1, periodic trials were associated with faster responses and lower error rates. This finding is similar to those reported in single target timing studies (Rohenkohl et al., 2012; Cravo et al., 2013). It suggests that rhythmic temporal expectations biased complex decision making by enhancing the perception and classification of the tones in each sequence. Importantly, when asked after the experiment about the timing of the sequences in experiment 1, no participant could recall with accuracy the types of rhythms that were used or the frequency of their occurrence. This highlights the automatic nature of the bias and its independence from conscious awareness.

In experiment 2, however, there was no effect of periodicity on choice accuracy, but still a main effect on response times (figure 4.3). Therefore, by presenting trials in blocks of periodicity and making participants explicitly aware of timing, the effect of rhythmic temporal expectation on decision accuracy was lost, but not the effect on speed of response. One possible explanation for this finding is that task instructions and practice conditions encouraged listeners to be more careful. This suggestion is supported by response times being on average slower in experiment 2 compared with experiment 1, and could be tested by emphasising the need for fast responses, and/or de-emphasising the need for accuracy in future experiments. An alternative, but complementary explanation is that prior knowledge enabled participants to deploy different attending strategies in order to complete the task. This idea is advocated by Jones and colleagues (Jones and Boltz, 1989; Jones et al., 2006), and more recently by Schroeder and Lakatos (2009) and Henry and Herrmann (2014) who both use the terms rhythmic and continuous processing modes, as illustrated in figure 4.10. Further investigation into this hypothesis would likely require EEG or MEG data and therefore it will not be focused on for the rest of this discussion.

Lastly, the auditory markers used in the practice session of experiment 2 may have been responsible for the difference in results. The markers were included in the design to help participants develop a clearer sense of the auditory space before doing the task. As the markers were not included in experiment 1, it is possible that they were the sole cause of the difference between the two experiments. In order to check for this, both experiments would ideally be repeated, either using markers in the first experiment or removing them from the design.

Effects of IOI variance: By removing the categorial distinction between periodic and aperiodic, and instead discussing the experimental findings from the perspective of IOI variance (i.e. periodic as zero IOI variance), the timing of the sequence appeared to affect responses in two seemingly independent ways: Firstly, when participants did not know in advance what the rhythm of the sequence would be (experiment 1), participants made more accurate decision when the stimulus did not contain IOI variance (periodic trials). This suggests that, under these conditions, the periodic presentation of auditory tones somehow enhanced the quality

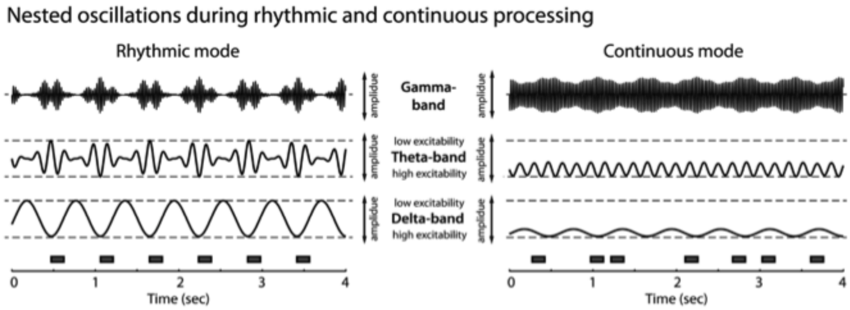


Figure 3. Schematic illustration of neural oscillations. Rhythmic (left) and continuous processing modes (right), and hypothesized nesting of neural oscillations. The phase of low-frequency oscillations (delta–theta band) reflects alternation between low and high excitability periods that affects gamma-band amplitude. In the continuous mode, low-frequency oscillations are suppressed (i.e., smaller amplitude), but remain in a high excitability state to facilitate processing of temporally unexpected events.

Fig. 4.10 A schematic illustration of the rhythmic and continuous processing modes of neural oscillations. From Henry and Herrmann (2014, p. 71). Having prior knowledge about the rhythmicity of a tone sequence in Expt. 2 may have allowed participants to deploy differing processing modes that were tuned towards the rhythmic structure of each trial. This may have reduced the effect of periodicity on choice accuracy in Expt. 2: rhythmic mode processing on periodic trials, continuous mode processing on aperiodic trials.

of the information being encoded and used in complex decision making. Secondly, the timing of the sequence regulated response latencies as a function of IOI variance (both in experiments 1 and 2) in a way that was separate from the quality of decision information. This is because, in both experiments, IOI variance on each trial positively correlated with response latencies but not error rates (figure 4.5). This effect was proportional to the uncertainty and did not affect the quality of the information being encoded. Sensory systems are already known to adapt their coding towards the statistics of attributes of their environment (Dahmen et al. (2010), De Gardelle and Summerfield (2011) and Michael et al. (2014) for recent examples). However, to the author's knowledge, the present finding is the first to show similar effects associated with the IOIs of randomly-timed acoustic events. These two functions of timing (enhanced decision information and RT regulation) should be classed as separate components within a cognitive model that describes how timing biases complex decision making.

Effects of spatial category: Spatial category [Diagonal, Cardinal] affected choice accuracy in experiment 1 but not experiment 2, and response times in both experiments. In experiment 1, diagonal trials were associated with fewer errors than on cardinal trials and in both experiments cardinal trials were responded to faster than diagonal trials. Whilst these effects were unexpected, they can be partially understood when considering the azimuth variance analysis reported in section 4.3.3.2. For example, in experiment 1, participants were more likely to incorrectly classify the sequence as belonging to the diagonal category if the spread of tones was limited to a small area in space (figure 4.6). A likely explanation for this is that the number of categorial axes to choose from was uneven, with three cardinal and only two diagonal axes within the decision space, and therefore an intuitive assumption is that there should be less azimuth variance on diagonal versus cardinal trials. This means that although the azimuth selection procedure comprised an evenly balanced number of $\boldsymbol{ID_k}$ distributions between both categories (see section 4.2.3.5), the asymmetry in on-screen axes may have biased the decision.

Participants also made more errors on high compared with low azimuth variance trials in experiment 2. Asked after the experiment, many reported finding it harder to classify tones within the sequence that were located close to the left/right cardinal axes (±90°) than those in more centralised positions. These accounts concur with psychophysical work investigating minimal audible angles (MAA) and point towards limitations in localisation and lateralization capabilities of the human auditory system (Moore, 2012). Mills (1958) demonstrated that for low frequency tones, MAA increased steadily with azimuth, from 1° for centralised sound sources (0° azimuth), up to 8° for azimuths of ±75° or more. If spatial resolution was weaker for tones associated with the ±90° axes, participants would have had weaker spatial resolution on trials with larger azimuth variability. This explanation does not, however, account for why there were faster response times on cardinal trials or why there were similar error rates on both cardinal and diagonal trials in experiment 2. As cardinal trials were likely to include widely spaced tones, cardinal trials should have resulted in greater error rates in both experiments. A different possible explanation, similar to one proposed by Cheadle et al. (2014), is that participants were most sensitive to the features of tones that shared a similar spatial location with previous tones in the sequence. Therefore the greater the variance among azimuths on each trial the lesser the feature detection. This could represent a process through which previous tones primed spatial attention and increased perceptual gain for specific areas of the auditory space.

Decision weights: In both experiments decision weights were smaller and more variable that those reported by Wyart et al. (2012, 2015) and Cheadle et al. (2014).

There was also no evidence of a recency bias nor heavier weighting of inlying versus outlying decision evidence. This departure from the visual literature implies that participants may have found it harder to extract detailed categorical information from each tone. One reason for this may have been due to the experimental design using a broad auditory decision space which forced participants to continuously reorientate their attention throughout each trial. There were, however, distinct patterns differing between the experiments which imply that different listening strategies were in operation (figures 4.7, 4.8 and 4.9). In experiment 1, the 1st tone was weighted more strongly than the 4th on both periodic and aperiodic trials resulting in a noticeable grouping pattern between the weights associated with tones 1:4 and tones 5:8. Although there was a high degree of between-participant variability, this suggests that participants may have perceptually grouped the sequence into two halves as a means of focusing on structurally important time points (see section 4.3.4.1 for further discussion). In experiment 2, decision weights were more uniformly distributed and no grouping pattern was found. The difference suggests that top-down knowledge about the periodicity of the sequence helped to remove bias and orient the decision process towards the structural properties of the stimulus.

The correlations between ranked ID_k value and decision weight, and their interaction with periodicity, signifies that a more complex classification process was in operation than that of a simple linear accumulation of ideal decision information. Not only were responses not reliant on an equal accumulation of information in experiment 1 (as assumed by sequential sampling models (SSMs) - section 2.2.2), but patterns of relative decision weights changed depending on prior knowledge and contextual features such as the periodicity and spatial location of the tone sequence. This sensitivity highlights that decision theoretic models of choice must be developed to describe how stimulus timing and other contextual factors affect the different components of a decision. A good example of an area in which SSMs could be improved is the assumption that decision evidence accumulates linearly and at a constant rate (Smith and Ratcliff, 2004; Gold and Shadlen, 2007). As subjective grouping patterns were observed in decision weights of experiment 1, information encoding may have at times been fluctuating rhythmically. This suggestion is supported by Wyart et al. (2012) and might represent a general sampling constraint that is not accounted for in the widely used "diffusion" or "race" models of choice (Smith and Ratcliff, 2004; Carpenter et al., 2009) - refer to section 2.2.2 for a more detailed discussion.

Finally, the interpretation of the decision weighting analysis may be complicated by the fact that, compared with high ability participants, low-ability participants needed to do the task with *ID* arrays whose means were close to the true categorical boundaries. Consequently, stimuli used for low-ability participants were more likely to contain tones close to or on the true cardinal and diagonal axes. These participants would only need to detect one of these tones in order to infer the sequence membership and could then stop paying attention to the rest of the sequence. Under harder conditions, the sequences would have contained smaller absolute ID_k values and more between-category variability. This would have forced high ability participants to allocate equal amounts of attention to all eight tones. As figures 4.7, 4.8 and 4.9 show decision weights averaged across both low- and high-ability participants, the results must be interpreted in light of this caveat. Between-subject variability could be reduced in future experiments by increasing the amount of practice time participants are given on the task.

4.5 Experiments 1 and 2: Summary

The experiments described in this chapter are the first attempts at implementing the new experimental approach described in chapter 3. They show that rhythmic presentation of information can systematically bias complex averaging decisions and that prior knowledge about the timing of the stimulus interacts with this process. Rather than periodicity affecting responses in categorically distinct ways, decisions were sensitive to degrees of temporal and spatial variance in the stimulus. For example, IOI variance affected responses in two seemingly distinct ways. The first was that under conditions of high temporal uncertainty (experiment 1), periodic sequences were easier to categorise than aperiodic sequences. This may have been due to periodicity enhancing the quality of decision information being encoded. The second was that decision latencies were proportionally affected by the degree of IOI variance in the stimulus in both experiments. This may suggest that humans systematically regulate the time it takes to compute and respond to decision information depending on how temporally variable it is. Having prior knowledge about the timing of a stimulus also affected responses by reducing bias and ensuring decision information was sampled uniformly. These findings highlight the importance of prior knowledge and context for complex averaging and question the modelled independence between sensory entrainment, top-down knowledge, and evidence accumulation.

A number of experimental design features limit the interpretability and generalisability of the current findings. Firstly, the rate of periodic stimuli used was fixed and did not vary. It is therefore unknown whether the findings replicate under varying presentational rates or are frequency specific. Secondly, the experimental design and task is unnecessarily complex making some of the data hard to interpret. A simpler design in future studies may limit the number of findings that were not anticipated, such as the difference in performance between the diagonal and cardinal spatial conditions. This could be easily achieved by better controlling the spatial presentation of targets, fixing the total duration of each stimulus regardless of condition and using shorter sequences. Finally, interaural cues are not the same as free-field sound localization and a relative, rather than absolute, sound lateralization task should produce clearer results. All three limitations act as the starting point to chapter 5.

Chapter 5

Stimulus rate and complexity (Experiments 3 and 4)

5.1 Experiments 3 and 4: Introduction

Chapter 5 reports two new behavioural experiments which investigate how auditory stimulus rate interacts with effects of rhythmic temporal expectations during complex averaging. The purpose of both experiments is to determine the generalisability of some of the experimental findings of chapter 4 and to refine the implementation of the experimental approach outlined in chapter 3. The aims of the chapter are: 1. To determine whether key findings of chapter 4 replicate under a different experimental paradigm that tests the same fundamental research question. 2. To determine whether the rate of rhythmic stimuli interacts with the effects of rhythmic temporal expectations during complex averaging. 3. To use a less complex experimental design than used in experiments 1 and 2. The experiments involved participants making relative and not absolute spatial averaging decisions on a task in which the interaural level differences (ILD) range and presentational order of the lateralised sound sequence were controlled.

Attentional entrainment models, such as Dynamic Attending Theory (DAT), state that rhythmic temporal expectations arise automatically via the coupling of non-linear biological oscillators to periodicities of similar frequencies in a sensory stream (Large and Jones, 1999). For this reason, from the perspective of DAT, rhythmic temporal expectations should arise during exposure to any periodic sequence, regardless of rate, as long as there exists a biological oscillator of a similar frequency to entrain to. It might, therefore, be assumed that perceptual enhancement associated with rhythmic temporal expectations will not be rate specific, but rather generalise to a wide range of stimulus rates and context. This is because the underlying cognitive mechanism is theorised to be adaptive to a range of different inputs. Indeed, Sanabria et al. (2011) showed that both fast (500 ms IOI) and slow paced (950 ms IOI) rhythmic precursors helped to reduce participants response times towards rhythmically expected, but not necessarily metrically aligned, auditory targets in a standard temporal expectation task.

Apart from a small number of studies that have explicitly tested the generality of DAT, the idea that DAT generalises across rates appears to have influenced previous experiments in the literature. Table 5.1 illustrates that 15 out of 20 reviewed experimental paradigms used only one periodic stimulus rate to investigate effects of rhythmic temporal expectations on reaction times and perception and therefore did not test whether experimental findings replicated at different rates within the same experimental design. This is problematic not only in that it makes it difficult to determine whether experimental findings associated with perceptual and decision enhancement generalise to different rates or are the result of specific features of an experimental design, but also because any attempt to assess from the literature whether stimulus rate does not matter entails experimental findings being compared adhoc across a range of independent experiments.

There is, however, evidence to suggest that the rate of a rhythmic sequence may influence the effects of rhythmic temporal expectation on choice. For example, Wyart et al. (2012) demonstrated that during extended categorical decision making the rate of evidence accumulation fluctuated rhythmically. This resulted in refractory periods in which new sensory information had a weaker impact on the same choice. These periods occurred for 250 ms after the onset of salient decision information within a rhythmic stream. This suggests that periodic sequences with an IOI of 250 ms or less will be harder to process and categorise compared with slower rhythms in which sequentially presented decision information does not fall within the refractory period. A less technical reason to question the assumption, however, is that humans demonstrate preferred rates for tapping in-time, walking, clapping and engaging in other natural movement. Moelants (2002) provides evidence that our average preferred tempo is around 500 ms (120 bpm) and that this corresponds to the average natural walking speed of adult males (513 ms - 117 bpm), as well as other activities such as rhythmically synchronised applause. If some stimulus rates align with common and preferred cognitive and motor processes, it is perhaps likely that information contained within the rhythmic sequence will be easier to detect, process and respond to. It may also be assigned more value. Drake and Botte (1993) support this idea by showing that

Experiment	*Periodic IOI ms*	*Aperiodic IOI ms*	*Domain*
Large and Jones (1999):			
- Expt. 2	600	NA	A
Barnes and Jones (2000):			
- Expt. 1, 2, 4, 5, 6	600	NA	A
Jones et al. (2002)	600	200 to 850	A
Doherty et al. (2005)	550	200 to 900	V
Ellis and Jones (2009)	500	NA	A
Iversen et al. (2009)	200	NA	A
Henry and Obleser (2012)	333	NA	A
Mathewson et al. (2012)	82.3	11.7 to 257.4	V
Rohenkohl et al. (2012)	400	200 to 600	V
de la Rosa et al. (2012)	800	400 to 1200	A
Cravo et al. (2013)	400	200 to 600	V
Lawrance et al. (2014)	250	150 to 350	A
Morillon et al. (2014)	667	NA	A
Curtanda et al. (2015)	800	400 to 1200	A
Hickok et al. (2015)	333	NA	A
Barnes and Jones (2000):			
- Expt. 3, 7	300, 500, 600	NA	A
Rohenkohl and Nobre (2011)	400, 800	300 to 900	V
Sanabria et al. (2011)	450, 950	NA	A
Marchant and Driver (2012)	100, 200, 300, 400	100 to 400	AV
Sanabria and Correa (2013)	450, 950	NA	A
Mean IOI:	482.69		

Table 5.1 The rates of rhythmic stimuli used in a sample of rhythmic temporal expectation experiments reported in the timing literature. The list represents a subset of experimental designs that have relevance to this thesis. "Domain" refers to the type of rhythmic stimuli used in each experiment: A = Auditory. V = Visual. AV = Auditory-visual. NA means that the experiment did not have an aperiodic condition.

humans make better tempo discrimination judgements when the IOI of a periodic sequence falls within a range of 300 to 800 ms.

As identified in table 5.1, there are a small number of studies that have incorporated multiple rates into a single experiment when studying rhythmic temporal expectations. Barnes and Jones (2000) used periodic sequences with different rates to determine whether harmonically related expectations resulted in the same perceptual effects across conditions. Marchant and Driver (2012) used four different rates as part of their periodic stimuli and averaged response data across these rates when testing for temporal expectations. Sanabria et al. (2011) and Sanabria and Correa (2013) demonstrated that both fast and slow rhythmic temporal precursors facilitate speeded responses times towards isolated auditory targets. Finally, Rohenkohl and Nobre (2011) used two different rates and compared participants' responses between them. They showed that responses to a rhythmically presented visual target were significantly faster when the target was preceded by a periodic versus aperiodic sequence with an IOI of 400 ms, but this distinction was not significant when a slower periodic rhythm was used (IOI of 800 ms).

In addition to stimulus rate, a major concern for this chapter is to ensure that the experimental task is less complicated than that of experiments 1 and 2. This is because, whilst experiments 1 and 2 act as a good starting point to the investigation, there are a number of specific features that limited the interpretability of the experimental data. Firstly, the range of ILDs used on a single trial was not constrained to a specific area in space and could occur in extreme left or right locations. This meant that participants needed to continuously reorientate their attention across the entire frontal spatial plane on every trial. This wide attentional reorientation was not controlled and therefore it may have interfered with processes associated with rhythmic temporal expectations. Some researchers in the visual domain avoid this problem by having participants make spatial decisions about the tilt of an orientated Gabor pattern positioned within a stationary circle (Cheadle et al., 2014; Wyart et al., 2012, 2015). This has the benefit of ensuring that selective attention is fixed to one area in space during each decision. Although this is harder to achieve in the auditory domain, the range of lateralized sounds used on every trial could be limited to a specific area in space. Ideally, this should be forward facing and avoid large absolute ILD values. This is because auditory spatial resolution is known to deteriorate the further a sound source moves away from mid-point (Mills, 1958).

Two other limitations are that the spatial order in which the lateralized tones were presented on each trial was randomised and not controlled and that the total duration of aperiodic sequences was not fixed. This assumes, perhaps falsely, that the presentational order of the spatial locations and/or small variations in the total duration of aperiodic sequences had no inhibiting or enhancing effect on choice. Intuitively, however, tracking a sound that moves sequentially from left-to-right is a very different task to tracking a sound that moves in a random order; even if the same absolute spatial locations are used in both sequences. One way that the issue associated with spatial order can be corrected is to develop a measure of movement complexity that quantifies the total change in ILDs between the sounds in a lateralized sequence. Once computed, only order permutations that have the same movement complexity score can be used as stimuli. One way to control the total duration of aperiodic sequences is to randomly sample IOIs whilst applying sampling constraints.

Lastly, other researchers have discussed problems that can arise when participants make absolute spatial judgements based on a lateralised sound signal. As discussed by Hopkins and Moore (2010), the perceived position of a lateralised acoustic signal can be influenced by factors other than interaural differences and therefore the mapping between these differences and perceived azimuth may not be linear under certain conditions of experiments 1 and 2. Additionally, neuronal ILD and interaural time difference (ITD) sensitivity, as well as the perception of auditory space, has been shown to change according to the statistical distribution of a preceding lateralized stimulus (Keating and King, 2015; Dahmen et al., 2010; Maier et al., 2012). One way to avoid this limitation is to require participants to make relative rather than absolute spatial decisions on each trial.

Experiments 3 and 4 advance the new experimental approach outlined in chapter 3 by implementing the above suggestions. Apart from stimulus rate, all other aspects of experiments 3 and 4 were identical so that cross-experimental comparisons could be made.

5.2 Experiment 3: When the pulse is fast

Experiment 3 used a 4 Hz periodic rhythm (IOI of 250 ms) to investigate whether periodicity affects complex auditory averaging decisions during fast stimulus presentation. This was 1 Hz faster than that the 3 Hz rate used in experiments 1 and 2 (IOI of 333 ms) and was the second fastest rate used by other auditory studies in the timing literatures (i.e. only Iversen et al. (2009) was faster in table 5.1). The task required making a relative auditory averaging decision between the perceived spatial location of a lateralized rhythmic sound sequence and a probe tone.

It was hypothesised that IOI variance would strongly affect decision latencies and that periodic sequences would be responded to faster than aperiodic sequences. Periodicity was also expected to enhance choice accuracy if the responses were similar to experiment 1. This is because participants were given no information about the timing of the rhythmic sequence prior to each trial, a situation similar to experiment 1. It was unknown, however, what effect the decision refractory period described by Wyart et al. (2012) would have on responses.

5.3 Experiment 3: Method

5.3.1 Participants

A total of 24 participants (14 female) took part. All were students aged between 18 to 40 (Mean = 23.5, SD = 4.59) and were paid £10 an hour. Two out of the 24 participants were excluded from the analysis because neither performed well enough to successfully complete the initial configuration procedure. All participants had tone detection thresholds of 15 dB HL or better as measured with a Grason-Stadler GSI 16 audiometer at octave frequencies between 250 Hz and 8000 kHz. At each tested frequency the thresholds for each ear differed by less than 10 dB HL. All but two participants were right handed and three were practicing musicians. The experiment received ethical approval from the Cambridge Faculty of Music Research Ethics Review Committee.

5.3.2 Auditory stimulus

The basic stimulus consisted of a train of six 40-ms bursts of broadband Gaussian noise, each including 5-ms cosine-squared ramps at both onset and offset. These were followed by a 40-ms 2 kHz probe tone with 5-ms cosine-squared ramps at both onset and offset. The noise was bandpass filtered using an 8th order Butterworth filter with cutoff frequencies at 300 Hz and 20 kHz. Both the bandpass filter and the frequency of the probe were used to target high-frequency sounds in which sensitivity towards interaural level difference (ILD) is optimal and mimicked stimuli used in other published sound localisation studies (Wenzel et al., 1993; Kumpik et al., 2010; Johnson and Hautus, 2010; Carlile et al., 2001). Each noise burst and probe tone was lateralised by presenting the sound to the right ear at ~70 dB sound pressure level (SPL) plus 0.5 times the total ILD, and the sound to the left ear at ~70 dB SPL minus 0.5 times the total ILD. As discussed by Wright

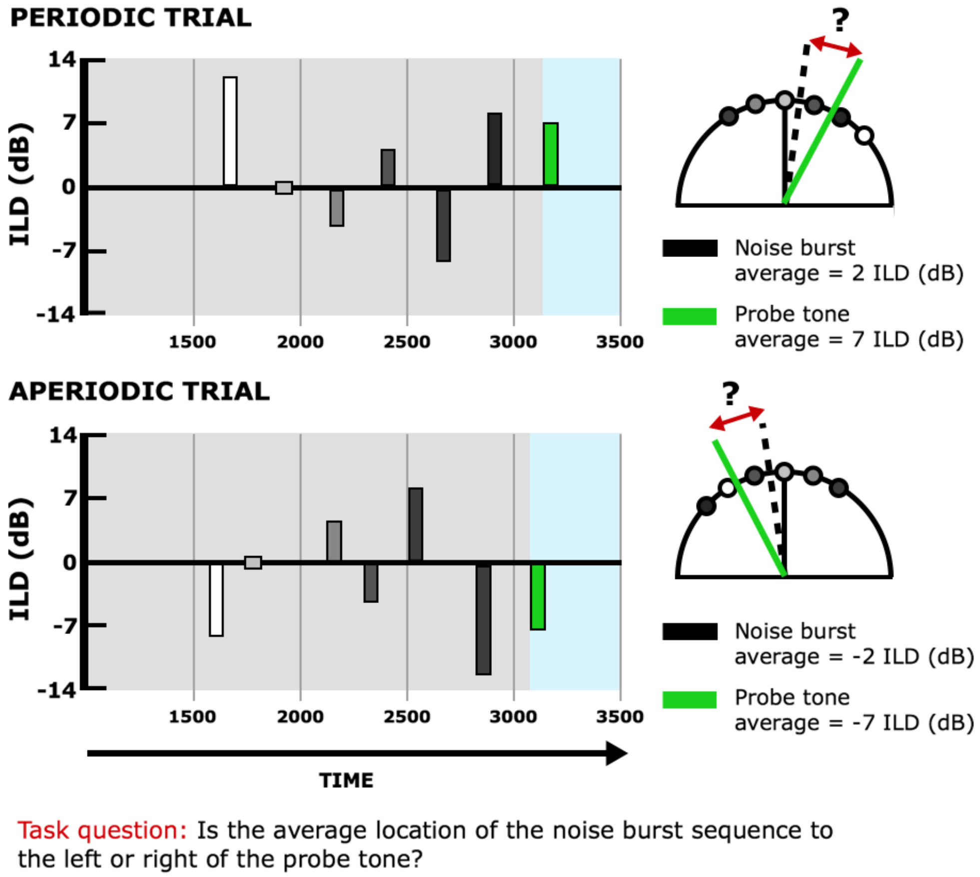


Fig. 5.1 Expt. 3. Schematic illustration of task structure: each trial began with silence lasting anywhere between 1500 and 1750 ms, randomly selected. This was followed by an auditory sequence containing six bursts of noise followed by a high pitched probe tone, periodic (top panel), aperiodic (bottom panel). The IOI between the final noise burst and probe tone was fixed throughout the experiment. Each individual noise burst and the probe tone were spatially lateralised via inter-aural level differences. The task was to decide whether the average spatial location of the noise burst sequence (represented by the dashed black lines in the right panels) was located to the left or right of the probe tone (represented by the solid green lines in the right panels). The rhythm of the noise burst sequence was periodic on 50% of trials and aperiodic for the rest, randomly sequenced.

and Fitzgerald (2001) this technique helped to keep the perceived overall level of the stimulus constant across different ILDs.

5.3.3 Noise burst lateralization

On every trial, ILDs were individually assigned to each of the six noise bursts in the stimulus sequence, such that the sequence's average ILD took one of five possible values, -4, -2, 0, +2, or +4 dB, with the value of its individual ILDs never exceeding ±14 dB, and the sequence ILDs always deviating from the average by -10, -6, -2, +2, +6 and +10 dB. Thus the ILD in the noise burst sequence were never closer than 4 dB, which exceeds by 3 dB the discrimination thresholds for detecting lateralisation differences (Moore, 2012). Sequences with each of the 5 ILD averages [-4, -2, 0, +2, or +4 dB] occurred with equal likelihood throughout the experiment. Balancing the average ILD around mid-point (0 dB ILD) ensured that any directional biases could be identified and separated from the effects of timing on choice. Directional biases have been previously reported to be caused by a deterioration in spatial resolution the further a sound source deviates from mid-point on the horizontal plane (Mills, 1958; Moore, 2012). Additionally, a spatial "aftereffect" is known to occur whereby attenuation towards a left or right lateralised acoustic signal biases the percept of the mid-point (Kashino and Nishida, 1998; Carlile et al., 2001; Dahmen et al., 2010).

Once the ILDs had been assigned to a trial, the six lateralised noise bursts were presented in a pseudo-randomised order. Piloting revealed that sequences with a simple presentation order, such as a sweep from left to right, were harder to average than those containing large spatial jumps and direction reversals. For this reason, on each trial, the presentation order was randomly selected from a list of 88 order permutations that controlled for spatial pattern complexity. Specifically, each presentational order contained three direction reversals and a movement complexity score of 13. Movement complexity measures the size of spatial jumps within a sequence and was calculated using the following equation:

$$Movement\_Complexity = \sum_{k=1}^{i-1} abs(S_{k+1} - S_k) \tag{5.1}$$

where, S = spatial location of noise burst (normalised 1-6 from left to right), k = current noise burst on the trial, i = number of noise bursts within a trial (set to 6 for the current experiment). A score of 13 represents the 75th quartile boundary of all possible movement complexity scores within the current design.

5.3.4 Timing

Temporal presentation of the sequence of noise bursts could be either periodic or aperiodic. Periodic IOIs were always 250 ms. Aperiodic IOIs were randomly selected from a range between 125 and 375 ms. This range was determined using the temporal jitter procedure described by Goupell et al. (2009) and Brown and Stecker (2011). For each trial, the IOI was drawn randomly from a uniform distribution centered at a IOI of 250 ms (4 Hz), with a dispersion equal to $2i$ times the centered IOI. The parameter i thus defined the degree of temporal jitter, with $i = 0$ corresponding to no jitter, i.e. perfect periodicity, and $i = 1$ to maximal jitter (individual IOIs ranged from 0 to 2 * centered IOI). Periodic sequences were thus generated with $i = 0$ while $i = 0.5$ was chosen for the aperiodic sequences. In order to reduce the likelihood of metrically related IOIs occurring in the aperiodic condition, aperiodic IOIs were resampled if their value was within ±25 ms of the preceding IOI in the sequence. This followed methods described in experiments 1 and 2 (see section 4.2.3.3 and Drake and Botte (1993)). Resampling ensured that the total duration of the aperiodic sequence was always the same as the total duration of the periodic sequence, with a variability of ±5 ms. Each trial used a new random jitter seed with resampling conducted before each trial until IOIs were found that fit the selection criteria. The IOI between the final noise burst of the sequence and the probe tone was always 250 ms on both periodic and aperiodic trials. This controlled against different foreperiods between the two events affecting responses between conditions.

5.3.5 Design and procedure

Practice trials: Participants practiced the task helped by visual feedback. On each trial they were shown an on-screen semicircle representing the auditory space. This was similar to those shown in the righthand panels of figure 5.1, but without the circles and lines representing averages. The task was to decide whether the average spatial location of the noise train was to the left or the right of the probe tone by pressing one of two response keys on the numeric pad of a computer keyboard. During each trial participants rested their index finger of their preferred hand on key 5, and responded by pressing the key immediately above (8) or below (2) the 5 (response keys were counterbalanced across participants). Use of vertical response keys avoided counterintuitive response mapping which would have arisen were horizontal response keys used.

Each trial began with silence that lasted anywhere between 1500 and 1700 ms. The noise train and probe tone then followed. To aid responses, a small white circle appeared on-screen at the same time and in the same spatial location as the probe tone. Participants could respond at any time from the onset of the probe tone and had a maximum of 5 seconds before the response period timed out. After a response, the true average location of the sequence was represented by a coloured line, which originated from the bottom centre of the semicircle and projected at the correct angle relative to the semicircular space. The line was green if the answer was correct and red if it was incorrect. A 200 ms pause followed either a response or the timeout period in the case of no response, before the next trial began. The practice session comprised 50 trials randomised by difficulty. Visual feedback (the coloured line) was given for the first 30 trials only; the last 20 had no feedback. For no feedback trials, participants were told they could either look at the screen or to close their eyes, depending on which felt most comfortable. Feedback was not given in any other portions of the experiment.

Calibration: After the practice session, participants were asked to adjust their headphones until the subjective location of a stream of broadband Gaussian noise with 0 dB ILD was perceived as centered relative to the mid-point on-screen. This was intended to minimise variations in the stimulus caused by misplacement of the headphones and replicated procedures used in other sound localisation studies (Hafter and Dye Jr, 1983; Yost and Dye Jr, 1988). Each participant then underwent a calibration session using a 3 up 1 down adaptive staircase procedure (Kaernbach, 1991). Its purpose was to estimate the difference in ILD between the probe tone and the average noise train location associated with correct classification 75% of the time (i.e. the decision threshold). This estimate was later used to tailor the location of the probe tone relative to the sequence average for each participant in the main experiment.

During the calibration, the difference between the probe tone ILD and the noise train average ILD was systematically varied from a starting difference of 11 dB between the two. To start with, the difference between the probe tone and average ILDs was decreased in steps of 2 dB after every correct response, and then set at the permanent default of 0.33 dB after the first incorrect response. For the rest of the calibration session, incorrect responses were followed by an increase in ILD of 3 * 0.33 dB, whereas correct responses were followed by a decrease in ILD of 1 * 0.33 dB.

As in the practice session, participants were shown an on-screen semicircle. As it was hypothesised that periodic sequences may be associated with more accurate decisions, the perceptual threshold was estimated using aperiodic trials only. To ensure that participants received equal exposure to both types of rhythms, the process contained two randomly interspersed staircasing tracks, one using periodic rhythms, the other aperiodic rhythms. The calibration session terminated after 10 reversals on aperiodic trials, following which the 75% decision threshold was estimated as the average of the last 6 aperiodic reversal points. This estimated was made due to 75% being the convergence point of the 3 up 1 down method described by Kaernbach (1991).

Experiment proper: The experiment proper began after the calibration session ended and when participants indicated that they were happy to continue. Each participant responded to three blocks of 140 trials (70 periodic, 70 aperiodic trials per block) in which the average location of the noise burst sequence was balanced. This was achieved using the method of constant stimuli whereby the probe tone took one of seven different positions relative to the average location of the noise train. Specifically, the probe tone could either be located at the true mean of the sequence, or ±k, ±2k, ±3k dB ILD away from it. This is referred in the rest of this chapter as "Probe tone ILD". The parameter k is participant-specific and represents the 75% correct ILD threshold that was estimated after the calibration procedure. In total, each probe tone position was repeated 30 times within each of the periodic and aperiodic conditions. In between the experimental sessions and blocks participants were allowed to take a short break and remove their headphones. If they chose to, however, they had to repeat the headphone placement procedure used in the calibration session before starting the following block.

Data cleaning: Trials that timed out before a response was made were removed from the dataset. Trials in which the response time (RT) was greater than +3 standard deviations from the mean RT for the particular participant were also removed. In total this trimming resulted in 2.04% of all experimental trials being removed from Expt. 3, and 2.01% from Expt. 4.

Psychometric model fitting: The psychometric data from each participant and rhythmic condition were fitted with sigmoidal Gaussian cumulative density functions. Each function was defined by three parameters: fitted threshold, fitted slope and fixed lapse rate. Guess rates were fixed at 0 (equivalent to the participant pressing the right key on every trial) across all subjects and conditions. The three parameters were fitted separately for each subject and periodicity condition (periodic or aperiodic). Fitted threshold was taken as being the predicted probe tone ILD that corresponded to 50% accuracy on each psychometric function. To test whether there was an effect of periodicity on choice accuracy, the fitted threshold and slope values for each participant were submitted to paired t-tests. This followed common analytical methods used throughout the timing and decision making literatures (Daw, 2011; Rohenkohl et al., 2012; Cravo et al., 2013) - see section 4.3.4 for an overview. The quality of fit for each subject was assessed by correlating predicted values from the best fitting psychometric function with observed accuracy. All fitted functions passed a deviance goodness-of-fit test, with periodic fits having a mean deviance of 4.46 (SD = 3.168) and aperiodic fits having a mean deviance of 5.02 (SD = 3.669). The analysis of the psychometric function was performed using the Palamedes toolbox for Matlab (Prins and Kingdom, 2009).

5.3.6 Apparatus

All phases of the experiment took place in a sound attenuated recording studio in the Centre for Music and Science, University of Cambridge. All stimuli were fully generated and behavioural responses recorded using Psychophysics-3 Toolbox (Brainard, 1997) in addition to custom scripts written for MATLAB (MathWorks). These scripts were written by the author of this thesis and can be downloaded from an online repository:

`https://github.com/dcgreatrex-phd/experiment_3`

Images were presented on a 22-inch Iiyama Prolite E2202WS screen with a vertical refresh rate of 60 Hz, which was positioned 100 cm in front of the participant. Responses were collected via an Apple keyboard with numeric keypad. Sound was heard through a Beyerdynamic DT 990 Pro headset.

5.4 Experiment 3: Results

Decision thresholds: Decision thresholds established during the calibration session ranged from 1.3 dB ILD to 3.3 dB ILD (Mean = 2.53, SD = 0.56). They did not correlate with participant age (r = -0.041, p = 0.858), gender (t = -0.219, p = 0.829), or handedness (t = -1.351, p = 0.193).

5.4.1 Effects on decision accuracy

Figure 5.2 shows the fitted psychometric curves for four out of the twenty-two participants across both the periodic and aperiodic conditions. Each fitted psychometric function regressed the probability that a participant responded Left over each of the seven probe-tone locations relative to the average location of the noise train. Each curve is characterised by a fitted threshold and slope value. The fitted threshold of the curve is the point on the x-axis where the proportion of Left responses is 0.5. A value less than zero represents a leftward bias in the perceived location of the average relative to the probe tone. A value greater than zero represents the opposite. The fitted slope was computed at the point on the curve where the proportion of Left responses is 0.5. Steeper fitted slopes represented more accurate classification of the average noise burst relative to the probe tone location.

As expected, the proportion of Left responses increased as a function of probe tone ILD in both conditions. Participants could thus reliably identify the average location of the noise burst sequence. This was supported by submitting fitted slope values for each periodicity condition to two one-sampled t-tests. Fitted slope values were significantly positive in both conditions indicating that there was a significant main effect of probe tone ILD on Left responses regardless of periodicity (Periodic: $t_{(21)}$ = 14.716, p < 0.001; Aperiodic: $t_{(21)}$ = 15.451, p < 0.001).

Effects of periodicity were assessed by submitting the fitted threshold and slope values for all participants' responses in the periodic and aperiodic conditions to two-tailed t-tests. Neither fitted threshold (Periodic: mean = -0.319, SD = 1.447; Aperiodic: mean = -0.076, SD = 1.681; $t_{(21)}$ = -1.437, p = 0.165) nor slope (Periodic: mean = 0.283, SD = 0.090; Aperiodic: mean = 0.284, SD = 0.086; $t_{(21)}$ = -0.0004, p = 0.999) differed significantly. This implies that the timing of the noise burst stimulus did not systematically bias the accuracy with which participants could compute the average location of the noise burst sequence. This is shown in figure 5.3 in that there is no group level bias towards either of the two periodicity conditions.

Participant ability: Figure 5.4 shows the fitted threshold and slope values by each participant's decision threshold estimate, as estimated during the calibration staircasing session. The differences between each level of probe tone ILD was set to be smaller for participants with a low decision threshold (represented by

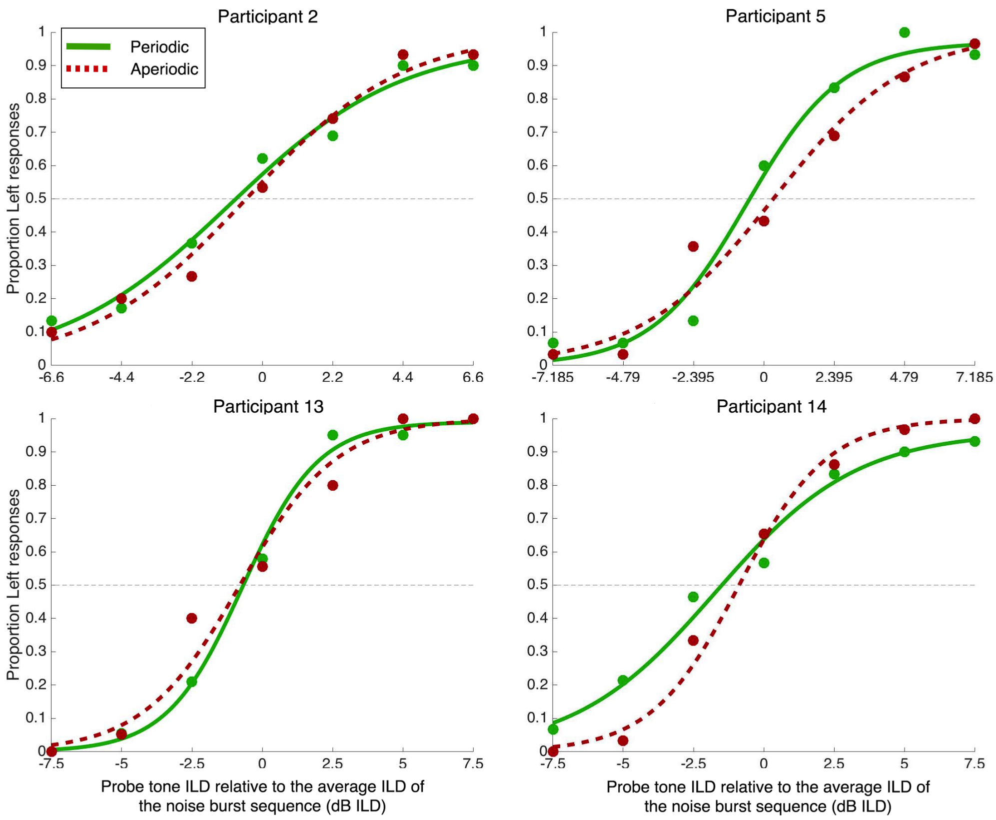


Fig. 5.2 Expt. 3. Fitted psychometric curves for 4 of the 22 participants (Ps 2, 5, 13, 14) showing proportion of Left responses relative to the difference between the ILDs of the probe tone and average of the noise bursts. Green solid curves: periodic stimuli. Red dashed curves: aperiodic stimuli. The four participants were chosen because they share similar fitted threshold values, which facilitates the visual comparison of fitted slopes.

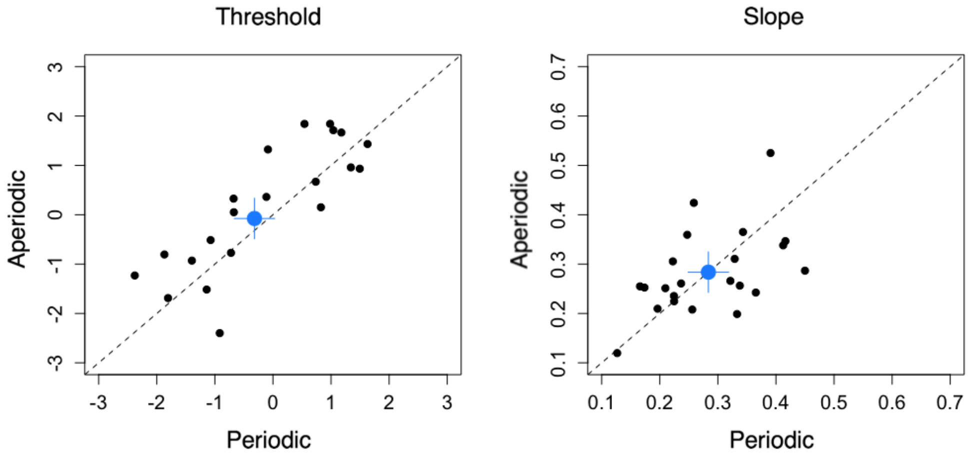


Fig. 5.3 Expt. 3. Scatterplots showing the fitted threshold and slope values for all participants by the periodic and aperiodic conditions. Left panel = fitted threshold values. Right panel = fitted slope values. Blue symbols mark the group average. Deviation from the dotted line indicates a bias towards one of the two periodicity conditions.

the red circles), whose discernment was more precise on the calibration task than others. To test whether overall ability on the task (indexed by decision threshold) interacted with sensitivity towards the periodicity of the sequence, a correlation analysis was made. This comprised the following stages: 1. The absolute difference between fitted slope values on periodic and aperiodic trials was computed for each participant, resulting in one value for each participant. 2. Spearman correlation coefficients were estimated between these absolute differences values and the decision-thresholds. A negative correlation would imply that the timing of the sequence [Periodic, Aperiodic] categorically affected the fitted slope value (and hence accuracy of the decision) for participants with smaller decision thresholds. The correlation was negative, but not significantly so (rho = -0.322, df = 20, p = 0.143). There was also no significant correlation between the fitted threshold values and decision thresholds (rho = 0.346, df = 20, p = 0.115).

Lastly, to test whether there was a bias in fitted slope values towards just one of the periodicity conditions, the above analysis was repeated except that rather than computing the absolute difference between the fitted slope values and the periodicity conditions, the signed difference was used. A negative correlation would imply that participants with smaller decision-thresholds were more accurate on periodic compared with aperiodic trials. Again, the relationship between the signed difference in fitted slope values and decision-threshold was negative, but not significantly correlated (rho = -0.135, df = 20, p = 0.549).

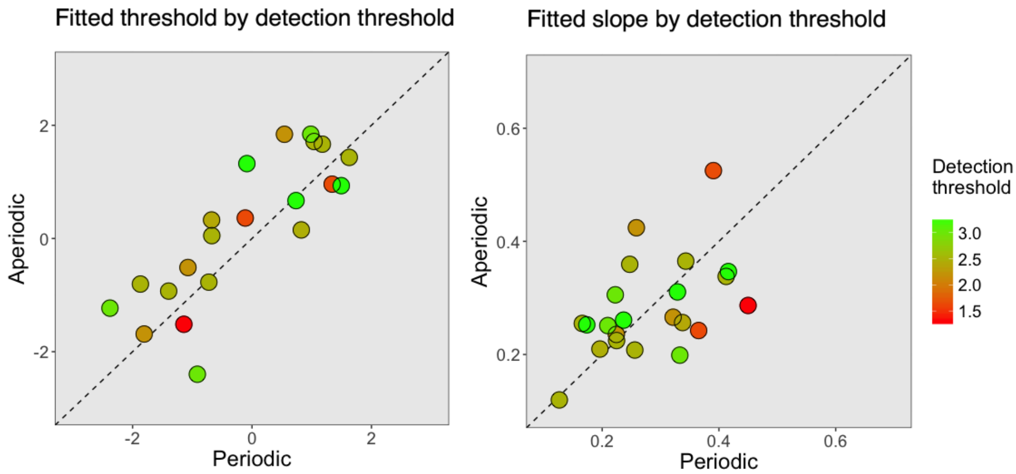


Fig. 5.4 Expt. 3. Scatterplots showing fitted threshold (left) and slope (right) values by decision threshold (colour spectrum). Red circles mark participants with a low decision-threshold. Green circles mark participants with a high decision-threshold. Deviation from the dotted line indicates a bias towards one of the two periodicity conditions.

5.4.2 Effects on response time

Figure 5.5 shows raw RTs by probe tone ILD. Log-transformed response times for correct trials were submitted to a repeated-measures 2 X 7 ANOVA with periodicity [Periodic, Aperiodic] and probe tone ILD [-3*k*, -2*k*, -1*k*, 0*k*, 1*k*, 2*k*, 3*k*] as factors. Note that although participants were forced to respond incorrectly at 0*k* these trials were not excluded from the analysis due to this being a feature of the experimental design and not participant error. As expected there was a main effect of probe tone ILD ($F_{(6,126)} = 19.909$, p $<$ 0.001, partial $n^2 = 0.83$), with response times increasing as the difference between the average ILD of the noise burst sequence and probe tone decreased. Response times were also shorter for sequences with a periodic versus aperiodic rhythm (Periodic: mean = 1.392, SD = 0.166, Aperiodic: mean = 1.421, SD = 0.186, $F_{(1,21)} = 9.356$, p = 0.006, partial $n^2 = 0.31$). The interaction between the factors was not significant ($F_{(6,126)}$ = 1.062, p = 0.389, partial $n^2 = 0.26$).

Rhythmic variability: To test whether response times were sensitive to varying degrees of IOI variance, each participant's data was divided into three IOI variance groups [Zero, Low, High]. These groups were formed by computing the SD of the five IOIs between the noise bursts on each trial. Periodic trials had zero variability by definition (50% of all data). Amongst aperiodic trials, "low IOI variance" trials were those with an IOI SD less than or equal to the 50th SD

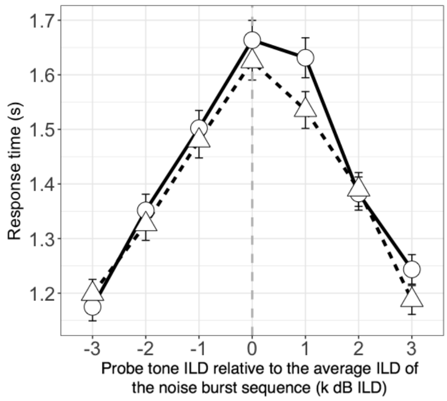


Fig. 5.5 Expt. 3. Mean response time data at each level of probe tone ILD (k dB) and periodicity. Error bars = standard error of the mean.

percentile for each participant (25% of all data), whilst those with SDs > 50th percentile were classed as "High IOI variance" (25% of all data). Figure 5.6 shows a positive correlation between response time and IOI variance, indicating that participants took longer to respond as IOI variance increased. As with the previous analysis log-transformed response times from correct trials were submitted to a one-way repeated-measures ANOVA with IOI variance [Zero, Low, High] as its factor. This was significant overall ($F_{(2,42)}$ = 4.350, p = 0.019, partial n^2 = 0.42), due to high variance sequences being associated with longer response times (Zero: mean = 1.391, SD = 0.782, Low: mean = 1.400, SD = 0.782, High: mean = 1.443, SD = 0.822). Posthoc pairwise comparisons with Bonferroni corrections highlighted however that the finding was restricted to the difference between the zero and high variability conditions and did not occur between other levels of IOI variance.

5.4.3 Mixed-effects regression analysis

A mixed-effect regression analysis was run to support the psychometric analysis described in section 5.4.1. Including all participants data in a single model, rather than modelling each participant's data separately, ensured that multiple sources of variability within the data were accounted for. The analysis was used to determine what effect the absolute location of the sequence had on choice and to confirm whether or not there were any interactions between periodicity and probe tone ILD across participants. The analysis followed procedures laid out by Knoblauch and Maloney (Knoblauch and Maloney, 2012). The analysis consisted of two stages: 1. Different generalized linear mixed effect models (GLMM) with the same fixed

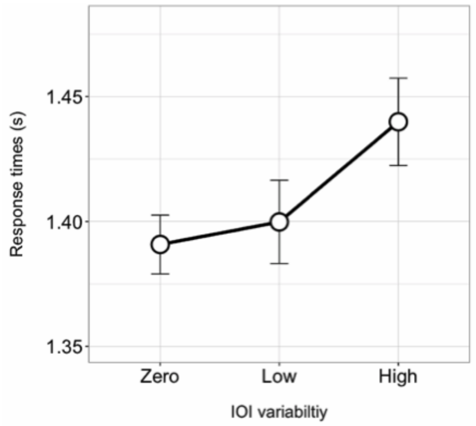


Fig. 5.6 Expt. 3. Mean response time data by IOI variance. Zero IOI variance = periodic trials (50% of all data). Low IOI variance = sequences with an IOI SD less than or equal to the 50th SD percentile for each participant (25% of all data) High IOI variance = sequences with an IOI SD greater than the 50th SD percentile for each participant (25% of all data). Error bars = standard error of the mean.

effects and a variety of random effects were compared to identify the best fitting random factors. 2. Models with the same random effects (determined by stage 1) but differing fixed effects were compared as means of identifying the best fitting overall model. Model selection was done using Likelihood ratio tests (LRTs) and Akaike Information Criterion (AIC) measurements.

Random effects: The starting model contained fixed effects of probe tone ILD [-3*k*, -2*k*, -*k*, 0, *k*, 2*k*, 3*k*], periodicity [Periodic, Aperiodic], average ILD of the noise burst sequence [-4, -2, 0, 2, 4] and all interactions. Random factors of participant, probe tone ILD and block were also included. LRTs were then conducted to determine the most appropriate combination of random factors in the model. The best fitting model was one in which both the intercept and slope of the probe tone ILD varied by participant ID (AIC = 7082.9).

Fixed effects: A backwards stepwise procedure was then used to determine the best fitting fixed effects in the model. The best fitting model had fixed effects of probe tone ILD and average ILD of the noise burst sequence. It also had random effects (both intercept and slope) of probe tone ILD by participant. This was selected using the following procedure: Firstly, the three-way interaction between the fixed effects was removed from Model 1 (= Model 2). A LRT showed that there was no significant difference between Models 1 and 2 (χ^2 = 0.559, df = 1, p = 0.455) and that Model 2 reduced AIC by 1.5. Model 2 was retained. All paired

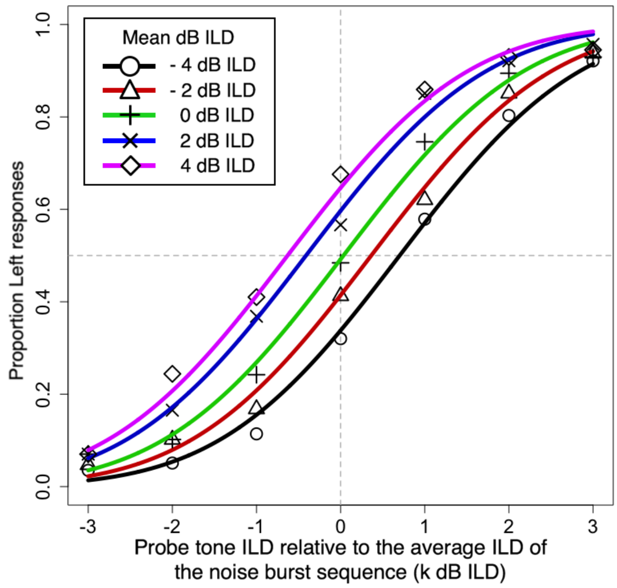


Fig. 5.7 Expt. 3. The predictions of the best fitting mixed-effects model overlaid on average proportion of Left responses by probe tone ILD (x-axis) and the average ILD of the noise burst sequence (colour). The different locations of the curves in relationship to the x-axis indicates that there was a spatial bias associated with the average ILD of the noise burst sequence.

interactions were then removed from Model 2 (= Model 3). There was again no significant difference between Models 2 and 3 (χ^2 = 2.664, df = 3, p = 0.446) and Model 3 reduced AIC by 3.3. Model 3 was retained. The main effect of periodicity was removed from Model 3 (= Model 4). Models 3 and 4 were not significantly different from one another (χ^2 = 1.618, df = 1, p = 0.203) and Model 4 reduced AIC by 0.4. Model 4 was retained. Finally, the average ILD of the noise sequence was removed (= Model 5). This time Model 5 was a significantly worse fit than Model 4 (χ^2 = 312.18, df = 1, $p < 0.001$) and AIC increased by 310.2. Model 5 was rejected and Model 4 retained.

Figure 5.7 shows the best fitting mixed effects model overlaid on the average proportion of Left responses (y-axis) by probe tone ILD (x axis) and average ILD of the sequence (colour). As noted at the start of section 5.4.1, the positive slopes of the curves show that participants were sensitive to the probe tone ILD relative to the noise burst average and could do the task. The different locations of the curves in relationship to the x-axis (i.e. different fitted threshold values) depict the main effect of average ILD of the noise burst sequence that was included in the best fitting model. This meant that there was not always a linear relationship

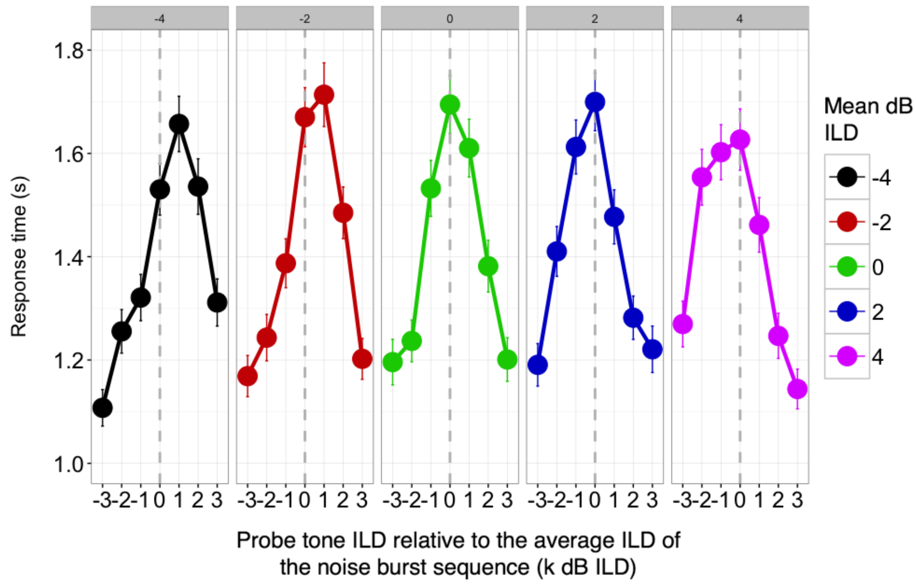


Fig. 5.8 Expt. 3. Mean response times by probe tone ILD (x-axis) and average ILD of the sequence (panels and colours). The reaction time curves appear to be displaced by average ILD of the noise burst sequence. Error bars = standard error of the mean.

between perceived azimuth and ILD and that this varied as a function of the true average ILD of the noise burst sequence. This spatial bias is similar to that previously described as a "spatial aftereffect" and is addressed in more detail in the discussion (Kashino and Nishida, 1998; Carlile et al., 2001; Dahmen et al., 2010).

To test whether RTs also reflected the above spatial bias, log-transformed RTs from correct trials were submitted to a repeated measures ANOVA. Figure 5.8 shows RTs by probe tone ILD and the average ILD of the noise burst sequence. The ANOVA was a 5x5 design containing factors probe tone ILD [-2k,-1k, 0k, k, 2k] and average ILD of the noise burst sequence [-4,-2,0,2,4]. The easiest conditions of probe tone ILD (-3k and 3k) were excluded from the analysis due to the dimensionality of the full model being too high for model stability. There was an expected main effect of probe tone ILD ($F_{(4,80)}$ = 12.824, p < 0.001, partial n^2 = 0.68) whereby RTs increased as probe tone ILD decreased relative to the average ILD of the noise burst sequence. There was also a significant interaction between probe tone ILD and average location ($F_{(16,320)}$ = 8.741, p < 0.001, partial n^2 = 0.91). This meant that negative probe tone ILD relative to the average ILD of the noise burst sequence resulted in shorter RTs when the average ILD of the noise burst sequence was negative and longer RTs when it was positive. Whilst pairwise comparisons indicated that only levels -2k,-1k of probe tone ILD reached statistical significance, a consequence of this pattern is that the longest RT is not always at 0 probe tone ILD.

5.5 Experiment 3: Discussion

The purpose of experiment 3 was to determine whether the key finding from experiments 1 and 2—that rhythmic temporal expectations bias complex averaging decisions—replicates under a less complex experimental design and a faster stimulus rate. The main findings were:

1. Responses were sensitive to IOI variance in the rhythmic sequence and were fastest on periodic trials. Periodicity did not, however, systematically affect response accuracy.

2. Responses were affected by a spatial bias that caused fitted threshold values to vary as a function of the average ILD of the noise burst sequence.

Effects of periodicity: As in experiments 1 and 2, RTs were not only significantly faster following periodic versus aperiodic sequences but were also sensitive to degrees of aperiodicity. Aperiodic sequences with low IOI variance were responded to faster than those with high IOI variance. This finding supports the idea discussed in section 4.4 that rhythmic uncertainty affects response times to a degree proportional to the uncertainty, but does not affect the quality of the information being encoded. This is because, contrary to the experimental hypothesis, periodicity did not affect the accuracy of the decision. This represents a departure from the findings of experiment 1 and suggests that, in experiment 1, the effects of rhythmic temporal expectation on complex decision making was influenced by either task complexity and/or stimulus rate.

If task complexity was the reason why periodicity affected choice accuracy in experiment 1 but not in experiment 3, it would have been caused by one or more of the following changes in experiment 3: 1. The rhythmic sequence was shortened to include 6 rather than 8 sounds. 2. ILDs were limited to a fixed range on each trial and did not include extreme values. 3. The presentational order of ILDs on each trial was controlled and not fully random. 4. The total duration of all sequences were controlled. This complexity explanation can be supported by considering the fact that periodicity did not influence choice accuracy in experiment 2. Experiment 2 differed from experiment 1 in that participants were told in advance about the timing of the rhythmic sequence before starting each trial.

If prior knowledge is considered as another way of reducing task complexity, the results of experiments 2 and 3 could be considered compatible.

A different explanation is that the fast 4 Hz rate interacted with the effects of periodicity and was the reason why periodicity did not affect choice accuracy in this experiment. As discussed in section 5.1, Wyart et al. (2012) showed that during extended categorical decision making the rate of evidence accumulation fluctuated rhythmically. This fluctuation resulted in a decision refractory period which lasted for 250 ms after the onset of salient decision information. This meant that information that fell within the refractory period had a weaker impact on the same choice. As periodic sequences in the current experiment had an IOI of 250 ms, a decision refractory period may have cancelled accuracy enhancing effects of periodicity that would have normally been observed with a lower stimulus rate. A simple way to test for this is to repeat the experiment using a slower stimulus rate, which is done in experiment 4 below (section 5.6). If periodic sequences enhance choice accuracy in experiment 4, the rate explanation will be supported. If not, the complexity explanation will be supported.

Spatial bias: Participants' responses were affected by a spatial bias. This was evident in that the fitted thresholds of the psychometric functions were sensitive to the average ILD of the noise burst sequence. This bias implies that there was not always a linear relationship between perceived azimuth and ILD and can be described as representing one of two things: Either the greater the absolute average ILD of the noise burst sequence, the greater the perceptual bias towards the midpoint (of x), or the greater the perceptual bias of the probe tone away from midpoint. This bias is similar to the "spatial aftereffect" mentioned earlier. Spatial aftereffects are known to occur after prolonged exposure to a preceding sound at a fixed location. This results in the apparent location of a test sound shifting away from its true location in relation to the adaptor sound (Kashino and Nishida, 1998; Carlile et al., 2001; Phillips and Hall, 2005; Maier et al., 2010; Dingle et al., 2010; Dahmen et al., 2010). Phillips and Hall (2005) theorised that aftereffects occur due to circuits serving laterality in one ear becoming selectively "fatigued" following exposure to a lateralised sound source at the same frequency and in the same ear. The authors showed this by having participants decide whether a probe tone was located to the left or right of mid-point (which was marked by three centralised clicks) both before and after an adaptation phase. The adaptation phase consisted of a 5 second long sine tone which had the same frequency as the target tone but was strongly lateralized to either the participant's left or right. The results showed that participants' percept of mid-point was biased towards the location of the adaptor sound after but not before the adaptation phase was administered.

Assuming that the aftereffect described by Phillips and Hall (2005) was the cause of the spatial bias observed in this experiment, it follows that the most likely explanation is that participants' percept of the probe tone was biased away from mid-point and that the noise burst sequence functioned as a spatial adaptor. If this is the case, the finding provides novel evidence that a spatial aftereffect can be induced by a short rhythmic sequence of lateralized noise bursts. As the experiment was fully balanced and used a relative rather than absolute decision task, the spatial bias should not have negatively affected the results of periodicity on choice.

5.6 Experiment 4: When the pulse is slow

Experiment 4 used a slower periodic rate than experiment 3, with the rest of the experiment remaining unchanged. The rate used, a 2 Hz periodic rhythm (IOI of 500 ms), was selected during piloting because it was comfortable to track and attend to. It was also half the frequency of that used in experiment 3 and is slower than the average IOI of 482.69 ms used by the studies reported in table 5.1. It was hypothesised that there would be little difference between experiments 3 and 4 apart from the effects of periodicity being more pronounced and perhaps affecting choice accuracy. This was hypothesised for two reasons: Firstly, a 2 Hz rate is slow enough to avoid the inhibiting effects of the decision refractory described by Wyart et al. (2012) from influencing sequence processing. Secondly, 2 Hz is close to the preferred tempo reported by Moelants (2002) and the average rates of a range of natural movements. For this reason the lateralized stimuli should have been easier to locate, process and respond to.

5.7 Experiment 4: Method

5.7.1 Participants

A total of 17 participants (9 female) took part. All were all students aged between 20 to 31 (mean = 25.18, sd = 3.78) and were paid £10 an hour. All participants had tone detection thresholds of 15 dB HL or better as measured with a Grason-Stadler GSI 16 audiometer at octave frequencies between 250 Hz and 8000 kHz.

At each tested frequency, the thresholds for each ear did not differ by more than 10 dB HL. All participants were right handed and none were practicing musicians. The experiment received ethical approval from the Cambridge Faculty of Music Research Ethics Review Committee.

5.7.2 Auditory stimulus and timing

All stimuli were identical to those used in Expt. 3. The only difference is that the presentational timing of the individual noise bursts were half the speed. Periodic IOIs were always 500 ms and aperiodic IOIs were randomly selected from a range between 250 and 750 ms. For each trial the IOI was randomly drawn from a uniform distribution centered at 500 ms (2 Hz). This distribution had a dispersion equal to 2i times the nominal IOI, whereby $i = 0$ was used on periodic trials and $i = 0.5$ was used on aperiodic trials. See section 5.3.4 for more information about the IOI selection process.

5.7.3 Procedure and design

As the total duration of each trial was longer than in Expt. 3, more blocks with fewer trials were used, to ensure that participants did not get too tired when doing the task and could take adequate breaks. Specifically, each participant performed 6 blocks of 70 trials (35 periodic, 35 aperiodic trials per block) in which the average location and complexity of the sequences were balanced. All other features of the procedure and design were identical to that of Expt. 3 (see section 5.3.5). The software for this experiment can be accessed here:

`https://github.com/dcgreatrex-phd/experiment_4`

5.8 Experiment 4: Results

Decision thresholds: Decision thresholds obtained during the calibration stage of the experiment ranged from 1.4 dB ILD to 3.3 dB ILD (Mean = 2.69, SD = 0.78). See section 5.3.5 for an explanation of decision thresholds. Decision thresholds did not correlate with participant age (r = 0.378, p = 0.135), gender (t = 0.464, p = 0.650), or handedness (t = 1.501, p = 0.156).

5.8.1 Effects on decision accuracy

Figure 5.9 shows individually fitted psychometric functions for four out of the seventeen participants across both the periodic and aperiodic conditions. As in

Expt. 3 the proportion of left responses increased as a function of probe tone ILD in both conditions, indicating that participants could reliably detect the average location of the sequence. This was confirmed using two one-sampled t-tests which showed fitted slope values were on average significantly positive in both the periodic and aperiodic conditions (Periodic: $t_{(16)} = 10.954$, $p < 0.001$; Aperiodic: $t_{(16)} = 14.257$, $p < 0.001$). Figure 5.10 shows the fitted threshold and slope values for all participants by the periodic and aperiodic conditions. As in Expt. 3, two additional two-tailed t-tests highlighted that neither fitted threshold (Periodic: mean = 0.318, SD = 1.532; Aperiodic: mean = 0.505, SD = 1.631; $t_{(16)} = -0.344$, p = 0.733) nor fitted slope (Periodic: mean = 0.334, SD = 0.126; Aperiodic: mean = 0.306, SD = 0.088; $t_{(16)} = 0.754$, p = 0.457) values were significantly different from one another. This suggests that the timing of the noise burst sequence did not systematically bias decision accuracy, regardless of the slower stimulus rate that was used in the experiment.

Participant ability: Figure 5.11 shows the fitted psychometric threshold and slope values by each participant's decision threshold estimate. To test whether task ability (estimated during the calibration session) interacted with the fitted psychometric functions, two correlation analyses were made. This followed the same procedure described in section 5.4.1. This relationship was negative but not significant (rho = -0.144, df = 15, p = 0.582). There was also no correlation when fitted threshold values were used (rho = 0.027, df = 15, p = 0.917). To test whether more able participants were more accurate on periodic versus aperiodic trials, the correlation analysis was repeated using the signed and not absolute differences between fitted slope values on periodic and aperiodic trials. Unlike in Expt. 3, there was a significant negative correlation between the signed difference in slope values and decision thresholds (rho = -0.687, df = 15, p = 0.002). Participants with a lower decision threshold (higher ability) were significantly more accurate on periodic versus aperiodic trials compared with participants with a high decision threshold (low ability). No correlation was found between the signed difference in fitted threshold values and decision thresholds (rho = 0.105, df = 15, p = 0.689).

5.8.2 Effects on response time

Figure 5.12 shows raw RTs by probe tone ILD. As in Expt. 3, log-transformed response times for correct trials were submitted to a repeated-measures ANOVA

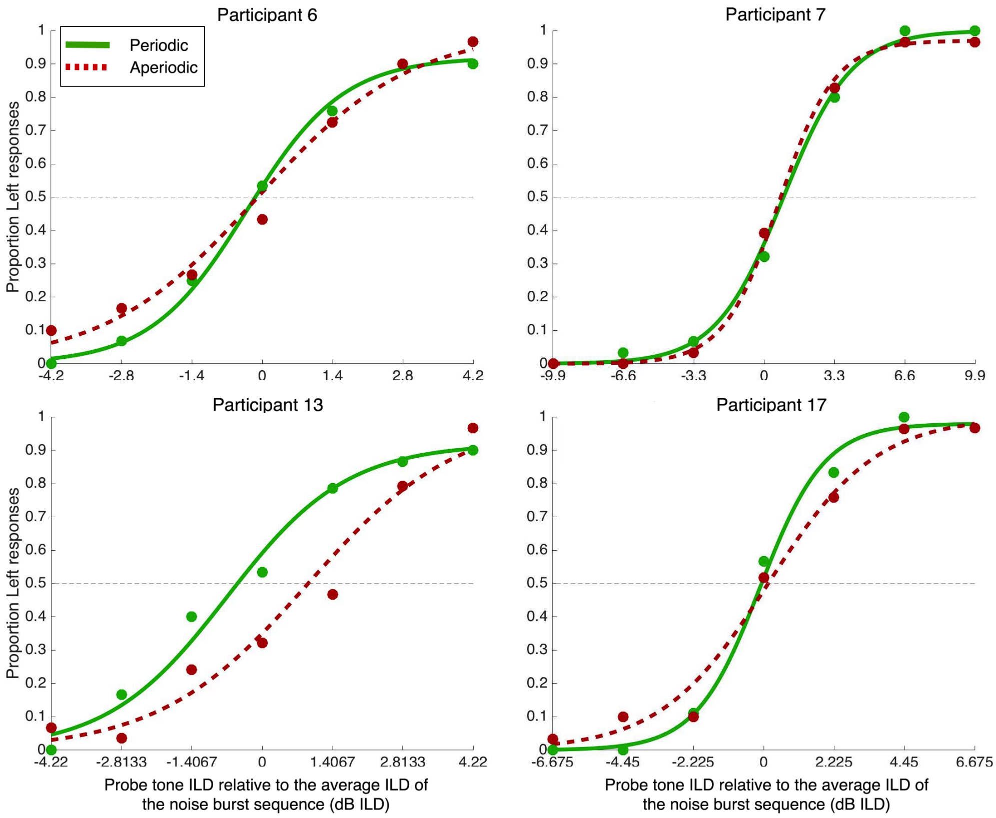


Fig. 5.9 Expt. 4. Fitted psychometric curves for 4 of the 17 participants (Ps 6, 7, 13, 17) showing proportion of Left responses relative to the difference between the ILDs of the probe tone and average of the noise bursts. Green solid curves: periodic stimuli. Red dashed curves: aperiodic. The four participants were chosen because they share similar fitted thresholds, which facilitates the visual comparison of fitted slopes.

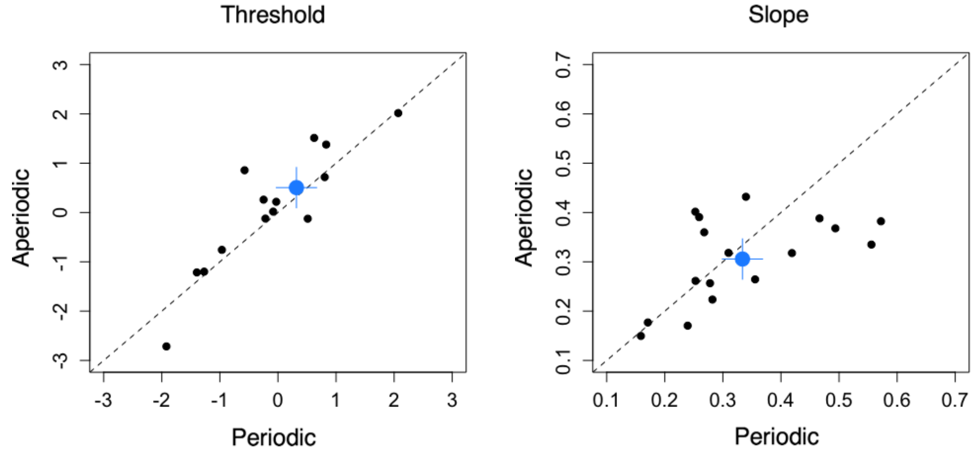


Fig. 5.10 Expt. 4. Scatterplots showing the fitted threshold and slope values for all participants by the periodic and aperiodic conditions. Left panel: fitted threshold. Right panel: fitted slope. Blue symbols mark the group average. Deviation from the dotted line indicates a bias towards one of the two periodicity conditions.

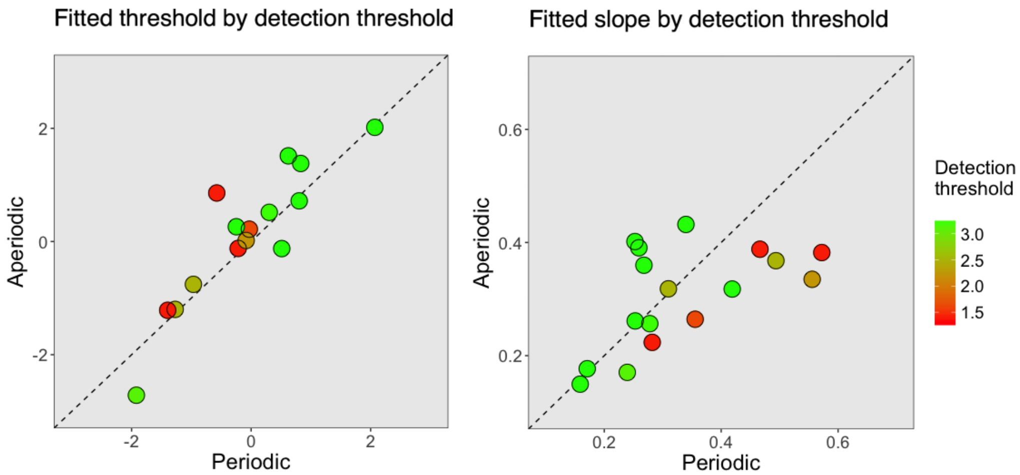


Fig. 5.11 Expt. 4. Scatterplots showing fitted threshold (left) and slope (right) values by decision threshold (colour spectrum). Deviation from the dotted line indicates a bias towards one of the two periodicity conditions.

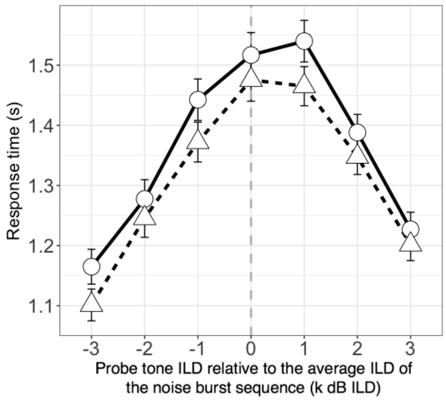


Fig. 5.12 Expt. 4. Mean response time data at each level of probe tone ILD (k dB) and periodicity. Error bars = standard error of the mean.

with probe tone ILD and periodicity as factors. There was again a main effect of both probe tone ILD ($F_{(6,96)} = 13.170$, $p < 0.001$, partial $n^2 = 0.96$) and sequence periodicity (Periodic: mean = 1.316, SD = 0.139, Aperiodic: mean = 1.365, SD = 0.145, $F_{(1,16)} = 16.367$, $p = 0.001$, partial $n^2 = 0.50$). Response times increased as the difference between the ILDs of the probe tone and the sequence average decreased, and they were significantly faster on periodic compared with aperiodic trials. The interaction between the factors was not significant ($F_{(6,96)} = 0.454$, $p = 0.840$, partial $n^2 = 0.17$).

Rhythmic variability: Using the same procedure described in section 5.4.2, each participant's data was divided into three groups according to IOI variance in each rhythmic sequence [Zero, Low, High]. Figure 5.13 shows a positive correlation between response time and IOI variance with longer RTs associated with higher IOI variance. Log-transformed response times from correct trials were submitted to a one-way repeated-measures ANOVA with IOI variance as its factor. As in Expt. 3 there was a main effect of IOI variance with longer responses occurring on high IOI variance trials (Zero: mean = 1.315, SD = 0.700, Low: mean = 1.352, SD = 0.735, High: mean = 1.374, SD = 0.736, $F_{(2,32)} = 6.495$, $p = 0.004$, partial $n^2 = 0.54$), however, this was again only significant between the zero and high variance conditions and not between the other stimulus levels (measured using pairwise comparisons with Bonferroni corrections).

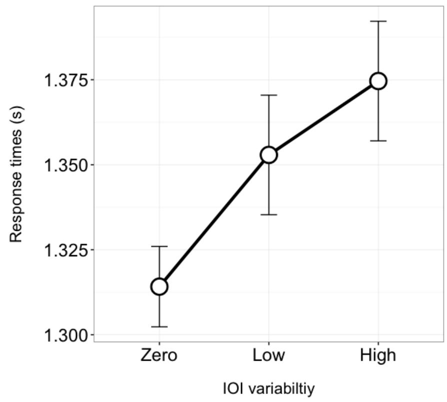


Fig. 5.13 Expt. 4. Mean response time by IOI variance. See figure 5.6 for definitions of IOI variance. Error bars = standard error of the mean.

5.8.3 Mixed-effects regression analysis

As in Expt. 3, a mixed-effect regression analysis was conducted to ensure that multiple sources of variability within the data were accounted for. The starting model contained fixed effects of probe tone ILD [-3k, -2k, -1k, 0, k, 2k, 3k], periodicity [Periodic, Aperiodic], average ILD of the noise burst sequence [-4, -2, 0, 2, 4] and all interactions. It also had random factors of participant, probe tone ILD and block.

Random and fixed effects: Fitted sub-models with different combinations of random factors were compared using likelihood ratio tests (LRTs). As in Expt. 3, the best fitting model was one in which both the intercept and slope of the probe tone ILD varied by each participant (Model 1: AIC = 5078.4). A backwards stepwise procedure was then used to test the significance of fixed effects in the model. The best fitting model contained fixed effects of probe tone ILD, average ILD of the noise burst sequence and a significant interaction between probe tone ILD and periodicity. This was selected using the following procedure: Firstly, the three-way interaction between the fixed effects was removed from Model 1 (= Model 2). A LRT showed that there was no significant difference between Model 1 or 2 (χ^2 = 1.196, df = 1, p = 0.274) and that Model 2 reduced AIC by 0.8. Model 2 was retained. All two-way interactions were then removed from Model 2 (= Model 3). Model 3 fit was significantly worse than Model 2 (χ^2 = 13.454, df = 3, p = 0.004) and AIC increased by 7.4. Fitted coefficients indicated that this was due to a significant interaction between probe tone ILD and periodicity. Model 3 was rejected. The two other interactions (average sequence location by

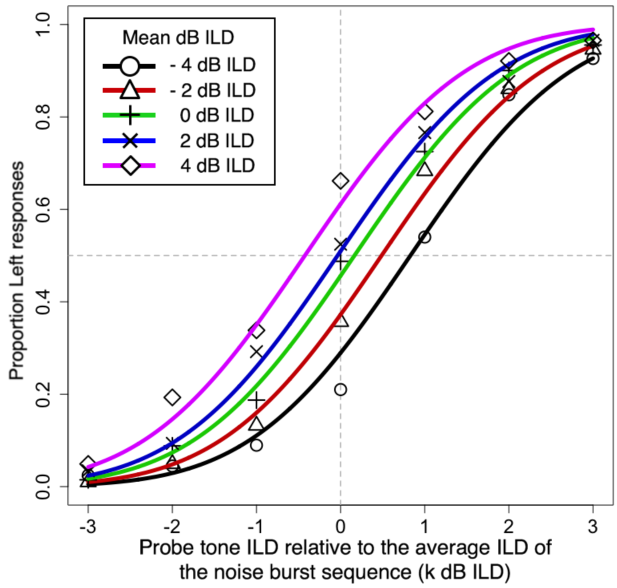


Fig. 5.14 Expt. 4. The predictions of the best fitting mixed-effects model overlaid on average proportion of Left responses by probe tone ILD (x-axis) and the average location of the noise burst sequence (colour). The different locations of the curves in relationship to the x-axis indicates that there was the same spatial bias as described in Expt. 3 - see section 5.4.3.

periodicity and average sequence location by probe tone ILD) were removed from Model 2 (= Model 4). Neither Model 2 or 4 were statistically different from one another (χ^2 = 4.977, df = 1, p = 0.060) and AIC reduced by 1.9. Model 4 was retained.

Figure 5.14 shows the predictions of the best fitting mixed-effects model on the average proportion of Left responses by probe tone ILD (x-axis) and average ILD of the noise burst sequence (colour). Due to the varying locations of the curves in relationship to the x-axis, the same spatial bias described in Expt. 3 occurred in Expt. 4 (see section 5.5). To determine the nature of the interaction between probe tone ILD and periodicity, follow-up comparisons were made. This involved fitting Model 4 separately to each level of probe tone ILD and comparing each fit against a null model in which the interaction with periodicity had been removed. LRTs were used to determine significance. The comparison showed that the interaction between probe tone ILD and periodicity was due to a difference in periodicity conditions when probe tone ILD was equal to zero relative to the average ILD of the noise burst sequence (χ^2 = 3.915, df = 1, p = 0.048) and 1k (χ^2 = 4.175, df = 1, p = 0.041). This meant that participants were more

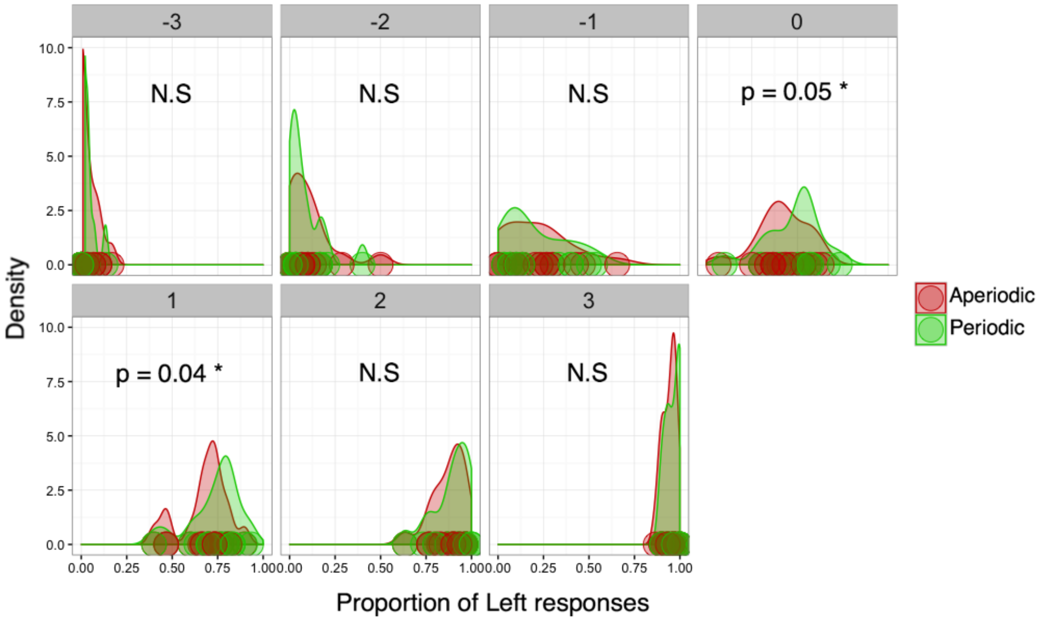


Fig. 5.15 Expt. 4. Density plots of proportion of Left responses by each level of probe tone ILD (panels) and periodicity (green = periodic, red = aperiodic). The difference between the periodicity conditions is only significant at probe tone ILD levels 0 (mid-point on x-axis) and +1.

likely to be correct on periodic compared with aperiodic trials when probe tone ILD was 1k relative to the average. They were also more likely to press the Left response button on periodic trials when the probe tone ILD was 0. As means of plotting a different view of this interaction, figure 5.15 shows probability density functions for all proportion of Left responses by probe tone ILD (different panels) and periodicity (colour).

Lastly, RTs were submitted to the same repeated measures ANOVA described in section 5.4.3. This was to test whether RTs were also sensitive to average sequence location. As in Expt. 3 there was a strong main effect of probe tone ILD ($F_{(4,64)} = 6.428$, $p < 0.001$, partial $n^2 = 0.88$) whereby RTs increased as probe tone ILD decreased relative to the average ILD of the noise burst sequence. There was also a significant interaction between probe tone ILD and average sequence location ($F_{(16,256)} = 3.598$, $p < 0.001$, partial $n^2 = 0.99$), however follow-up comparisons showed this not to be significant at any level of probe tone ILD.

5.9 Experiment 4: Discussion

Experiment 4 used a slow stimulus rate and the same experimental design as experiment 3 to investigate whether rhythmic temporal expectations are influenced by stimulus rate during complex averaging decisions. The main findings with respect to those of experiment 3 were:

1. RTs were not affected by rate of stimulus presentation. In both experiments, they were sensitive to IOI variance and fastest on periodic trials.
2. Decision accuracy was not affected in a simple way by periodicity. When the rate was slow (experiment 4) high ability participants made more accurate decisions on periodic trials and accuracy increased generally on a subset of difficult stimulus conditions. When it was fast (experiment 3) periodicity had no effect on accuracy.
3. The same spatial bias was observed at both rates of presentation.

Effects of periodicity: Periodicity affected response times in much the same way as in experiment 3. RTs were fastest on periodic trials and sensitive to IOI variance in the rhythmic sequence. This strengthens the findings of experiments 1, 2 and 3. It implies that rhythmic uncertainty affects response times, proportional to the uncertainty, but is robust to changes in average stimulus rate (at least within the range of 2 Hz and 4 Hz). The major difference between experiments 3 and 4, however, was that in experiment 4 the mixed effect model identified an interaction between periodicity and two levels of probe tone ILD [0k and +1k] with regards to choice selection. This suggests that periodicity did affect the outcome of participants decisions, but that it was limited to only a small subset of difficult stimulus conditions. One explanation for this findings is that the rate of the stimulus in experiment 4 was longer than a decision refractory period described by Wyart et al. (2012) and therefore more rhythmically weighted information was processed in experiment 4 compared with experiment 3.

A second difference between experiments 3 and 4 was that there was a significant correlation between the signed difference of fitted slope values between the periodicity conditions and participants' decision thresholds. This meant that participants with low decision thresholds were more likely to make more accurate decisions on periodic versus aperiodic trials on experiment 4. A low decision threshold meant that a participant successfully completed the task under hard conditions during the calibration session and that the ILD difference between a single level of probe tone ILD was set to be relatively small in the main experiment. As this effect was not found in experiment 3, a slower stimulus rate was presumably the cause of this difference. If true, a good question to ask is why does the data not look like that of experiment 1? In experiments 1, a 3 Hz rhythm was used and there was a main effect of periodicity on choice accuracy. One reason may have been the difference in task complexity between experiments 1 and 4. As discussed in section 5.5, experiments 3 and 4 were designed to be less complex than that of experiments 1 and 2. Therefore, a reduction in task complexity may have helped to diminish the effects of periodicity on decision accuracy, but the slower stimulus rate may have ensured that periodicity still affected choice outcome on difficult trials. One way to test this in a future experiment is by replicating experiment 4 except with increased task complexity. The effects of periodicity on choice accuracy should disappear if task complexity is the cause of experiment 4's finding. Task complexity could be increased by either including more noise bursts in each stimulus sequence, or by presenting the noise bursts in more complicated orders.

Spatial bias: The same spatial bias was observed in experiment 4 as in experiment 3 - see section 5.5 for a discussion. Importantly, this suggests that the spatial bias is robust enough not to be a consequence of the rate of the rhythmic sequence.

5.10 Experiments 3 and 4: Summary

Experiments 3 and 4 tested two research questions: Firstly, whether key findings of chapter 4 replicate under a different experimental paradigm that tests the same fundamental research question. Secondly, whether the rate of rhythmic stimuli interact with the effects of rhythmic temporal expectations during complex averaging. Both questions were investigated using a less complex experimental design than used in experiments 1 and 2 in an attempt to increase the interpretability of the experimental data. The data showed that, in both experiments, the effects of periodicity on RTs replicated that of experiments 1 and 2. Participants made faster decisions on periodic trials and slower decisions on aperiodic trials with high IOI variance. RTs also increased as probe tone ILD decreased relative to the average ILD of the noise burst sequence indicating sensitivity towards task difficulty. The only large difference between RTs in experiments 3 (fast rate) and 4 (slow rate) is that they were on average slower in experiment 3 than in experiment 4 (Periodic mean: Expt. 3 = 1.391 s, Expt. 4 = 1.316 s; Aperiodic mean: Expt. 3 = 1.421 s, Expt. 4 = 1.365 s). This suggests that participants found information contained in the slower 2 Hz rhythmic rate easier to process and respond to compared with the faster 4 Hz rate.

Unlike in experiment 1, periodicity did not strongly bias decision accuracy in either experiment 3 or 4. There was no effect of periodicity on decision accuracy in experiment 3, and periodicity only affected decision accuracy for a subset of high ability participants and hard stimulus types in experiment 4. This difference suggests that stimulus rate and/or task complexity interacted with the effects of rhythmic temporal expectations during complex averaging. The fact that there was a difference between experiments 3 and 4 in this measure implies that the fast 4 Hz stimulus rate may have cancelled the effects of periodicity on choice accuracy. This is because, apart from a difference in stimulus rate and consequent changes in block length, the two experimental designs were identical. An unexpected spatial bias was also observed in both experiments regardless of periodicity whereby decisions were biased depending on the average ILD of the noise burst sequence. The most likely explanation, based on publications that investigate "spatial aftereffect", is that participants' percept of the probe tone was biased away from mid-point the greater the absolute average ILD of the noise burst sequence. If true, this finding provides novel evidence that a short rhythmic sequence of lateralized noise bursts with onsets that are perceptually distinguishable can function as an adaptation stimulus at multiple rates.

To conclude, the timing of the rhythmic stimulus had a large effect on decision latencies during complex averaging but did not appear to systematically affect the quality of the information being encoded. That said, choice accuracy was enhanced on periodic trials in experiment 4 for a subset of high ability participants, as well as on a subset of stimulus levels. These findings suggest that the slower presentational rate increased the likelihood that periodicity would affect choice accuracy, but that it did not influence how rhythmic temporal expectations affected response times.

Chapter 6

Rhythmic variability (Experiment 5)

6.1 Experiment 5: Introduction

Whereas experiments 1 to 4 used a complex decision task and controlled for the precursor's overall duration, experiment 5, reported in this chapter, was designed with three aims in mind. 1. To determine whether RTs are sensitive to IOI variance in a rhythmic precursor during an experimental task that requires a simple rather than a complex decision. 2. To determine whether the duration of the rhythmic sequence interacts with effects of IOI variance on response time. 3. To determine whether effects of IOI variance are robust or diminish with task repetition.

The first aim assesses the generality of the IOI variance findings of experiments 1 to 4, and is expected to contribute towards determining whether the cognitive processes responsible seem general to a range of decision types, or just to complex decisions. Whilst there have been no explicit attempts in the timing or decision making literatures to answer this question, there is a strong probability that the findings will generalise to simple decisions. This is because Mathewson et al. (2012) and Herrmann et al. (2016) have both shown that degrees of aperiodicity affect responses in simpler psychophysical tasks than experiments 1 to 4 (see section 3.2.4 for a discussion), albeit using methods that differ from those of this thesis.

The second aim should lend insight into the range of precursor durations required for sensitivity towards IOI variance to be observable. This question is investigated because the duration of the rhythmic sequence used on each trial has not yet been addressed in this thesis, yet would enhance its aim of achieving ecological validity. Natural rhythmic stimuli do not usually share the same duration and there is a large degree of variability between the durations of rhythmic events. A secondary reason for investigating precursor duration is that it has been widely shown that the duration between a warning signal and response stimulus, known as the "foreperiod", strongly affects reaction times (see Niemi and Näätänen, 1981, for a review). When the foreperiod is constant across experimental trials, short (though not extremely short) foreperiods are generally associated with faster reaction times and long foreperiods with slower reaction times (Woodrow, 1914; Klemmer, 1956). The opposite is true when the foreperiod is unknown and varies across experimental trials (i.e. in situations of high temporal uncertainty) (Woodrow, 1914; Davis, 1965). In these situations short foreperiods generally correlate with longer reaction times and long foreperiods with shorter reaction times. These findings highlight the strong effect that temporal uncertainty, prior knowledge and the conditional probability of the stimulus occurring over time has on reaction times. If these effects are associated with processes of motor preparation and response, there is a possibility that the total duration of the precursor sequence may determine what degree participants are sensitive to IOI variance.

The third aim assesses whether task learning reduces or removes effects of IOI variance on response times. This is important to test because it will help to determine how robust effects of IOI variance are on motor preparation and whether it can be suppressed or unlearned. Whilst IOI variance is a novel topic in this experimental setting, effects of task repetition and learning on response times can be easily tested by including experimental block as a factor in the analysis.

The simpler task comprised pressing a response key when a high-pitched tone was heard, following a rhythmic precursor sequence of noise bursts whose overall duration was relatively unpredictable. Thus experiment 5 was a simple detection task making minimal demands on memory and cognitive processing. If participants' RTs are affected by IOI variance it will imply that the findings of experiments 1 to 4 were caused by processes that are relatively general to perception-action tasks. If an effect of IOI variance is not found, it will suggest that the findings of experiments 1 to 4 are specific to complex decision making. It was hypothesised that IOI variance would affect response times in ways similar to that of experiments 1 to 4, and that the duration of the rhythmic sequence used on each trial would interact with effects of IOI variance on response. The latter is compatible with attentional entrainment models (see section 2.2.1) and implies

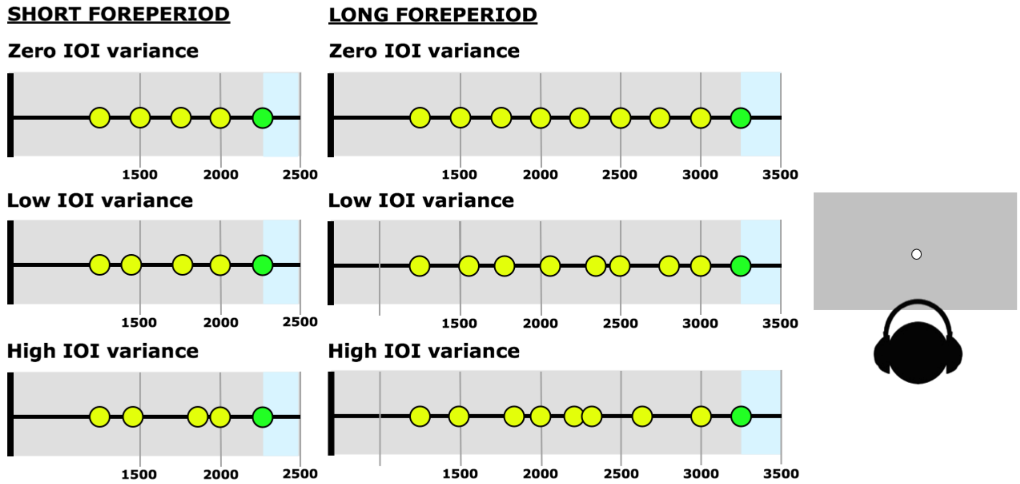


Fig. 6.1 Expt. 5. Schematic illustration of task structure: each trial began with silence lasting anywhere between 1000 and 1500 ms. This was followed by an auditory sequence of centrally located noise bursts that was between 4 to 10 bursts long (yellow circles), followed by a 2 kHz probe tone (green circles). The IOI between the final noise burst and probe tone was fixed at 250 ms. Participants were to respond by pressing a response key as soon as they heard the probe tone. IOI variance was manipulated in three levels: zero (top panels), low-aperiodic (middle panels), high-aperiodic (bottom panels). For the duration of each trial participants were shown a grey screen with a fixation point at mid-point (rightmost panel).

that prolonged exposure towards a rhythmic sequence is required in order for its statistical qualities to influence response.

6.2 Experiment 5: Method

6.2.1 Participants

A total of 13 participants (9 female) took part. All were students aged between 21 and 28 (Mean = 24.3, SD = 2.46) and were paid £10 an hour. All but one were right handed and none were practicing musicians. All reported to have normal hearing. The experiment received ethical approval from the Cambridge Faculty of Music Research Ethics Review Committee.

6.2.2 Auditory stimulus

Noise burst sequence: A rhythmic sequence of short noise bursts was used on each trial. This followed the same principles described in Expts. 3 and 4 (see section 5.3.2), except that (as explained in the introduction of this chapter) the total duration of the sequence was not fixed and the number of sounds in each sequence varied across trials. Each sound was centrally located and was not spatially lateralized.

Probe tone: The probe tone was identical to that used in Expts. 3 and 4 but was not spatially lateralised. It always sounded 250 ms after the final noise burst on each trial.

6.2.3 Timing

Foreperiod: The duration between the onset of the first noise burst in the rhythmic sequence and probe tone was manipulated across trials by varying the number of noise bursts in each sequence. Seven durations were used in total with each sequence containing between four to ten bursts of noise. Sequences containing four, five and six noise bursts were classified as having short foreperiods (the term used in this chapter to represent total precursor duration). Sequences containing seven, eight and nine noise bursts were classified as having long foreperiods. To prevent participants learning the maximum number of possible noise bursts on each trial and then using this knowledge to anticipate, rather than to react to the onset of the probe tone, there were also sequences that contained 10 noise bursts. These ensured that all test foreperiods (sequences containing four to nine noise bursts) were associated with a temporal uncertainty of 50% or less. Each of the seven foreperiods was presented with equal likelihood and randomised within each experimental block. Selecting foreperiods from a rectangular distribution increased the likelihood that the subjective probability for each foreperiod was equal at the beginning of every trial (Niemi and Näätänen, 1981).

Rhythm: Three distinct rhythm types were used to investigate the effect of IOI variance on reaction times: "Zero IOI variance" (periodic), "Low IOI variance" and "High IOI variance". Periodic IOIs were always 250 ms (4 Hz), the same as the periodic IOI in Expt. 3. Aperiodic IOIs were randomly selected from a range between 125 and 375 ms. This range was determined using the temporal jitter procedure described by Goupell et al. (2009) and Brown and Stecker (2011). This meant that for each trial, each IOI was randomly drawn from a uniform distribution centered at an IOI of 250 ms, with a dispersion equal to $2i$ times the 250 ms. Periodic sequences were thus generated with $i = 0$ and aperiodic sequences were generated with $i = 0.5$. The number of IOIs selected for each sequence was determined by the length of the foreperiod on each trial.

To differentiate between low and high IOI variance, a second constraint was imposed on aperiodic stimuli after an array of IOIs, used to construct the foreperiod sequence, had been selected. For low IOI variance trials, the SD of the IOIs within the selected array needed to be within -2.5 and -1 SD of all possible SD values that could be generated using this method. Likewise, for high IOI variance the SD of the IOIs within the selected array needed to fall within +1 and +2.5 SD of all possible SD values. The method used to determine which values of SD corresponded to these low and high variance categories is as follows: 1. randomly generate 10,000 aperiodic IOI arrays for each foreperiod length, 2. compute the SD for each IOI array, 3. select the SD values corresponding to -2.5, -1, +1 and +2.5 SDs of the underlying SD distribution (10,000 data points), 4. repeat this process for each foreperiod length used in the experiment to create an array of SD cut-off values. 5. take the average of each corresponding cut-off value. The resulting average values were 0.035 (-2.5 SD cut-off), 0.059 (-1 SD cut-off), 0.090 (1 SD cut-off) and 0.114 (2.5 SD cut-off). Using this method, low IOI variance arrays were resampled until one was selected that had a SD of between 0.035 and 0.059. High IOI variance arrays were resampled until one was selected that had a SD of between 0.090 and 0.114.

Three additional controls were used to ensure that the stimuli shared similar qualities to that of Expt. 3 and 4. These were the same as those used in Expt. 3 in that: 1. Aperiodic IOIs were resampled if their value was within ±25 ms of the preceding IOI in the sequence. 2. Resampling ensured that the total duration of the aperiodic sequence was always the same as the total duration of its corresponding periodic sequence, with a variability of ±5 ms. 3. The IOI between the final noise burst of the sequence and the probe tone was always 250 ms on both periodic and aperiodic trials.

6.2.4 Design and procedure

The task was administered within a sound attenuated recording studio located in the Centre for Music and Science at the University of Cambridge. The experimental delivery and data collection were controlled by a Matlab program written by the author of this thesis:

`https://github.com/dcgreatrex-phd/experiment_5`

Participants were told that the experiment tested their ability to detect and respond to high pitched sounds and required listening to rhythmic sequences via headphones. They were instructed that on each trial they would hear a rhythmic sequence followed by a high pitched probe tone. Their task was to press a response key as quickly as possible as soon as they heard the tone.

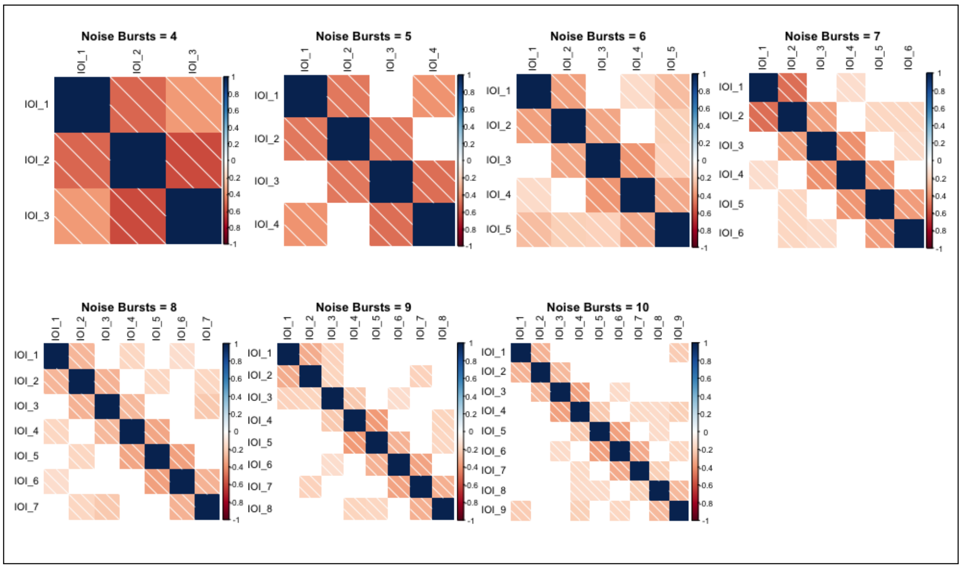
On each trial of the experiment participants were shown a grey screen containing a small white circle located at mid-point (radius = 10 px). They were told to fixate on the white circle and to avoid looking at different areas of the room whilst doing the task. Participants were then instructed to rest their index finger of their dominant hand on the response key ("B" on an Apple computer keyboard) and to maintain this position throughout the entire experiment. All trials began with a period of silence lasting anywhere between 1 and 1.5 seconds. This method mimicked the task structure of similar experiments in the decision making and timing literatures (Carpenter et al., 2009; Lawrance et al., 2014). It was long enough for participants to anticipate the forthcoming sequence, yet variable enough not to elicit accurate expectations concerning the onset of the rhythmic sequence. The noise burst sequence, probe tone and response period then followed. As soon as the participant pressed the response key, the central fixation point flashed green for 50 ms and then back to white. The next trial began 200 ms later. If participants responded before the onset of the probe tone, they were shown the message "You responded too early!" and the trial was aborted. This message appeared on-screen for 1.3 seconds and was followed by the same 200 ms pause as above. It was included to highlight to participants that they anticipated the onset of the target rather than reacted to it. Aborted trials were not replaced due to the relatively large number of trials in the experiment.

The experiment proper contained 8 blocks of 63 trials (21 periodic, 21 aperiodic low IOI variance, 21 aperiodic high IOI variance) and lasted on average 35 minutes. These were preceded by 21 practice trials. The practice session consisted of three repetitions of the six test foreperiods and boundary condition: rhythmic variability [Zero, Low, High] x foreperiod [3–Short, 3–Long, 1–Boundary], presented in randomised order. Participants were aware that the practice trials were not part of the main experiment and were allowed to take a short break prior to starting the experiment proper. All experimental conditions in both the practice and experiment proper were fully randomised within each block.

6.2.5 Random sampling

As random sampling was used to generate the aperiodic rhythmic stimuli, it is important to check whether this method contributed to unexpected correlations

Low IOI variance trials



High IOI variance trials

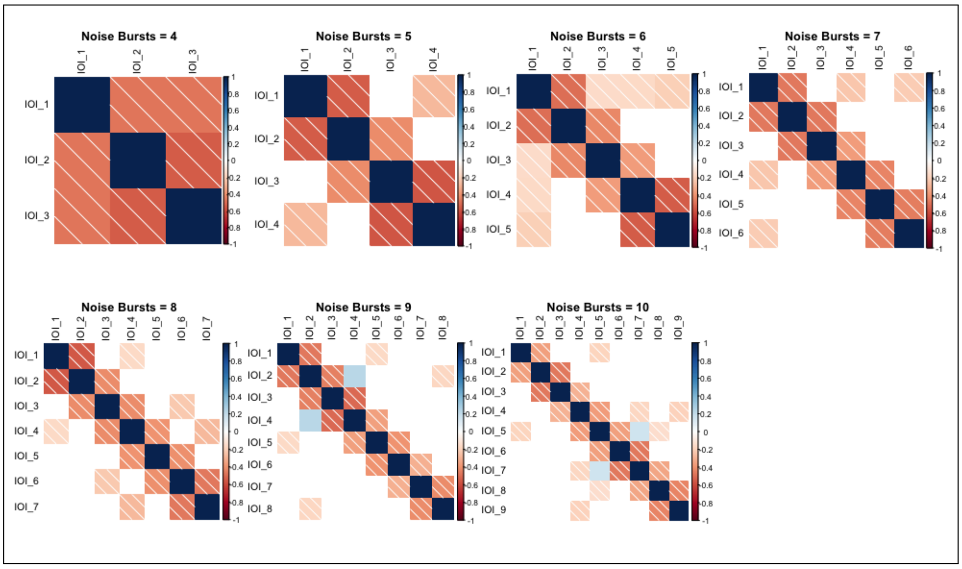


Fig. 6.2 Expt. 5. Correlation matrices showing relationships between the randomly sampled IOIs used in Expt. 5's precursor sequences. (IOI_1 = the first IOI in the sequence). Correlations with a p-value less than or equal to 0.01 are represented by the coloured squares: red = negative correlation, blue = positive correlation. The difference in red hues represents the strength of the negative correlation: dark hue = strong correlation. light hue = weak correlation. Top panel - Low IOI variance trials. Bottom panel - High IOI variance trials.

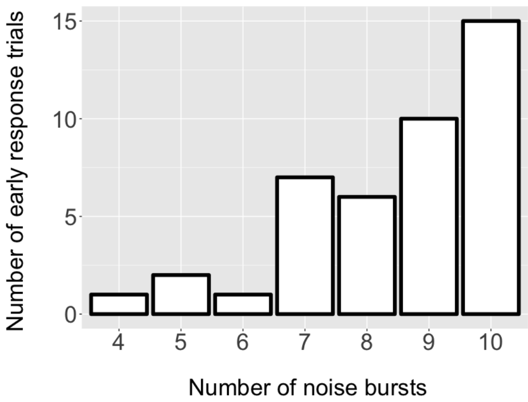


Fig. 6.3 Expt. 5. Number of early response trials by number of noise bursts in the precursor sequence for all participants. Participants were more likely to respond early as the number of noise bursts in the precursor sequence increased.

in the stimulus that may have negatively affected the response data. Figure 6.2 shows correlation matrices for all of the IOI values used in both low IOI variance (top panel) and high IOI variance (bottom panel) trials in the experiment. Correlations with a p-value less than or equal to 0.01 are represented by coloured squares: red = negative correlation, blue = positive correlation, with darker reds representing stronger negative correlations compared with light reds. The matrices show significant correlations between IOIs in all of the aperiodic conditions. These were expected and caused by the IOI sampling rules used when generating the stimulus (see section 6.2.3). Specifically, there were mostly negative correlations between neighbouring IOIs (evidence in the diagonal pattern running from top left to bottom right in each panel), and there were relatively more correlations on shorter precursor sequences. The most important point is that the patterns are similar between the low and high aperiodicity conditions (top versus bottom panels). This suggests that any differences found in responses between the two conditions will likely be the result of IOI variance and not trial-by-trial stimulus sampling.

6.2.6 Data preparation and outliers

Of the original 6552 trials, 65 (1%) were removed. Of these 65 rejected trials, 42 were removed because participants pressed the response key before the onset of the probe tone. Figure 6.3 shows that participants were more likely to respond early as the number of noise bursts in the sequence increased. 7 trials were removed due to RTs exceeding 1 second and 16 trials removed due to RTs being less than 150 ms. RTs longer than 1 second were assumed to represent a loss of attention and to be the result of a different cognitive process than the one under investigation (Ratcliff, 1993). RTs less than 150 ms were assumed to be the result of well timed anticipation rather than reaction towards the response stimulus. This cut off value was used because 150 ms is claimed to be the minimum RT threshold when foreperiod duration is unknown (Niemi and Näätänen, 1981). In the remaining 6487 trials, the minimum and maximum response times across all participants were 151 ms and 868 ms respectively.

As Shapiro-Wilk tests revealed that no participant's RT distributions were normally distributed, all RTs in the data were transformed using both the log-normal ($log(RT)$) and inverse transformation ($-1/RT$) for the purpose of statistical analysis. Further normality testing indicated that the distributions of the inverse transformation were closer to normal than the log-normal, so the inverse transformation was retained as the dependent variable in the analysis. Regardless of this transformation, six of the participant-specific distributions remained not normally distributed. To allow for this, two separate datasets were made for comparison in the ANOVA, following Ratcliff (1993) and Baayen and Milin (2015) who both recommend using the inverse transformation and multiple datasets in RT analysis: The first dataset (DS1) contained all remaining trials and used the inverse RT transformation as a dependent variable. The second dataset (DS2) was a subset of the first, containing only the 6276 trials in which the inverse RT transformations were within 2.5 SD of the mean for each participant. The 211 (3.22%) trials were removed manually so that participants' RT distribution best met the assumptions of normality.

6.3 Experiment 5: Results

6.3.1 Reaction time analysis

Each of the two datasets (DS1 and DS2) were subjected to a 2x3x2 repeated measures ANOVA, with factors foreperiod [Short, Long], IOI variance [Zero, Low, High] and session [Blocks 1:4, Blocks 5:8]. As the same pattern of significant F-values were found in both datasets (section 6.2.6) only the statistical results associated with the reduced dataset are reported. This is because the findings associated with this dataset should be more reliable due to participant's data being normally distributed. In order to aid interpretability of the data, however, all charts and summary statistics relating to the analysis were made using non-transformed RT data.

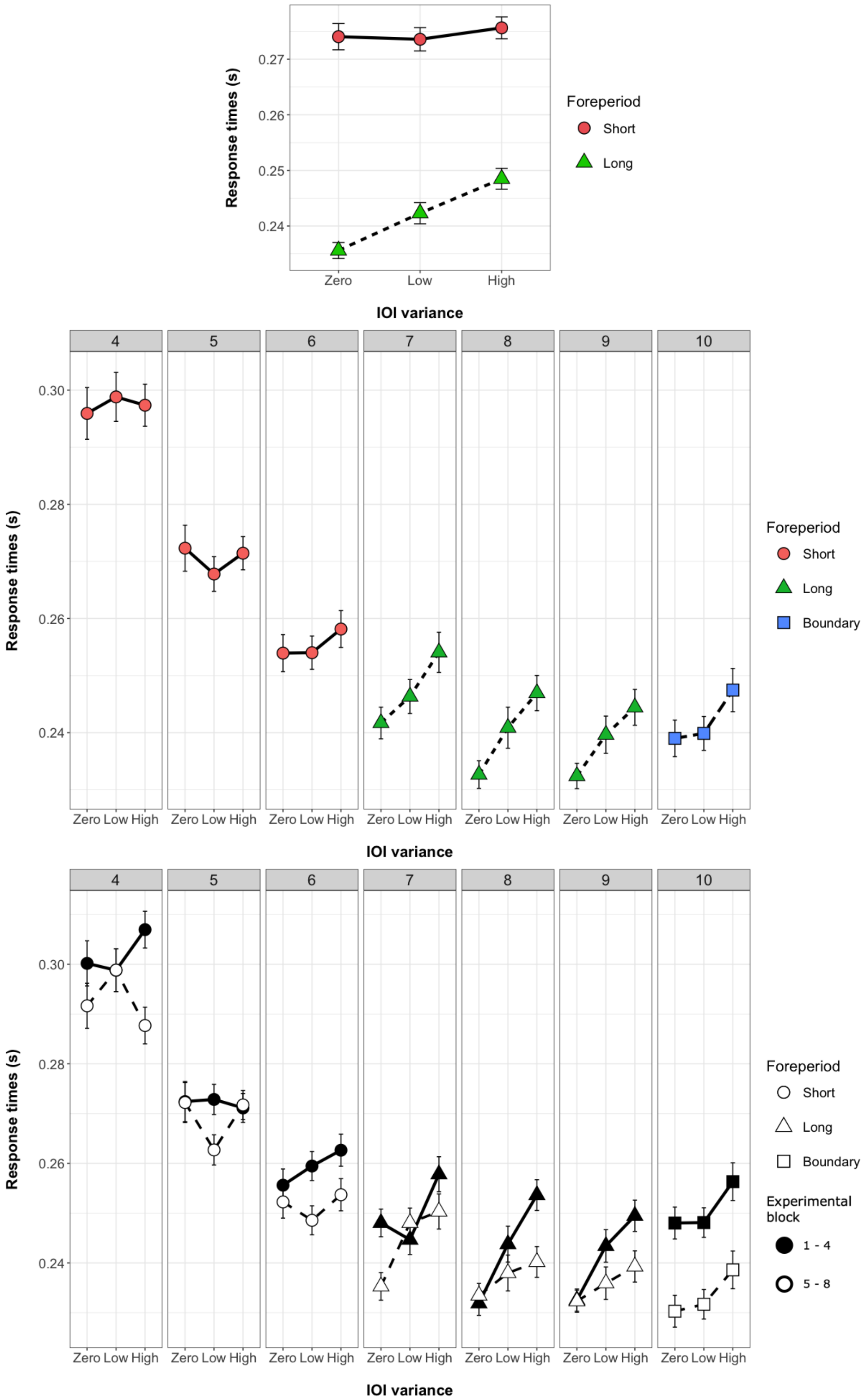


Fig. 6.4 Expt. 5. Three groupings of averaged non-transformed RT data from dataset 1: Top panel: mean RTs by IOI variance (x-axis) and foreperiod (colour and shape). Middle panel: mean RTs by IOI variance (x-axis), foreperiod (colour and shape) and number of noise bursts in the precursor sequence (panels). Bottom panel: as in the middle panel, with the addition of experimental session (black fill = blocks 1:4, white fill = blocks 5:8).

Figure 6.4 shows three groupings of the averaged non-transformed RT data from DS1. Significant differences were found in all of the main experimental conditions. RTs were longer on short compared with long foreperiod trials (Short: mean = 0.27, SD = 0.06, Long: mean = 0.24, SD = 0.05, $F_{(1,12)} = 149.358$; $p < 0.001$, partial $n^2 = 0.92$). RTs also positively correlated with IOI variance (Zero: mean = 0.25, SD = 0.06, Low: mean = 0.255, SD = 0.06, High: mean = 0.26, SD = 0.06, $F_{(2,24)} = 13.905$; $p < 0.001$, partial $n^2 = 0.69$), and were shorter in the second half of the experiment compared with the first half (Session 1: mean = 0.255, SD = 0.025; Session 2: mean = 0.247, SD = 0.020, $F_{(1,12)} = 8.250$; $p = 0.014$, partial $n^2 = 0.40$). There was also a significant interaction between foreperiod and IOI variance ($F_{(2,24)} = 4.581$; $p = 0.021$, partial $n^2 = 0.44$), but no interaction between foreperiod, rhythmic variability and session ($F_{(2,24)} = 2.959$; $p = 0.071$, partial $n^2 = 0.07$). Two follow-up one-way ANOVAs highlighted that this was caused by RTs being sensitive to rhythmic variability on long ($F_{(2,24)} = 13.859$; $p < 0.001$, partial $n^2 = 0.64$) but not short foreperiod trials ($F_{(2,24)} = 2.412$; $p = 0.111$, partial $n^2 = 0.29$). On long foreperiod trials, pairwise comparisons with Bonferroni corrections showed that RTs slowed significantly at each level of IOI variance (Zero variability: mean = 0.235, SD = 0.021; Low variability: mean = 0.239, SD = 0.022, High variability: mean = 0.244, SD = 0.025; Zero/Low p = 0.05, Zero/High $p < 0.01$, Low/High p = 0.04). Finally, although the boundary condition was excluded from the statistical analysis (square points in the middle and bottom panels of figure 6.4), it appears to have been responded to differently compared with the other foreperiod conditions, thus supporting the inclusion of the boundary condition in the experimental design (section 6.2.3).

6.3.2 Mixed effects analysis

One limitation with the above ANOVA is that it overlooks temporal discrepancies that may exist between trials. RTs are known to correlate with those of recently completed trials and participants develop response strategies as the experiment progresses (Baayen and Milin, 2015). Figure 6.5 provides evidence of these effects in the current data. The plot shows that there were significant autocorrelations at short trial lags across the majority of participants (all except 2, 4, 9, 10) and across a much wider span of lags for participants 3, 5, and 11. This means that for some participants, some observed RTs were not independent of earlier ones. A second limitation of repeated measures ANOVAs is that they do not account for the same degree of between participant variability as a linear mixed effect model

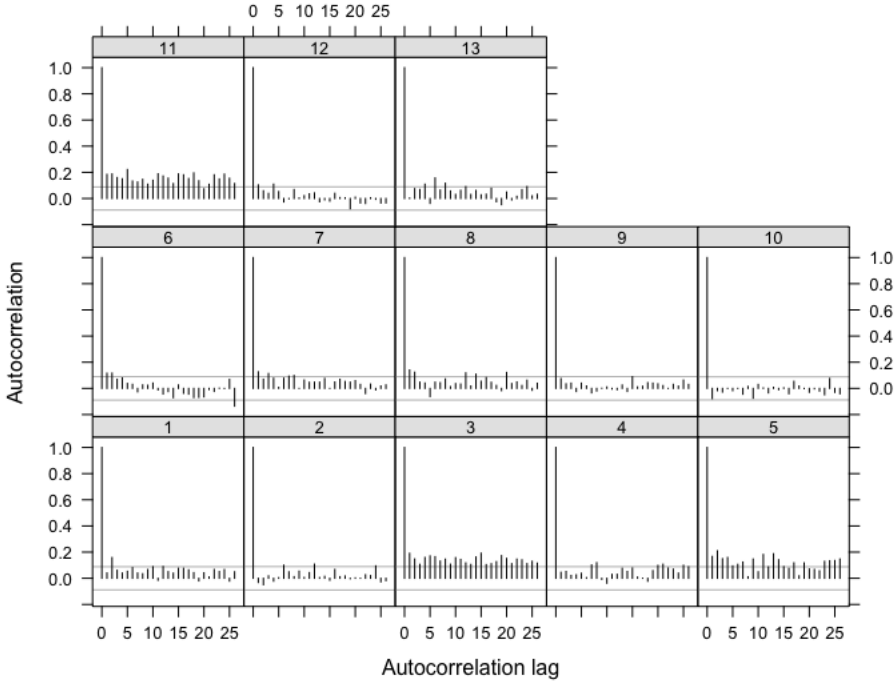


Fig. 6.5 Expt. 5. A trellis plot showing autocorrelation functions for participant-specific trial-by-trial sequential dependencies in response latencies. x-axis: the mean autocorrelation lag between RTs and those of previous trials in the experiment. Panels: individual participants. The grey horizontal line in each panel represents the upper bound of a 95% confidence interval around zero. Any autocorrelation greater than the upper bound represents a significant correlation between RTs and those on the lagged trial. The functions were generated using the acf.fnc function from the "languageR" package in R (Baayen, 2013).

due to the need for prior averaging. For these reasons a mixed effect analysis was run. This involved finding the best fitting random and fixed effects using a backwards stepwise procedure using inverse RTs as the dependent variable. The same model selection methodology was used as in chapter 5 (section 5.4.3).

Starting model: The starting model contained fixed effects of foreperiod [Short, Long], IOI variance [Zero, Low, High], trial index and lagged inverse RT. Trial index describes the order with which all of the experimental conditions were presented to each participant throughout the experiment. Lagged inverse RT is the inverse RT of the preceding trial in each experiment (except for the initial trial of each participant's dataset for which the mean of the RT distribution was used). Both trial index and lagged inverse RT functioned as control covariates. They were included in the model due to the non-independence that was found in the autocorrelation plots. Each helped to satisfy the independence assumption of the linear model and to reduce the residual error of the model fit (Baayen and Milin, 2015). Including trial index also meant that the predictor "session" (used in the initial ANOVA - see section 6.3.1) was no longer required. This is because there was a high correlation between factors trial index and session. Three random effects were also included in the model: 1. Participant ID (ID). 2. Number of noise bursts in each sequence (No. noise bursts). 3. Duration between trial onset and the start of the precursor sequence (Randomised start interval). These allowed each corresponding value to be regarded as drawn from a larger population.

Backwards stepwise selection: To determine the best fitting random effects, three initial models were compared using a likelihood ratio test (LRT) and Akaike information criterion (AIC) measurements: Model 1 contained only a random intercept for ID. Model 2 contained random intercepts for ID and No. noise bursts. Model 3 contained random intercepts for ID, No. noise bursts and randomised start interval. Model 2 had the best fit (χ^2 = 339.89, df = 1, p = < 0.001) and the lowest AIC measurement compared with Models 1 and 3 (Model 1 AIC = 9215.2; Model 2 AIC = 8877.3; Model 3 AIC = 8879.1). Model 2 was retained as it was simpler and better fitting than Model 3. Model 2 was then compared with a fourth model. This contained random intercepts and slopes for No. noise bursts by participant ID as well as a random intercept for No. noise bursts. It did not include randomised start interval. Model 4 provided a better fit (χ^2 = 339.89, df = 2, p = < 0.001) and reduced AIC by 56.2. Model 4 was retained.

	Estimate	Std. Error	t value
(Intercept)	-3.5048	0.1412	-24.8240
Foreperiod-Long	-0.4793	0.1556	-3.0810*
IOIVar-Low	0.0114	0.0514	0.2220
IOIVar-High	0.0465	0.0517	0.9010
Trial	-0.0003	0.0001	-2.0650*
LaggedInverseRT	0.0768	0.0123	6.2470*
Foreperiod-Long : IOIVar-Low	-0.0215	0.0724	-0.2970
Foreperiod-Long : IOIVar-High	0.1844	0.0725	2.5440*
Foreperiod-Long : Trial	0.0000	0.0002	0.2790
IOIVar-Low : Trial	0.0000	0.0002	0.2090
IOIVar-High : Trial	0.0001	0.0002	0.5220
Foreperiod-Long : IOIVar-Low : Trial	0.0003	0.0002	1.0400
Foreperiod-Long : IOIVar-High : Trial	-0.0004	0.0003	-1.541

Table 6.1 Expt. 5. Estimated coefficients, standard errors, and t-values for the best fitting mixed-model (Model 4) fitted to inverse transformed reaction-times elicited for a probe tone following a rhythmic sequence. Absolute t-values greater than 2 for experimental factors are marked by asterisks.

A second stepwise selection process was then performed on all fixed effects and interactions in Model 4. Single term deletions (achieved by using the "dropterm" function in the "MASS" package in R (Venables and Ripley, 2002)) showed that there was a significant three-way interaction between foreperiod, rhythmic variability and trial and a main effect of preceding inverse RT. To explore the influence of this interaction, a fifth model (Model 5) was constructed in which the three-way interaction was removed. A LRT showed that Model 4 was a slightly better fit than Model 5 ($\chi^2 = 6.763$, df = 2, p = 0.034) and had a lower AIC value (Model 4 AIC = 8821.1; Model 5 AIC = 8823.9) regardless of it being more complex in design.

Model estimates: Table 6.1 lists model estimates (intercept and slopes) for the experimental factors and interactions of the best fitting model: Model 4. As discussed by Baayen et al. (2008), Baayen and Milin (2015) and Lo and Andrews (2015), an absolute t-value exceeding 2 is an excellent indicator of significance in the model. This rule was adopted as a measure of significance due to the problems of associating p-values with mixed-effect models (see Baayen et al. (2008) for a discussion). The predictors with t-values above 2 were: 1. Lagged inverse RT, t = 6.2470 2. Foreperiod, t = -3.0810, 3. Foreperiod-Long : IOIVar-High, t = 2.5440, 4. Trial, t = -2.0650 As in the ANOVA (section 6.3.1) there was a significant interaction between foreperiod and IOI variance. When the foreperiod was long, participants took longer to respond on trials with high IOI variance compared with zero IOI variance (periodic) trials. The interaction between low IOI variance and periodic trials was, however, not significant. Interestingly, both control covariates were significant predictors which indicates that RTs were influenced by the previous response in the experiment and decreased as the experiment progressed. This finding is common in a broad range of RT studies and is characteristic of the type of task used (Taylor and Lupker, 2001; Baayen and Milin, 2015). Lastly, although the t-value for the three-way interaction was larger than most, it did not achieve significance and therefore is not a model predictor. This is similar to the strength of the 3-way interaction found in the ANOVA ($F_{(2,24)} = 2.959$; p = 0.071, partial $n^2 = 0.07$) - section 6.3.1.

The difference between the findings of the ANOVA and mixed effects analysis was that sensitivity towards IOI variance was weaker in the mixed effects model compared with the ANOVA and that RTs were dependent on past responses in the mixed effects analysis.

6.4 Experiment 5: Discussion

Experiment 5 aimed at identifying whether response time sensitivity towards IOI variance in a rhythmic sequence is specific to complex decision making or the result of preparatory motor activity. It also investigated what effect the duration of the rhythmic sequence had on RTs and whether this interacted with effects of IOI variance. The main findings were:

1. RTs were sensitive to IOI variance in the rhythmic sequence on trials with long foreperiods, being slowest when the precursor sequence contained high IOI variance and fastest when it was periodic.

2. Trials with long foreperiods were responded to faster than trials with short ones.

These results confirm the experimental hypotheses. They show that the IOI variance findings of experiments 1 to 4 generalise to a very simple task. The implication is that sensitivity towards IOI variance results from general cognitive processes rather than being specific to complex decision making.

By demonstrating that IOI variance affects RTs in a task that did not require complex decision making, the findings of experiment 5 suggest that IOI variance influences behaviour in a way that is relatively general to perception-action tasks.

This suggestion has been previously associated with periodic stimulation (Nobre et al., 2007; Rohenkohl and Nobre, 2011; Morillon et al., 2016), but this is the first explicit demonstration that a critical factor is the degree of aperiodicity rather than a categorical periodic-aperiodic distinction. For example, Rohenkohl and Nobre (2011) showed that periodic compared with aperiodic stimulation not only reduced RTs but led to earlier onsets and increased amplitude in the lateralized readiness potential (LRP). The LRP isolates lateralized activity in the premotor and motor cortex and is known to be negatively correlated with RTs (Osman et al., 1995). The current finding may suggest that oscillatory mechanisms located in the premotor and motor cortex regulated this process. As discussed in section 2.1.3, this is because oscillatory coupling is known to passively strengthen or weaken depending on the degree of phase alignment between interacting systems, resulting in windows of readiness that reflect this coupling. It therefore follows that a rhythmic signal with low IOI variance should produce stronger coupling in an attending oscillatory system than one with relatively higher IOI variance. As a result, RTs towards a target should decrease with decreasing IOI variance. The fact that similar mechanisms are also proposed during early sensory processing (Cravo et al., 2013) implies that temporal expectations may modulate different brain areas through similar mechanisms. This also raises the possibility that the experimental findings may not have been purely the result of motoric processes.

In addition to IOI variance, the duration of the rhythmic sequence affected RTs. Participants responded faster to the onset of the probe tone as the duration of the rhythmic sequence increased. This replicates a well known phenomenon whereby temporal uncertainty negatively affects RTs when the total duration of the sequence is unknown (Woodrow, 1914; Davis, 1965). Interestingly, however, the sequence foreperiod also interacted with IOI variance. This highlights that sensitivity to IOI variance requires prolonged exposure to a rhythmic sequence in order to affect responses. This finding may again be indicative of oscillatory processes, consistent with Large and Jones' (1999) claim that the window of readiness or "attentional pulse" associated with the coupling of biological oscillators reflects an accumulated effect of expectancy violations rather than sequence variability. Therefore, there may not have been enough accumulated expectancy violations on short duration trials for motor responses to have been affected.

Finally, the three-way interaction between periodicity, foreperiod and session, whilst positive, was not significant. This means that IOI variance affected long foreperiod response times in both experimental sessions and that the effect of IOI variance was robust throughout the duration of the experiment. This implies that sensitivity towards IOI variance is an endogenous, stimulus-driven, process that makes minimal demands on higher-order cognitive functions such as memory and conscious awareness.

6.5 Experiment 5: Summary

Chapter 6 reports a new experiment in which participants made simple rather than complex decisions whilst responding to acoustic targets preceded by rhythmic precursors. Whilst the experimental design does not conform to the new experimental approach outlined in chapter 3, it was necessary in order to determine the generalisability of a replicated finding reported in earlier chapters. In experiments 1 to 4, participants' RTs were sensitive to the degree of aperiodicity in the experimental stimulus during complex averaging. By reducing the complexity of the task, yet maintaining common features of the previous experiments, it was possible to determine whether the cognitive processes responsible for this bias seem general to a range of decision types, or just to complex decisions. The results show that, as in experiments 1 to 4, RTs were sensitive to degrees of aperiodicity in the rhythmic stimuli. As IOI variance increased, RTs became longer. This suggests that the previous findings were caused by cognitive processes that are not specific to complex decision making. There was also a significant interaction between IOI variance and sequence duration whereby the effects of IOI variance only occurred on trials that had long precursors. This suggests that sensitivity towards IOI variance requires prolonged exposure to a stimulus in order to affect response. Finally, RTs were much faster on long compared with short duration trials and decreased as the experiment progressed whilst remaining sensitive to IOI variance. The thesis now turns towards an important yet unexplored question relating to complex decisions. Namely, what effect does timing have on complex decisions when the participant, and not the experimental environment, determines how much decision evidence to sample prior to responding.

Chapter 7

Evidence accumulation (Experiment 6)

7.1 Experiment 6: Introduction

A limitation of experiments 1 to 4 and indeed many psychology experiments is that participants had to wait until the end of the stimulus before responding. This makes it difficult to determine at what point in time participants would have responded naturally were they to have the freedom to do so, raising the possibility that stimulus features biased the decision process in a way that was not recorded in the experimental data.

Experiment 6 reported in this chapter investigates how rhythmic temporal expectations bias complex averaging when participants choose how much of a stimulus to listen to. Critically, this procedure distinguishes the time it takes to reach a decision from responses that are restricted to a response period wholly after the stimulus has finished. It also increases the ecological validity of the task by investigating decisions in which people, and not the experimental environment, determine how much of a stimulus to attend to, thus allowing the investigation to be tailored towards behaviours in which deliberation time can be costly. An added benefit of this approach is that Drift Diffusion Models (DDMs), a form of sequential sampling model, can be fitted to the data. DDMs use response distributions to estimate how different components of the decision function under varying experimental conditions. (The key components of the model are the drift rate v, boundary separation a, non-decision time T_{er} and response bias z: see figure 7.1 C for a schematic reminder). Specifically, experiment 6 was designed with three aims in mind: 1. To investigate what effect rhythmic temporal expectations have on complex decision making when participants determine how much of the stimulus to listen to before responding. 2. To determine which components of a DDM are affected by rhythmic temporal expectations during complex averaging. 3. To increase ecological validity and reduce task complexity compared with experiments reported previously in the thesis. The task required deciding whether a long rhythmic sequence of lateralized noise bursts were drawn from a spatial distribution with a mean that was either to the left or right of mid-point. Using the mid-point as the decision criterion meant that, unlike experiments 3 and 4, probe tones were no longer required in the experimental design.

Unlike switching T.V channels, the decision maker does not always control the duration of the stimulus. Decisions must sometimes be based on missing evidence, or responses must be postponed until all of the available information has been acquired (for example, as in judges rating a performance). Regardless of these environmental constraints it is unlikely that decisions are based on all the observed information all of the time. This is because perceptual sampling is thought to be a costly process, helping to penalise decisions that are not timely or that make excessive demands on cognitive resources (Gold and Shadlen, 2007; Drugowitsch et al., 2012; Forstmann et al., 2016). One consequence of this cost is that the accumulation of evidence can terminate well before the end of an observed stimulus stream (Kiani et al., 2008).

As described in section 2.2.2, DDMs propose that perceptual decisions follow the noisy integration of information over time until a threshold level is reached and a response is made. This framework allows for the quantitative modelling of separate decision components under varying experimental conditions. Surprisingly, however, there have been few attempts to apply this approach to the investigation of temporal expectations. Rohenkohl et al. (2012) and Cravo et al. (2013) provide an exception to this. Both studies used a DDM to determine what effect rhythmic temporal expectations had on decision model parameters (see section 2.3). They showed that the periodic presentation of a stimulus increased the rate of evidence accumulation in the model, compared with aperiodic presentation, but that timing did not affect other model parameters. However, Jepma et al. (2012) failed to replicate this latter finding. They used static visual cues to elicit temporal expectations and showed that it was the non-decision time component of the model that was affected by the expectation. These discrepancies suggest that the type of task greatly influences how temporal expectations bias the decision making process. As all three studies required participants to make simple classification decisions, the use of a DDM in this study should provide novel insights into the effects of temporal expectations on complex decision making.

Based on previous findings in this thesis it was hypothesised that periodicity would decrease the time it took for participants to respond during complex averaging (and thus the amount of sampled evidence required to reach a decision), but not necessarily increase the accuracy of the decision. As a result, participants would need to listen to less of the noise burst sequence on periodic trials before making a decision. Forecasting which components of the decision processes would be affected by rhythmic temporal expectation was less clear due to the lack of information in the literature. As sensory entrainment is generally claimed to increase the quality of sensory information, the drift rate of the model should be affected (Rohenkohl et al., 2012; Cravo et al., 2013). However, if true, periodicity should also help to increase the accuracy of the decision (Ratcliff and McKoon, 2008). As this has not generally been the case in previous chapters, this hypothesis appears to lack support. An alternative scenario is that periodicity will bias both the decision boundary and drift rate components of the model. Whilst this effect may be specific to complex averaging, it would ensure that additional accumulated evidence (and hence time) is required on aperiodic versus periodic trials without necessarily improving the accuracy between conditions. Finally, based on a common finding in the decision making literature, task difficulty was expected to affect the drift rate of the model.

7.2 Experiment 6: Method

7.2.1 Participants

A total of 20 participants (9 females) took part. All were students aged between 19 and 35 (Mean = 26.8, SD = 3.7) and were paid £10 an hour. All participants had tone detection thresholds of 15 dB HL or better as measured with a Grason-Stadler GSI 16 audiometer at octave frequencies between 250 Hz and 8000 kHz. At each tested frequency the thresholds for each ear differed by less than 10 dB HL. All but two participants were right handed and none were highly trained practicing musicians. The experiment received ethical approval from the Cambridge Faculty of Music Research Ethics Review Committee.

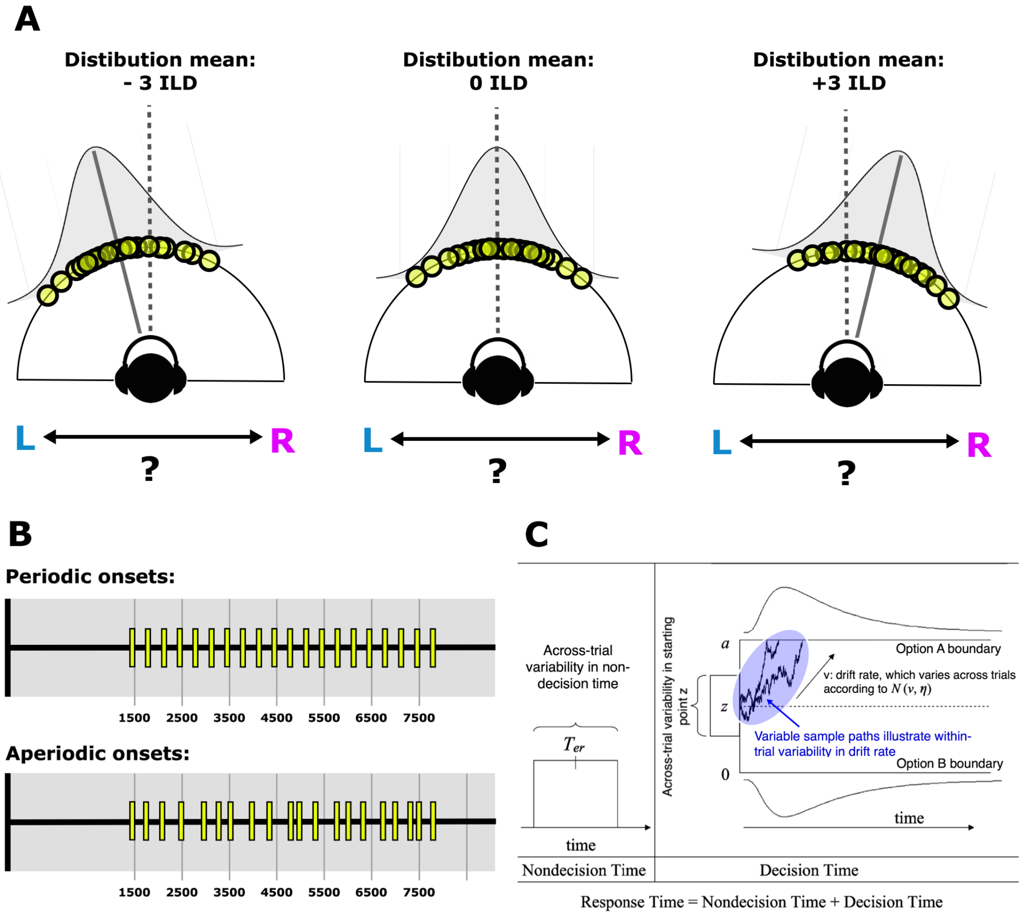


Fig. 7.1 Expt. 6. Schematic illustration of task structure and analytical methods. A: example of three trial conditions. On each trial participants heard a sequence of lateralized noise bursts that had been drawn from a normal spatial distribution. Their task was to decide whether the mean of the distribution was located either to the left or right of mid-point. Yellow dots represent the locations of the lateralized noise bursts. Grey distributions represent the Gaussian spatial distribution of ILDs from which the locations were sampled. Participants did not need to listen to the entire sequence and could respond as soon as they had an answer. B: timing of noise bursts in two trial sequences. Sequences formed either a periodic (top panel) or aperiodic (bottom panel) rhythm on each trial. C: a schematic illustration of the Drift Diffusion Model adapted from Wagenmakers (2009). DDMs are characterised by the drift rate of evidence accumulation v, the decision threshold/boundary a, the non-decision time T_{er} and a response bias z. See section 2.2.2 for a full description of the model.

7.2.2 Auditory stimulus

The auditory stimulus consisted of a train of twenty 40-ms bursts of broadband Gaussian noise, each including 5-ms cosine-squared ramps at both onset and offset. Each noise burst was spatially lateralized using interaural level differences (ILDs) and was bandpass filtered using an 8th order Butterworth filter with cut-off frequencies at 300 Hz and 20 kHz. The bandpass filter was used to target high-frequency sounds in which ILD sensitivity is optimal. Sound lateralization was achieved as in experiments 3 and 4, namely by presenting sound to the right ear at ~70 dB SLP plus half the total ILD, and sound to the left ear at ~70 dB SPL minus half the total ILD.

Preceding every noise burst sequence were two 50-ms 2 kHz sine tones, presented in mid-line with 80 ms IOI, and followed by a silence whose duration ranged randomly between 1400 - 1733 ms. The purpose of these tones was to centre participants' attention on the mid-point and interrupt priming biases associated with recently completed trials. Each tone included 5-ms cosine-squared ramps at both onset and offset.

7.2.3 Noise burst lateralization

On each trial, twenty ILD values were randomly sampled from one of seven ILD distributions. Each distribution was normally distributed, had a mean of either -4k, -2k, -k, 0k, k, 2k, 4k dB ILD and was sampled with equal likelihood throughout the experiment. k was set during the calibration session (see section 7.2.5). It represented the mean ILD of the ILD distribution for which each participant was able to correctly classify the sequence 75% of the time. Each spatial distribution also had a SD of 5 which ensured that the spread of ILD values on each trial was not too narrow. This value was selected during piloting. Once a twenty element ILD array had been selected from one of the seven distributions on each trial, resampling was conducted until the resulting mean and SD of the ILD array differed by no more than ±0.1 from that of the underlying distribution from which it was drawn. This ensured that each noise burst sequence had similar statistical qualities to the distribution from which it was sampled. The elements in the resulting ILD array were then randomly assigned to one of the twenty noise bursts in the sequence. This process was repeated until the ILD that was assigned to the first noise burst in the sequence fell within one SD of the mean of the underlying ILD distribution.

7.2.4 Timing

Temporal presentation of the noise burst sequence could be either periodic or aperiodic. Periodic IOIs were always 333 ms (3 Hz). This rate mimicked that used in experiments 1 and 2 as well as in similar timing studies (Henry and Obleser, 2012; Hickok et al., 2015). It was slow enough to avoid the effects of a decision refractory period (Wyart et al., 2012) but fast enough to ensure that the total duration of the experiment was less than than 1.5 hours. Aperiodic IOIs were randomly selected from a range between 166.5 ms and 499.5 ms. The timing of the individual noise bursts were determined using exactly the same temporal jitter procedure and sampling constraints as used in experiments 3, 4 and 5 - see section 5.3.4 for a full description. Briefly: 1. Aperiodic IOIs were resampled if their value was within ±25 ms of the preceding IOI in the sequence. 2. The total duration of the aperiodic sequence was always within ±10 ms of the total duration of the periodic sequence. Each trial used a new random jitter seed and resampling was conducted until IOIs were found that fitted the selection criteria.

Feedback: Visual feedback was administered on every trial as soon as a response was made. This consisted of a small green or red dot appearing at 0° on an on-screen semi-circle for 200 ms (see section 7.2.5). Green indicated that the response was correct. Red indicated that it was incorrect. Trials in which the spatial distribution had a mean of zero dB ILD (1/7 of all trials) were associated with pseudorandom feedback that was positive on 60% of trials. This mimicked the method used in a similar perceptual averaging task (Wyart et al., 2012).

7.2.5 Design and procedure

The task was administered via headphones within a sound attenuated recording studio located in the Centre for Music and Science at the University of Cambridge. The experimental delivery and data collection were controlled by a Matlab program written by the author of this thesis:

`https://github.com/dcgreatrex-phd/experiment_6`

The experiment consisted of three stages [practice, calibration and experiment proper] in which the task was identical. Participants heard a long sequence of spatially lateralized noise bursts that were drawn from a normal spatial distribution and were required to decide as accurately and as fast as possible whether the mean spatial location of the distribution was either to the left or right of midpoint. Participants could respond at any point before the end of the sequence as soon as they had an answer. On every trial they were shown an on-screen semi-circle presented on a grey background which represented the auditory space. The trial began with a white dot that flashed on-screen at 0° on the outer perimeter of the semi-circle for 200 ms. This dot started at the same time as the first sine tone of the auditory sequence (see section 7.2.2) and functioned as a visuo-spatial reference for the decision. The silent interval and noise burst sequence then followed. Participants could respond at any point from the onset of the first noise burst by pressing one of two vertically aligned response keys (numbers "2" and "8" on the number pad of a computer keyboard). They were to respond using the index finger of their dominant hand and to keep their response finger rested on a separate key ("5") when not responding. The resting position was always equidistant from both left/right response keys and the response key mapping was counterbalanced across participants. The trial timed out 10 seconds from the onset of the first noise burst. Visual feedback appeared 200 ms after a response or the timeout and was always negative following a timeout (see section 7.2.4). The next trial began 600 ms after the visual feedback ended, regardless of whether a response was made.

Prior to starting the experiment proper, participants read task instructions and were shown the left and right images of figure 7.1 A as an illustrative example of the task. They then completed 26 practice trials (13 periodic and 13 aperiodic, randomised), which varied in difficulty. After the practice session participants underwent the same headphone adjustment and staircasing procedures used in experiments 3 and 4 (section 5.3.5). The only difference was that the 3 up 1 down staircase procedure, used to determine a particular participant's 75% accuracy on the task, varied the mean ILD of the underlying spatial distribution and not the difference in ILD between the probe tone and the average noise train location. This estimate of the decision threshold k was used in the experiment proper to tailor the mean location of the ILD distribution in relation to mid-point on each trial. The experiment proper began 5 minutes after the calibration session ended. Each participant performed 8 blocks of 70 trials, 35 periodic and 35 aperiodic, presented in random order, with ILDs within a noise burst sequence randomly sampled as described in section 7.2.3. In total, each of the seven ILD distributions was used 40 times within each of the periodic and aperiodic conditions. Participants were allowed to remove their headphones after the calibration session and again after the fourth block of the experiment proper but were to re-calibrate headphone placement before proceeding.

7.2.6 Hierarchical Drift Diffusion Model (HDDM)

Drift Diffusion models (DDMs) were fitted to each participant's accuracy and RT distribution to determine which aspects of choice were modulated by the timing and location of the auditory sequence. As described in section 2.2.2, DDMs are widely used in the decision making literature and provide a computational account of how separate elements of the decision process vary under different experimental conditions. The hierarchical drift-diffusion model (HDDM) toolbox written by Wiecki et al. (2013) was used to fit the models to the data. The HDDM assumes that parameters for individual participants are random samples from group-level distributions and uses Bayesian statistics to produce estimates of the DDM parameters. These assumptions help to optimise the trade-off between fully independent models that are applied to each participant and fully dependent models that are applied to all participants, such that individual estimates are constrained by group-level distributions (Wiecki et al., 2013). This approach has two benefits over traditional DDM implementations: Firstly, rather than just providing the most likely value for each model parameter, uncertainty in the estimation is quantified by generating posterior distributions. Secondly, it has been shown to produce more accurate parameter estimates compared with other DDM implementations when fitting the model to less than 600 trials (Wiecki et al., 2013; Cavanagh et al., 2011, 2014; Ratcliff and Childers, 2015).

The fitting procedure involved comparing variants of the HDDM to identify which combinations of model parameters best accounted for the response data. This required testing whether choices and response times could be captured by a HDDM in which the drift rate v, boundary separation a and non-decision time T_{er} (or combinations of these parameters) varied as a function of the experimental conditions. In total 19 variants of the HDDM with different parameter constraints were compared in the primary analysis. Following Ratcliff and McKoon (2008), all models included group level trial-by trial variability in the drift rate s_v, response bias s_z and non-decision time s_t, as well as bias z estimates for each participant. Prior to modelling the data, fast response outliers were removed for each subject using the exponentially weighted moving average (EWMA) package in D-MAT, a toolbox for fitting diffusion models (Vandekerckhove and Tuerlinckx, 2008). This followed methods used by Dunovan et al. (2014) and ensured that the estimated non-decision time component T_{er} was never longer than the fastest response. In total, 1.72% of trials were eliminated based on subject-specific EWMA estimates.

Markov Chain Monte Carlo (MCMC) sampling was used to estimate the joint posterior distribution of all models in each HDDM. For each, 10,000 posterior samples were generated with the first 5000 discarded prior to computing the parameter estimation. Discarding the initial samples of a Markov Chain (commonly referred to as "burn in") is a standard technique that assumes the initial samples are too unreliable for use in estimating an unbiased statistic. This followed similar methods used by Zhang and Rowe (2014) and was recommended by Wiecki et al. (2013). Model convergence was assessed by inspecting traces of model parameters, their autocorrelation and by using the Geweke statistic (Gelman and Rubin, 1992). Parameter estimates in all HDDM models converged after 10,000 samples, as assessed by the Gelman-Rubin convergence statistic run on four chains of the three best fitting HDDM models. For all model parameters, this statistic was between 0.990 and 1.004, suggesting 10,000 samples was sufficient for achieving model convergence (Gelman and Rubin, 1992; Wiecki et al., 2013). The Deviance Information Criterion (DIC) was used for model comparison as recommended by Gelman and Rubin (1995), Wiecki et al. (2013) and Frank et al. (2015). DIC applies a degree of penalty for additional free model parameters whereby lower values represent models with the highest likelihood and best fit. Once the best fitting model was identified, the effects of the experimental conditions on estimated parameters were tested.

7.3 Experiment 6: Results

7.3.1 Decision accuracy and latency

Figure 7.2 shows the mean proportion of errors and response times for all participants on the task. See figures B.1 and B.2 in appendix B for participant-specific plots. Neutral trials (in which the underlying ILD distribution had a mean of 0 dB ILD) were omitted from the proportion of errors plot (left panel) due to participants being forced to respond incorrectly on these trials. The figure shows that participants' responses were highly sensitive to the average position of the underlying spatial distribution. As the distribution moved closer to mid-point, participants took longer to respond and made more errors. Response times were also affected by the periodicity of the sequence and were on average longer on aperiodic trials. This means that participants tended to listen to less of the sequence before responding on periodic trials.

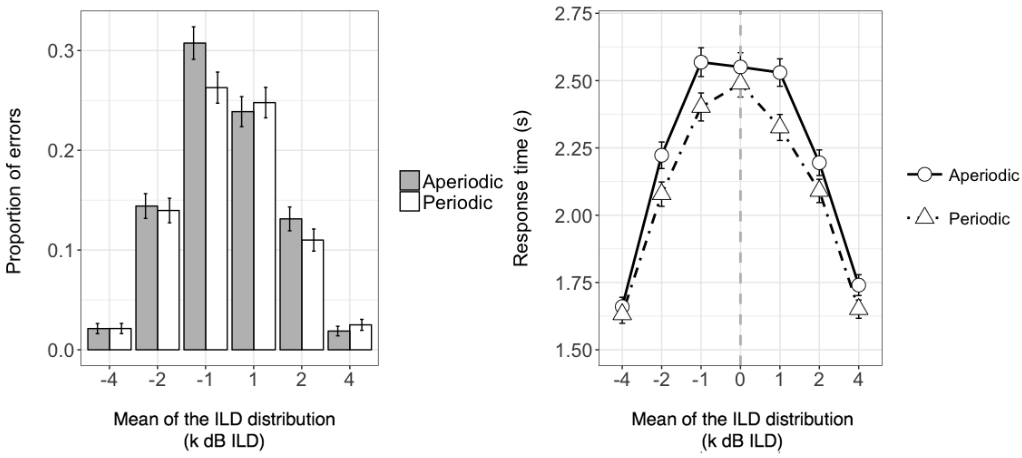


Fig. 7.2 Expt. 6. Mean proportion of errors (left) and response time (right) data at each of the ILD distribution locations relative to mid-point, for periodic and aperiodic stimuli. Error bars = standard error of the mean.

These informal observations were largely confirmed in two separate repeated measures ANOVAs on response times and error rates. For response times, factors were 2 periodicity [Periodic, Aperiodic] x 7 mean ILD distribution levels [-4*k*, -2*k*, -*k*, 0*k*, *k*, 2*k*, 4*k* dB ILD]. Participants responded faster on periodic compared with aperiodic trials (Means: 2.09 versus 2.21 s, SD: 0.76 versus 0.81 for periodic and aperiodic sequences respectively, $F_{(1,19)} = 21.10$, $p < 0.001$, partial $n^2 = 0.53$). The slowing of responses (over a range of about 0.9 s) as the mean ILD approached zero, which is clear in the right hand panel of figure 7.2, was strongly significant (ILD distribution, $F_{(6,114)} = 28.15$, $p < 0.001$, partial $n^2 = 0.83$). No post-hoc tests between individual ILD levels were conducted, as the results would be uninformative for current purposes. The periodicity x ILD distribution interaction was not significant. Similar results were found for log-transformed RTs from correct trials only (Periodicity: $F_{(1,19)} = 19.15$, $p < 0.001$, partial $n^2 = 0.50$; ILD distribution: $F_{(5,95)} = 26.08$, $p < 0.001$, partial $n^2 = 0.84$), and a non-significant interaction.

The ANOVA on the proportion of errors used 2 periodicity x 7 ILD factors. There was a main effect of ILD distribution on proportion of errors, with increasingly more errors at levels closer to zero ILD, but no effect of periodicity, and their interaction was again not significant. [ILD distribution: $F_{(6,114)} = 398.445$, $p < 0.001$, partial $n^2 = 0.99$; Periodicity: means 0.135 and 0.144, SD 0.130 and 0.141 for periodic and aperiodic stimuli respectively, $F_{(1,19)} = 3.475$, $p = 0.078$, partial $n^2 = 0.15$].

Unlike the analysis of the previous experiments, the effects of IOI variance on response was not analysed statistically for Expt. 6. This is because averaged results would have been uninterpretable owing to varying stimulus durations across trials caused by participants responding before the end of some sequences.

7.3.2 Psychometric model fitting

To further examine effects of periodicity on participants' responses, two sigmoidal cumulative distribution functions were fitted to each participants' data by the periodicity condition. This was the same method used in Expt. 3 and 4 - see sections 5.3.5, 5.4.1 and 5.8.1. Briefly, each curve was defined by three parameters: fitted threshold α, fitted slope β, and a fixed lapse rate λ as implemented in the Palamedes toolbox (Prins and Kingdom, 2009). Guess rates were fixed at 0 across all participants and conditions and fitted threshold fixed at 50% accuracy. Each parameter was fitted separately for each subject and periodicity condition. A difference between the periodicity conditions in the slope of the function would indicate enhanced choice accuracy for the steeper slope, whereas a difference in the 50% threshold would highlight a directional bias in favour of one or other periodicity condition. All functions passed goodness of fit (GoF) tests and had a deviance that ranged between 0.56 - 9.29 (Mean = 4.01, SD = 2.31). See Table B.1 in appendix B for all fitted thresholds and slopes values, estimated standard errors, deviance and GoF parameters for each participant and periodicity condition.

Figure 7.3 shows the mean psychometric curves for all participants (top panel) and four participant-specific psychometric curves (bottom panel). See figure B.3 in appendix B for participant-specific plots. These four participants were chosen because they share similar fitted threshold values which facilitates the visual comparison of fitted slopes. Each curve had a positive slope. This indicates that participants could reliably identify the mean location of the underlying noise burst distribution on the majority of trials. This was confirmed using two one-sample t-tests that compared fitted slope values in each periodicity condition against zero (Periodic: mean = 0.67, SD = 0.21, $t_{(19)} = 14.15$, $p < 0.001$; Aperiodic: mean = 0.66, SD = 0.18, $t_{(19)} = 16.07$, $p < 0.001$). The fitted threshold and slope values for all participants were then submitted to two paired t-tests (two-sided) to test for effects of periodicity. The tests revealed that neither thresholds (Periodic: mean = -0.093, SD = 0.483; Aperiodic: mean = -0.146, SD = 0.530; $t_{(19)} = 1.088$, $p = 0.290$), nor slopes (Periodic: mean = 0.674, SD = 0.213; Aperiodic: mean = 0.658, SD = 0.183; $t_{(19)} = 0.828$, $p = 0.418$) were significantly different from one another. As with the initial ANOVA described in section 7.3.1, this finding con-

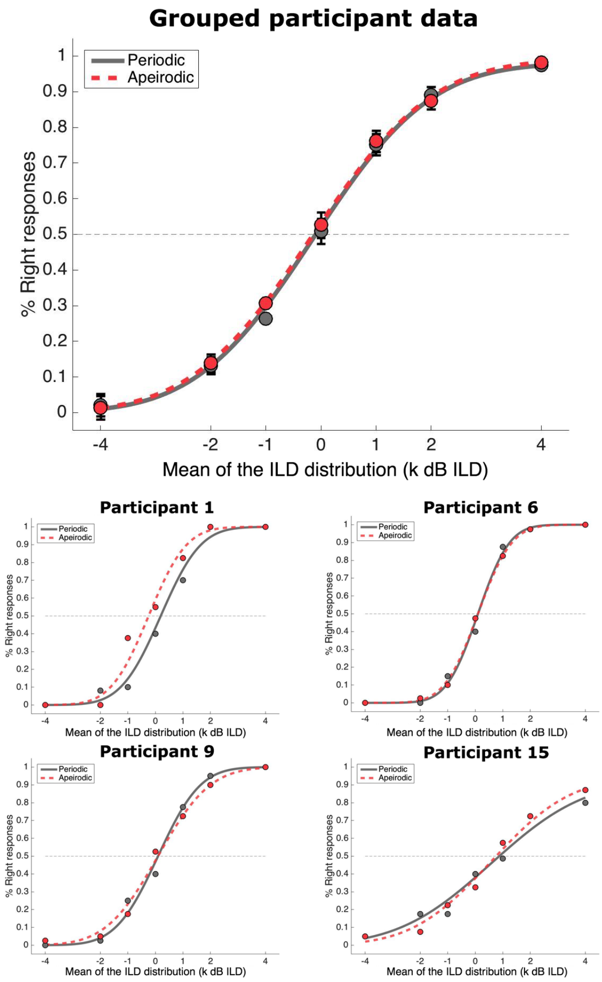


Fig. 7.3 Expt. 6. Fitted psychometric curves showing proportion of Right responses relative to the mean of the underlying ILD distribution. *k* refers to each participant's decision threshold value that was estimated during the calibration session. Group average: top panel. 4 of the 20 participants: bottom panel. Grey solid curves: periodic stimuli. Red dashed curves: aperiodic stimuli.

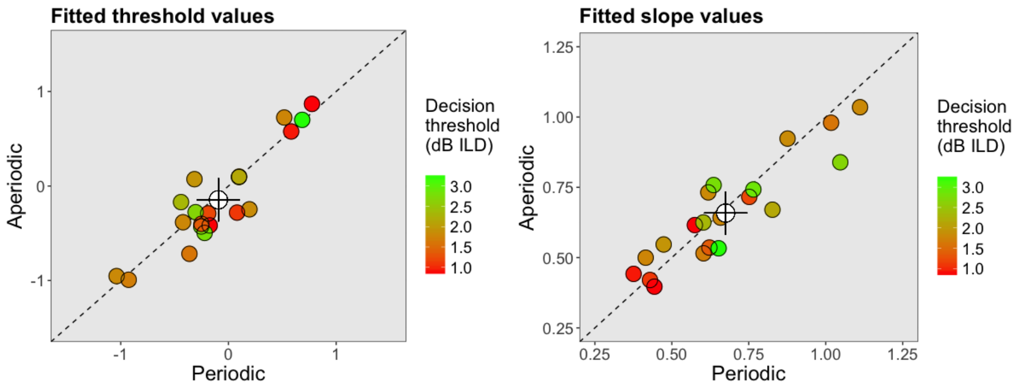


Fig. 7.4 Expt. 6. Scatterplots showing fitted threshold and slope values by decision threshold (colour spectrum). Left panel = Fitted thresholds. Right panel = Fitted slopes. Red circles mark participants with a low decision threshold. Green circles mark participants with a high decision threshold. Crosses mark average values for all participants. Deviation from the dotted line indicates a bias towards one of the two periodicity conditions.

firms that the periodicity of the sequence did not significantly bias the accuracy of the decision.

7.3.3 Decision threshold

Figure 7.4 shows fitted threshold and slope values from the psychometric curve analysis by each participant's decision threshold that was estimated during the calibration session (colour spectrum). In each panel, the centre of the cross marks the average value for all participants. Absolute decision thresholds ranged from 0.825 dB ILD to 3.425 dB ILD (Mean = 1.828, SD = 0.690) and did not correlate with participant age (r = 0.033, p = 0.890), gender (t = 0.659, p = 0.519), nor handedness (t = 0.397, p = 0.696). Unlike in Expt. 4, but similar to Expt. 3, decision thresholds were not predictive of whether a participant was sensitive towards the periodicity of the sequence, according two Spearman's test of correlation and the same method reported in sections 5.4.1 and 5.8.1 (absolute difference between periodicity slopes and decision thresholds: rho = 0.399, p = 0.081; Signed difference between periodicity slopes and decision thresholds: rho = 0.086, p = 0.719).

7.3.4 HDDM selection

The mean ILD of the spatial distribution was used to regroup the data into two new factors: "Perceived average direction" (Left: $k < 0$, Centre: $k = 0$, Right: $k > 0$) and "Difficulty" (High: $[-k, k]$, Medium: $[-2k, 2k]$, Low: $[-4k, 4k]$ db ILD).

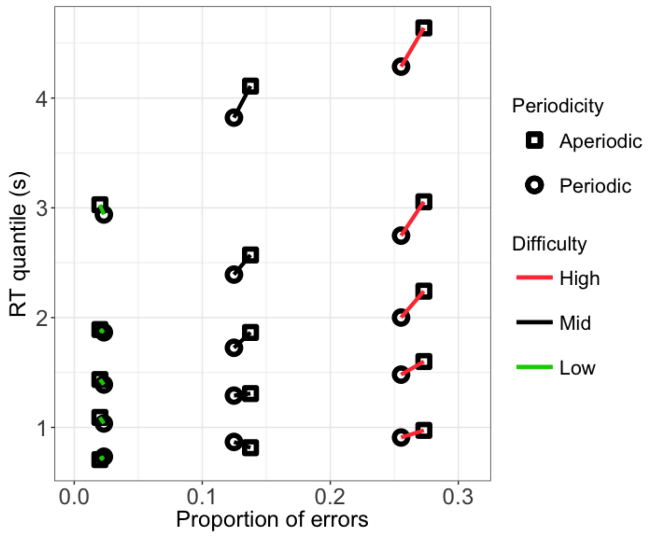


Fig. 7.5 Expt. 6. Quantile probability plots showing RT quantile information and averaged proportion of errors on the same graph split by the periodicity and difficulty conditions. y-axis: vertically stacked quantile information (quantile centres: [0.1, 0.3, 0.5, 0.7, 0.9]) for RT distributions associated with each experimental condition. x-axis: average proportion of errors.

This was done to make it easier to detect more general differences in decision components during the DDM analysis and to increase the number of trials associated with each group. Figure 7.5 shows a quantile probability plot of the difficulty and periodicity groups, excluding neutral trials ($k = 0$). Specifically, it shows vertically stacked quantile information (quantile centres: [0.1, 0.3, 0.5, 0.7, 0.9]) for averaged RT distributions split out by the experimental conditions. Quantile probability plots are commonly used in the decision making literature to visualise differences in response distributions prior to running a DDM analysis (Ratcliff and McKoon, 2008). The plot shows that periodicity affected response distributions on medium and hard difficulty trials and implies that decision components did not vary on trials classed as easy.

To find the DDM that best described the experimental data, 19 variants of the HDDM with different parameter constraints were fitted to the data. This relatively large number of model comparisons was due to the lack of evidence in the literature to strongly support specific hypotheses. Each model varied by whether the three key model parameters (drift rate v, boundary separation a and non-decision time T_{er}) were invariant or varied across the task conditions. The aim of the analysis was to determine which features, if any, of the decision making process were affected by rhythmic temporal expectations. The HDDM analysis

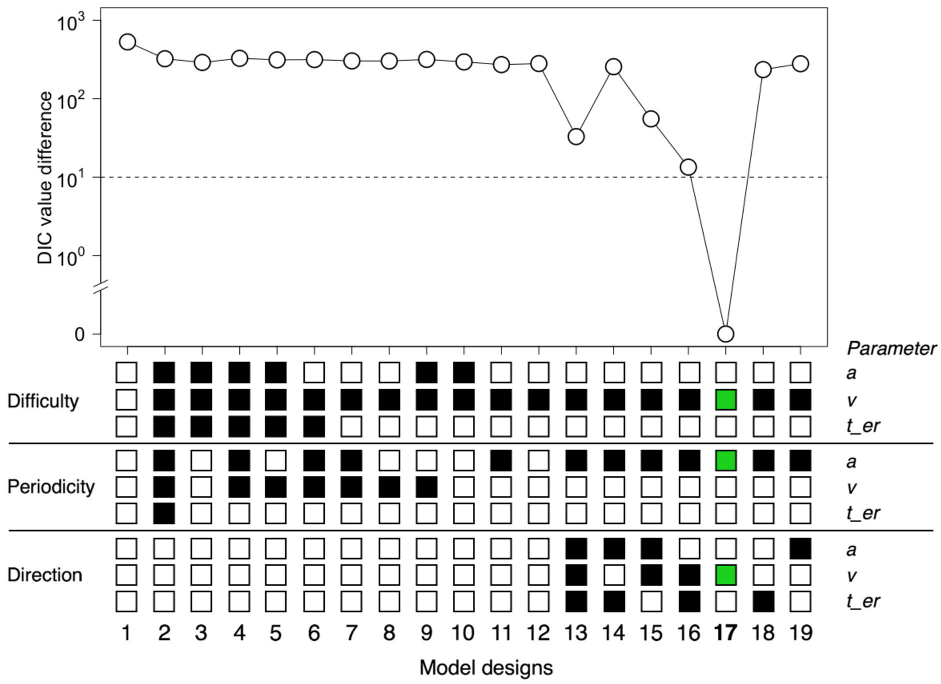


Fig. 7.6 Expt. 6. The deviance information criterion (DIC) value differences between nineteen variants of the drift-diffusion model. The models varied by whether the drift rate v, boundary separation a and non-decision time T_{er} were invariant or varied across the task conditions. The horizontal dashed line represents a significance cut-off point for difference in DIC value. The model structures are shown below the top panel. Filled squares indicate that the particular parameter (labelled at the right) varied by the corresponding experimental condition (labelled at the left). Empty squares indicate that the parameter value was constant in the corresponding experimental condition. Green squares indicate the best fitting model (Model 17, DIC = 22066.22).

was restricted to trials in which the mean ILD of the noise burst sequence was either $\pm 2k$ or $\pm k$. As explained above, this was because the lack of an effect of periodicity on easier trials ($\pm 4k$ - see figure 7.5) meant that changes in model parameters would be restricted to harder task conditions. All models included trial-by-trial variability parameters and participant-specific response bias z as described in section 7.2.6. Model fits were assessed by comparing the DIC values.

Figure 7.6 shows the fitted DIC values for each of the 19 models. The best fitting model (the one with the lowest DIC value) to describe the data across task conditions was Model 17. In this model the boundary separation a was allowed to vary across the periodicity conditions and the drift rate v to vary across both the difficulty and direction conditions. The non-decision time T_{er} was set to a constant in all experimental conditions. The second best fitting model (Model 16) was the same as Model 17 apart from the non-decision time T_{er} component was allowed to

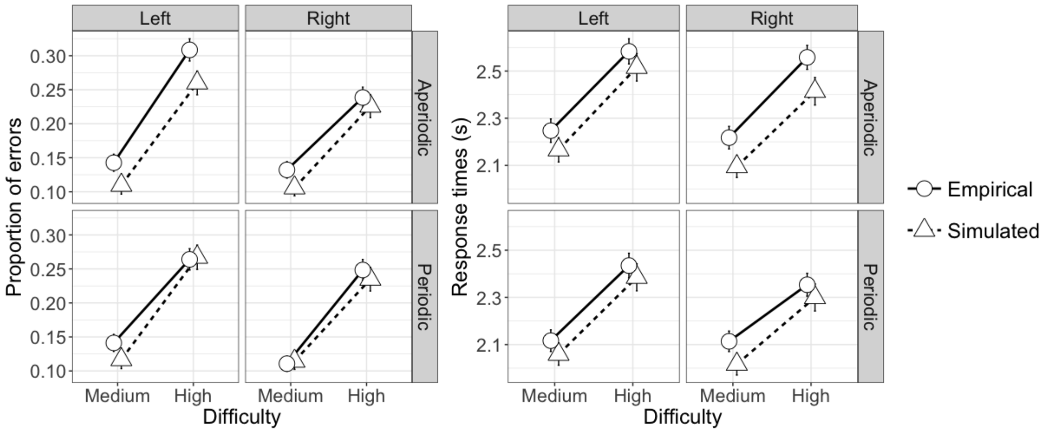


Fig. 7.7 Expt. 6. Proportion of errors (left) and RT plots (right) for medium difficulty [$\pm 2k$] and high difficult [$\pm k$] trials arranged by periodicity (rows: top row = aperiodic sequences) and direction relative to $0k$ (columns). Circles and solid lines are empirical responses (error bars are standard error of the mean). Triangles and dashed lines are posterior predictive simulations from Model 17 (error bars are SDs of posteriors).

vary by the direction condition. This model had a DIC value 13.40 larger than the best fitting model. A difference of 10 or more between DIC scores is interpreted as significantly different (Zhang and Rowe, 2014; Dunovan et al., 2014). The worst fitting model (Model 1) was one in which the three main parameters were invariant in all of the experimental conditions. This had a DIC value 529.24 greater than the best fitting model. The process that was used to select which models to compare within the analysis was as follows: 1. Qualitatively compare empty and complete models (Models 1 and 2) to determine which decision components were most likely to be influenced by the periodicity and difficulty conditions. The average direction of the noise burst sequence was not considered at this stage because it was not originally hypothesised to influence the decision. 2. Based on this assessment, compare a selection of reduced models that are likely to improve the model fit (Models 3 to 12). 3. Once the best model was found (Model 11), systematically add variants of the direction condition to the model design (Models 13 to 19) to determine if direction significantly improved DIC.

To evaluate how well the best fitting model predicted the experimental data, methods recommended by Cavanagh et al. (2014), Zhang and Rowe (2014) and Dunovan et al. (2014) were used. Posterior predictive simulations were made using Model 17. This involved using the posterior estimates of the model parameters for each participant and experimental condition to simulate the same amount of predicted data as observed in the experiment. This procedure stimulated datasets to be compared with the observed data in order to ensure that the estimated model parameters are capable of recreating the main trends in the experimental data. Posterior predictive checks fulfil a similar function to cross-validation and were used because cross-validation is very time-consuming when associated with MCMC sampling. Figure 7.7 shows participants' data plotted against the simulated data produced from Model 17. It shows that there was reasonable agreement between the observed data and the simulated data with the majority of the empirical trends being replicated in the model, although with consistently fewer errors and slower responses estimated. All simulated estimates occurred within 3 standard errors of the experimental means - (see Table B.2 in appendix B for the underlying data).

Finally, to ensure that no better fitting models were excluded from the analysis, a further 12 models were run. These used Model 17 as a base and allowed previously excluded variants of the decision components to vary across the periodicity and difficulty conditions. None of these models provided a statistically better fit to the data than Model 17 - see figure B.4 in appendix B for differences in fitted DIC values.

7.3.5 HDDM analysis

The analysis of Model 17 compared the means of the participant-level posterior distributions using repeated-measures ANOVAs, and where relevant, the group-level posteriors using Bayesian methods as recommended by Wiecki et al. (2013) and described by Kruschke (2014). The first ANOVA on the drift rate v used 2 difficulty [Medium, Hard] x 2 direction [Left, Right] factors. The second ANOVA on the boundary separation a used 2 periodicity [Periodic, Aperiodic] x 2 difficulty [Medium, Hard]. In the following results, p is used to refer to the classical p-value from ANOVA and $P_{P|D}$ to refer to the proportion of the posteriors supporting the testing hypothesis at the group level.

Figure 7.8 shows the group-level posterior mean and SD and posterior probability density distributions for the hierarchical model parameters. The drift rate v was approximately halved for high compared with medium difficulty trials [High: mean = 0.414, SD = 0.229; Medium: mean = 0.809, SD = 0.247; $F_{(1,19)} = 231.488$, $p < 0.001$; $P_{P|D} = 1$]. The boundary separation a was also smaller on periodic (Mean = 2.740, SD = 0.603) than aperiodic (Mean = 2.862, SD = 0.632) trials at both levels of difficulty [$F_{(1,19)} = 9.341$, p = 0.006; $P_{P|D} = 0.767$]. The effect of direction on drift rate and the interaction between direction and difficulty were not significant [Direction: $F_{(1,19)} = 0.409$, p = 0.530; Direction * Difficulty: $F_{(1,19)}$

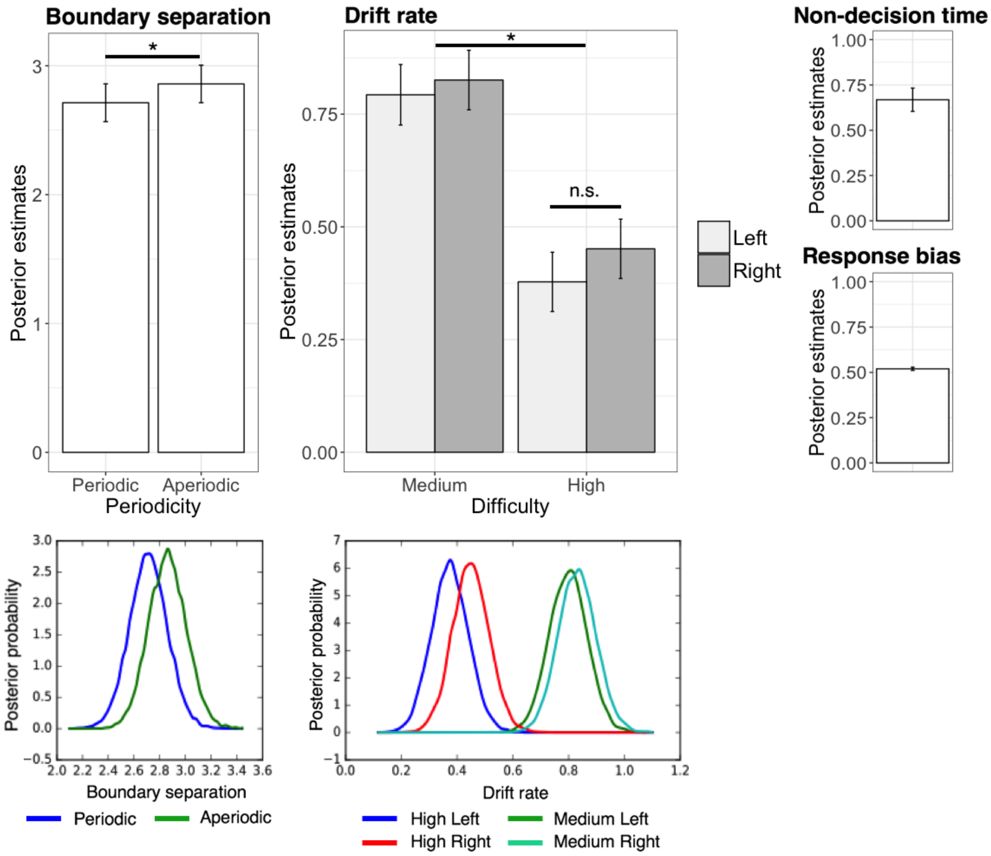


Fig. 7.8 Expt. 6. Group-level posterior estimates of the hierarchal drift-diffusion model parameters from the best fitting Model 17: Top row: boundary separation a (left), drift rate v (middle), non-decision time T_{er} and response bias z (top right). The bars are the sampled mean posterior estimates and the error bars are standard deviations from sampled posterior distributions. Bottom row: the distributions are the Bayesian posterior density functions for each group-level parameter. Asterisks indicate differences significant at p = 0.05 in the ANOVA of individual participant estimations.

= 0.755, p = 0.396; $P_{P|D}$ Left High < $P_{P|D}$ Right High = 0.792; $P_{P|D}$ Left Medium < $P_{P|D}$ Right Medium = 0.642].

7.4 Experiment 6: Discussion

Experiment 6 aimed at determining what effect rhythmic temporal expectations have on complex averaging when participants decide how much of a stimulus to listen to before responding. This is an important change from experiments 1 to 4 as the task provided greater insight into the duration of evidence accumulation on each trial. A hierarchical drift diffusion model (HDDM) analysis was used to determine which components of the decision process are affected by rhythmic temporal expectations. The main findings were: 1. Overall, RTs were sensitive to the periodicity of the sequence with participants listening to more of the sequence on aperiodic versus periodic trials before responding. There was, however, no overall difference in proportion of errors and therefore periodicity did not significantly bias the accuracy of the decision (also found in experiments 2 - 4). 2. Both RTs and proportion of errors were strongly influenced by the location of the underlying spatial distribution (also found in experiments 1 - 4). As the spatial distribution moved closer to the mid-point, participants took longer to respond and made more errors. 3. The periodicity of the noise burst sequence affected the boundary separation a of the HDDM, increasing the boundary separation on medium and hard difficulty aperiodic trials. See figure 7.9 A (left panels) for a schematic illustration of the observed changes in the model. This finding contradicts that of Rohenkohl et al. (2012), Cravo et al. (2013) and Jepma et al. (2012) who showed that temporal expectations either bias the drift rate or non-decision time components of simple decisions. As these authors used experimental tasks that are very different from that employed in experiment 6, this may indicate that the type of task strongly influences how temporal expectations bias decision making processes. 4. Task difficulty affected the drift rate v of the model by increasing the drift on easier trials. This finding is similar to those of a wide range of perceptual decision making studies (e.g. Ratcliff and McKoon, 2008).

From a behavioural perspective, it can be thought that temporal expectations enhanced participants' decisions. Participants made faster decisions based on less evidence on periodic trials without making significantly more errors. This suggests that temporal expectations aided the construction of ensemble representations, perhaps by enhancing the quality of decision information, so that the average location of the noise burst sequence could be more easily inferred from less decision

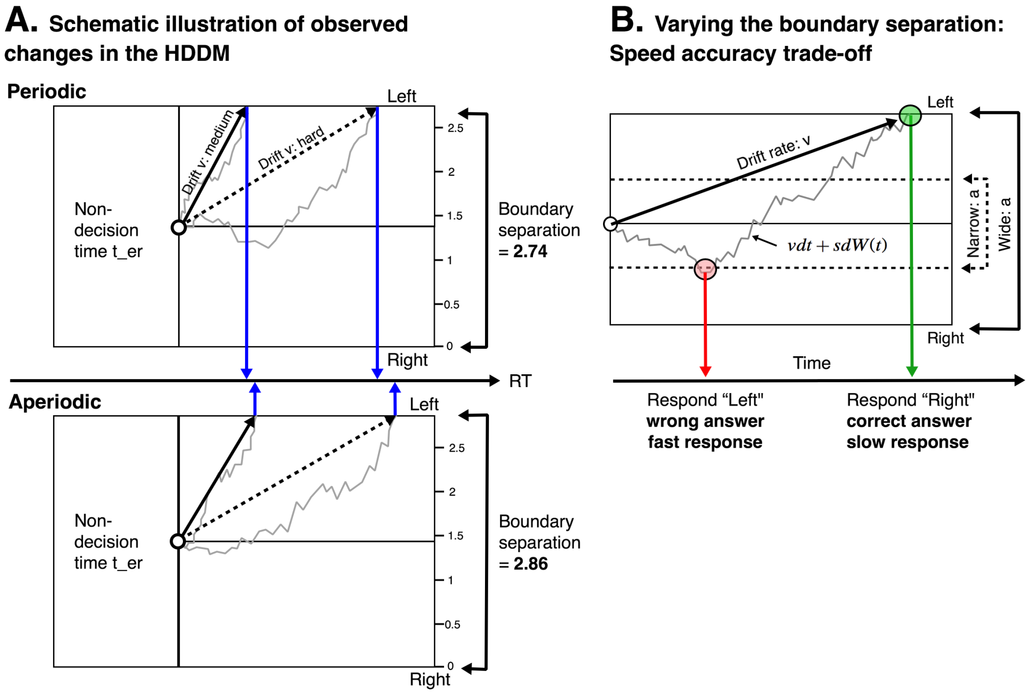


Fig. 7.9 Expt. 6. A: schematic illustration of the changes in the components of the best fitting HDDM model applied to experimental data. Top panel = periodic trials. Bottom panel = aperiodic trials. Grey jagged line = Weiner diffusion process on a single trial in which the participant responds "Left". Drift rate = solid and dashed tilted arrows. Blue vertical lines = moment that the participant initiates a response. Key differences are: 1) Periodic trials have a smaller boundary separation a than aperiodic trials. 2) Hard difficulty trials have a smaller drift rate v than medium difficulty trials but drift rate does not vary by periodicity. B: schematic illustration of the effect that narrow and wide boundary separation a has on the speed and accuracy of a decision. Narrow boundary separation = dashed horizontal lines. The diffusion process is defined by the stochastic differential equation: $dX(t) = vdt + sdW(t)$, where $dX(t)$ = change in accumulated evidence for a small time interval dt, v = drift rate and $sdW(t)$ are zero-mean random increments with infinitesimal variance s^2dt (Wagenmakers, 2009). When boundary separation is narrow, participants respond quickly, but with increased likelihood of errors due to random variance causing the accumulating signal to terminate on the boundary opposite to v. When the boundary separation is wide, there is a reduced likelihood of errors in exchange for longer accumulation time.

evidence. This behaviour differs from the widely reported speed-accuracy tradeoff in which task instructions to respond either "quickly" or "accurately" result in a negative correlation between RT and proportion of errors.

From the perspective of the HDDM, however, the conclusion that temporal expectations enhanced the perceived quality of decision-relevant information is not supported. This is because signal enhancement is characteristic of an increase in the drift rate v and not a change in the boundary separation a. As participants did not make more accurate decisions, but simply responded sooner on periodic trials, the model behaviour is inconsistent with the observed behaviour. That is, a reduction in the boundary separation a on periodic trials implies that participants should have responded faster but with more errors. To understand why, further description of the DDM is required. In a DDM, trial-by-trial evidence accumulation is described as following a Wiener diffusion process in which decision evidence fluctuates randomly over time (Smith and Ratcliff, 2004; Ratcliff and McKoon, 2008). Figure 7.9 B (left panel) shows a schematic illustration of the diffusion process and how random variance can trigger fast error responses. A narrow boundary separation thus increases the likelihood that the diffusion process will terminate quickly (i.e., fast RTs), but simultaneously increases the likelihood that it will terminate on the boundary opposite to the direction of the drift due to random variation in the signal (i.e., a higher likelihood of errors) (Starns and Ratcliff, 2010). Why, therefore, was there a decrease in the boundary separation without an increase in error rates?

Starting with the simplest idea, there is a possibility that the effect of periodicity on boundary separation was not robust and would disappear with a different participant set. For example, although the ANOVA showed that participants' individual boundary separations were quite strongly significantly smaller on periodic compared with aperiodic trials ($F_{(1,19)} = 9.341$, $p = 0.006$), the probability of the posteriors in which the boundary separation for aperiodic trials was greater than the boundary separation for periodic trials was 0.767. This means that according to hierarchical Bayesian hypothesis testing there is a 0.233 probability that the finding would not replicate were a different subset of participants to have been drawn from from the wider population. Whilst this indicates that a decrease in the boundary separation without a change in error rates (as described above) may have occurred due to chance, there are reasons to look for different explanations. Firstly, it is common to assess the explanatory power (i.e. validity) of DDM parameters by comparing DIC and ANOVA significance levels of individual parameter estimations and to then contextualise the strength of these findings using Bayesian hypothesis testing (Jahfari et al., 2013; Zhang and Rowe, 2014; Nie et al., 2016). Secondly, and more importantly, there should theoretically be no instances in which periodicity does not feature in the HDDM if there is behavioural evidence to show that it strongly biases RTs. This is because even were periodicity not to have biased the decision process it should have impacted non-decision time.

A different explanation for why there was a difference in the boundary separation without a change in error rates is that the boundary separation was disproportionately large compared with the variance of the accumulating signal. In this scenario it becomes increasingly unlikely that random variance will cause the signal to revert to the decision boundary that is opposite to its drift (see the red shaded circle in figure 7.9 B for an example of how high variance paired with a small boundary separation can lead to an incorrect response). This is because when the boundaries are widely spaced the drift should eventually determine which boundary the accumulating signal terminates on. Therefore, a small increase in their spacing (caused by periodicity) in the latter stages of the accumulation process should result in increased RTs but not necessarily decreased accuracy. This idea leads to an intriguing hypothesis: if participants were to force themselves to respond on aperiodic trials as quickly as they do on periodic trials they would have responded with the same accuracy. This would be because the additional time associated with widening the boundary separation on aperiodic trials is unlikely to have been large enough to change the course of the accumulating signal and thus affect error rates. According to this argument, aperiodicity therefore made participants less confident and unnecessarily conservative in their deliberation time.

Finally, the use of DDMs in a complex auditory averaging task is novel and therefore potential limitations of the approach need to be discussed. Firstly, RTs in experiment 6 (Mean periodic = 2.09 s) were longer than that of most decision making tasks on which the DDM was originally developed (RTs usually feature within the range of 0.2 - 1.5 s). This could indicate that the decision was too complex to be fully explained using a DDM, but it is unknown at this stage what model would be a better fit. Secondly, the task required averaging a dynamic and not fixed source of information. Decision evidence is therefore likely to have accumulated rhythmically rather than in a linear fashion which is one of the base assumptions of the DDM. Neither caveat is reason to discard the use of DDMs in timing research, but rather, they offer incentives to develop new, more suitable models.

7.5 Experiment 6: Summary

Experiment 6 investigates how rhythmic temporal expectations bias complex averaging when participants choose how much of a stimulus to listen to. The task required deciding whether a long rhythmic sequence of lateralized noise bursts had been drawn from a spatial distribution with a mean that was either to the left or right of mid-point. Unlike in experiments 1 to 5, participants were allowed to respond as soon as they had an answer and were not required to wait until the end of the stimulus. This had the benefit of affording greater insight into the process of evidence accumulation, whilst simultaneously increasing the ecological validity of the task.

The data showed that participants were sensitive to the periodicity of the noise burst sequence, as well as to the location of the underlying spatial distribution. Participants made faster decisions based on less evidence on periodic trials without making significantly more errors. They also responded faster and with fewer errors the further the spatial distribution was located away from mid-point.

A hierarchical implementation of the DDM revealed that the periodicity of the noise burst sequence affected the boundary separation component of the model on medium and high difficulty trials. This implies that participants required less decision evidence before responding on periodic compared with aperiodic trials and that periodicity did not strongly enhance the quality of encoded decision information. As participants did not know before hearing the stimulus whether the sequence would be periodic or aperiodic, this parameter change is likely to have happened in real time during sensory processing on each trial. One hypothesis that accounts for these findings is that aperiodicity made participants less confident and unnecessarily conservative in their deliberation time. In order to validate this idea, additional testing would be required to first replicate the current findings and secondly, expand the context of the task to include visual, tactile or cross-modal averaging decisions.

Having established that aperiodicity increases decision time but not errors, probably by decreasing decision confidence, the thesis now moves away from complex averaging to address the final outstanding area of the new experimental approach described in chapter 3. Namely, do rhythmic temporal expectations bias subjective value decisions?

Chapter 8

Subjective value (Experiment 7)

8.1 Experiment 7: Introduction

Experiment 7 reported in this chapter expands the investigation beyond perceptual decision making to test whether rhythmic temporal expectations bias subjective value representations. This is the last unexplored feature of the new experimental approach described in chapter 3 and is key to expanding the generalisability of the thesis to everyday decision making. A large proportion of everyday decision making is concerned with not only the detection and identification of a stimulus (as in experiments 1 - 6), but predictions about the benefits of outcomes that are yet to be experienced. Understanding, therefore, whether the presentational timing of a stimulus biases the attractiveness of options will afford new insights into the effects of expectation and timing on choice. Most previous research into subjective value is associated with tasks in which participants make preference decisions between items of food. Experiment 7, in contrast, used non-appetitive audio-visual stimuli relating to music in order to increase the generalisability and ecological validity of value-based decision making research. Thus experiment 7 was designed with three aims in mind: 1. To determine whether rhythmic temporal expectations bias subjective value representations. 2. To explore whether results from the literature on subjective value generalise to non-appetitive stimuli, specifically musical preferences. 3. To test whether preference decisions associated with music are any different to those associated with tangible substances such as food.

As discussed in sections 2.1.4 and 3.2.2, there is good reason to believe that temporal expectations may influence subjective value representations. Krajbich and colleagues showed that visual selective attention biases subjective value in that the longer one looks at one of two equally liked food items (with liking determined prior to the experiment), the more likely participants are to select the option from the pair (Armel et al., 2008; Krajbich et al., 2010; Krajbich and Rangel, 2011; Krajbich et al., 2012). This is thought to be caused by the value of the unattended option being discounted during the decision relative to the value of the attended option. As rhythmic temporal expectations are also theorised to affect selective attention (Jones, 1976, 2010; Nobre et al., 2007; Nobre and Rohenkohl, 2014; Henry and Herrmann, 2014), temporal expectations should influence subjective value if both literatures are referring to the same cognitive processes when they interpret results in terms of selective attention. If they do not, it will mean that either the findings of Krajbich and colleagues do not generalise to dynamic multi-modal situations, or that the term selective attention is used by the two literatures to describe distinct cognitive processes.

The experiment incorporated elements from two series of studies, one investigating perceptual modulation associated with neural entrainment (Rohenkohl et al., 2012; Cravo et al., 2013) and the other investigating the role of visual selective attention in complex decision making (Krajbich et al., 2010; Lim et al., 2011). The perceptual modulation studies contributed the use of periodically versus aperiodically presented visual stimuli in a perceptual classification task. The complex decision making studies contributed the use of binary preference comparisons and dynamic stimuli in a subjective value task. The novel aspect of experiment 7 was in combining these methods to test the effects of temporal expectations on abstract valuation processes which concern musical preference and are common in many cultures. It was hypothesised that if rhythmic temporal expectations boosted subjective value representations, participants would be more likely to choose to listen to a piece of music if information relating to the music was temporally aligned versus temporally mis-aligned with a periodic rhythmic context. Temporal expectations were manipulated by having participants watch either a periodic or aperiodic visual flash that immediately preceded the sequential presentation of two musical items, after which participants decided which musical track they would prefer to listen to. Each musical item comprised a very short sound clip of music paired with an image of its corresponding CD cover.

Whilst experiment 7 is the last experiment to be reported in the thesis, it was the first experiment to be run during the PhD. It is placed last because after running the experiment it was decided that further ground work was required into the effects of temporal expectations on perceptual decision making and complex averaging before subjective value could be properly investigated. For this reason the experimental design was not informed by the findings of experiments 1 to 6 and in hindsight could have been improved. These improvements are discussed in section 8.4 alongside ideas for further testing.

8.2 Experiment 7: Method

8.2.1 Participants

A total of 25 participants (14 females) took part. All were aged between 18-32 (mean = 21.9, SD = 3.4) and reported having normal or correct-to-normal hearing and vision. 2 participants were left handed, 7 were actively performing musicians and the average time spend listening to music per day was reported as 2.54 hours (DS = 1.22). Participants were paid £10 for their time. An additional 14 participants (6 females) took part in a pilot study but did not take part in the main experiment. All were aged between 20-60 years (mean = 32.4, SD = 13.48) with an average time spent listening to music per day of 1.75 hours. The experiment received ethical approval from the Cambridge Faculty of Music Research Ethics Review Committee.

8.2.2 Stimulus

Stimuli were 55 audio-visual items comprising a 350-ms sound clip of culturally familiar music accompanied by the image of its corresponding CD cover - see Table C.1 in appendix C for a list of all the musical samples used to create the stimuli. 350 ms was long enough to trigger an emotional response, yet short enough to avoid participants entraining to a musical pulse (Filipic et al., 2010; Gjerdingen and Perrott, 2008; Krumhansl, 2010). To increase the chances of a song being familiar, each auditory clip was taken from either the first chorus of the song, or from central motifs in cases where there was no chorus (Krumhansl, 2010). The 55 stimuli were selected from a larger list of 70 well-known musical recordings: 26 popular songs used by Krumhansl (2010), 24 from the top 100 of Rolling Stone's top 500 songs of all time [1], and 20 from well known jazz and classical recordings. These 70 clips were rated by the 14 pilot participants using methods similar to that of the familiarity task of the main study via a custom built webpage [2]. The

[1]http://www.rollingstone.com/music/lists/the-500-greatest-songs-of-all-time-20110407
[2]http://dcgreatrex-phd-experiment-7.s3-website-us-east-1.amazonaws.com

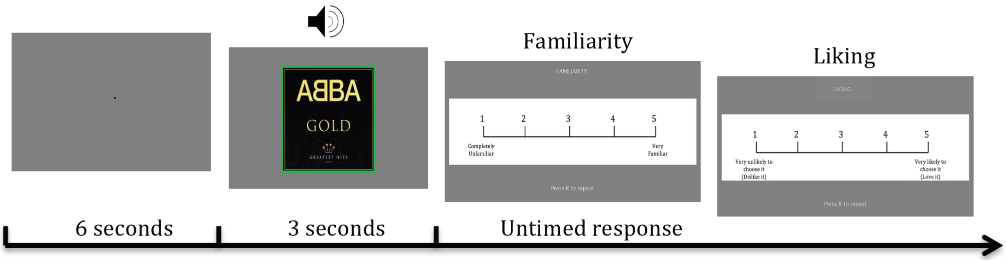


Fig. 8.1 Expt. 7. A schematic illustration of the familiarity and liking task (Task 1).

15 tracks that were least familiar to the pilot participants were omitted from the experimental stimulus set.

8.2.3 Task preparation

One day before the experiment, the 25 experimental participants were asked to prepare by reviewing all of the experimental stimuli under non-timed conditions. This required visiting a webpage that contained images of the 55 CD covers as well 30 s samples of corresponding music (this was similar to the website used by the pilot participants). They were to review each musical item and to listen to the music over headphones. To standardise the time between reviewing this material and doing the experiment proper, participants were again shown the same webpage within the laboratory before doing the experiment. This time, they were told to re-familiarise themselves with the musical items by spending around 5 s listening to parts of each track whilst viewing its CD cover. The presentational order of the on-screen items was randomised across participants and all participants reviewed every stimulus before starting the main experimental tasks.

8.2.4 Task 1: Familiarity and liking

Task 1 established each participant's individual rating of familiarity and liking for each stimulus. Each participant was instructed to imagine that they were flicking through radio stations. Their task was to answer how well they already knew the music and then how likely they would be to tune into it because they liked it.

Figure 8.1 shows a schematic illustration of the task. On each trial participants were shown an image of a CD cover which appeared in the centre of a grey screen (6.7°x 6.7°visual angle) and its corresponding 350-ms sound clip was played simultaneously. This was preceded by a 6 s interval of silence during which a centrally located dot flickered on the screen. The purpose of this interval was to reduce trial-by-trial carry-over effects and to allow participants to focus their attention on the centre of the screen. After the sound clip and visual presentation of the CD cover (3 s), the first of two untimed response windows appeared - the first for familiarity and the second for liking. The familiarity response window contained a 5-point scale ranging from "Completely unfamiliar" to "Very familiar". The response was recorded by pressing one of five corresponding response buttons on a computer keyboard, which produced the next untimed response window: a 5-point "liking" scale relating to the likelihood that the participant would choose to listen to that clip on the radio. The end points were "Very unlikely to choose it - Dislike it" to "Very likely to choose it - Love it". The next trial followed immediately after a liking response. If at any point in the response period participants wanted to see or hear the stimulus again, they could repeat the presentation stage by pressing a marked button on the computer keyboard. Familiarity was always rated before liking, but the button responses were counterbalanced across participants and the order of the audio-visual stimuli randomised on the screen.

After a participant completed Task 1, stimuli that he or she had rated as 4 or 5 on the familiarity scale were grouped into pairs such that there was a difference in liking of between 0 and 2 between the members of each pair. These pairs were to be used in Task 2 for exploring preferences. Thus, each participant's preference choices in Task 2 were based on their actual preferences as established in Task 1, and for any given participant, all stimuli in Task 2 would be more or less equally familiar, but some members of a pair would be more likeable than the other. In principle, if there were not enough familiar items (rated 4 or 5) in Task 1 to generate 8 blocks of unique pairs, the participant would not be able to continue to Task 2. However, all participants passed this condition, thus all completed both Tasks 1 and 2.

8.2.5 Task 2: Binary preference task

As described at the end of section 8.2.4, Task 2 comprised a series of paired stimulus presentations. On each trial, participants were shown the CD covers of a pair of stimuli along with corresponding sound clips, and were to decide which of the two musical items was most preferred. A singe trial comprised a sequence of "pre-target light flashes", followed by presentation of the two musical items and a response period. Details are described separately below.

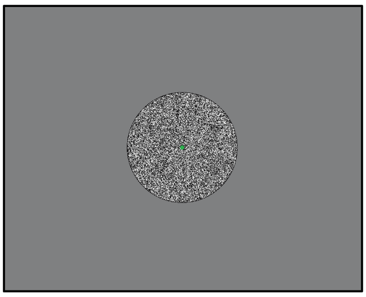


Fig. 8.2 Expt. 7. Example of a Gaussian noise patch used on each trial (Task 2). The black rectangle highlights the size of the noise patch relative to the computer monitor.

8.2.5.1 Pre-target flashes

Each trial began with a sequence of rhythmically flashing noise patches (hereafter termed pre-target flashes). The purpose of these flashes was to vary the rhythmic context in which the task decision was made. The pre-target flashes were identical to those used by Rohenkohl et al. (2012) and Cravo et al. (2013) and were used because Cravo et al. (2013) provide EEG evidence that they induce delta-band entrainment within the visual cortex. As illustrated in figure 8.2, the pre-target flashes consisted of circles (each with a diameter of 4°of visual angle) containing Gaussian noise with a constant root mean square contrast of 10%. In the center of each circle was a clearly visible green fixation point (diameter 20 pixels with rgb value of [0 190 0]). Each flash appeared on-screen for 50-ms.

8.2.5.2 Timing

Temporal presentation of the pre-target flashes could be either be periodic or aperiodic and lasted for a total of 6.8 s (17 flashes on periodic trials). This was long enough to induce sensory entrainment on periodic trials, whilst making it difficult to predict when the paired stimulus would appear in absolute time. Periodic interonset intervals (IOIs) were always 400 ms. Aperiodic IOIs were randomly selected from values 200, 300, 400, 500 and 600-ms in which two identical values never preceded one another. These parameters and constraints were similar to that used by Rohenkohl et al. (2012) and Cravo et al. (2013). Importantly, to control for effects of foreperiod on RTs, the IOI between the final noise patch and paired stimulus presentation was always 400-ms.

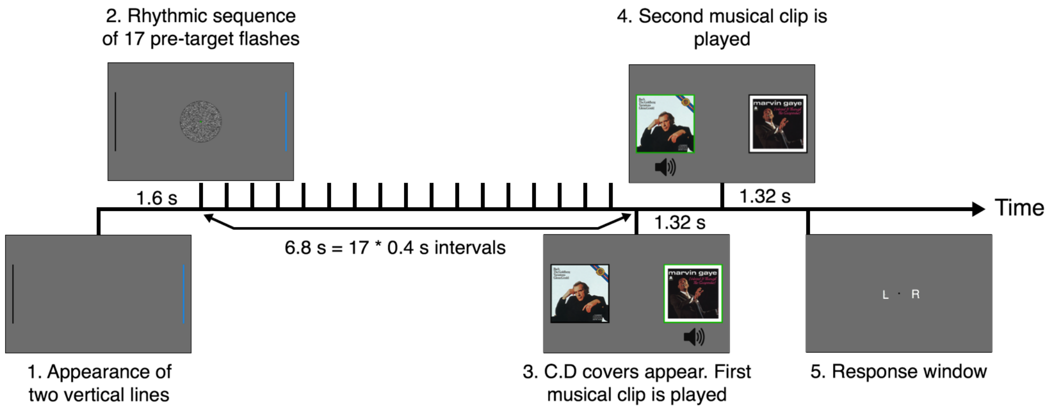


Fig. 8.3 Expt. 7. A schematic illustration of a periodic trial in the binary preference task (Task 2). Aperiodic trials used an identical structure apart from the 6.8 s of pre-target flashes comprising randomly selected time intervals of 0.2, 0.3, 0.4, 0.5, and 0.6 s. The duration between the final pre-target flash and the onset of the CD covers was always 0.4 s on both periodic and aperiodic trials. If participants took longer than 3 s to respond, the trial timed out and the next one began.

8.2.5.3 Structure of a single trial

Figure 8.3 illustrates the structural sequence of two binary preference trials. Each trial began with the appearance of two vertical lines (width 15 pixels), one black and the other blue (rgb value of [0 160 255]), set 16.3°apart. Both were positioned in the same location as the outer edge of each CD cover image. The blue line indicated the side of the screen (left or right) that the participant should attend to when the two CD cover images appeared. After 1.6 seconds, the pre-target flashes started and continued for 6.8 s. Participants were instructed to look steadily at the green fixation point and to pay attention to the flashes until the CD covers appeared. When this happened a green frame surrounded the cover on the side of the blue vertical line and its corresponding sound clip was played. At this moment participants were to look at the cover with the green frame, whilst listening. After 1.32 s the green frame moved to surround the second image and the second sound clip was played, whilst the participant looked at that cover. This duration of 1.32 s was used because it had minimal metrical relationship with the periodic precursors and was long enough for participants to consider each option. After another 1.32 s, a response window appeared. Participants had a maximum of 3 s to indicate which item they preferred to listen to by pressing the relevant one of two response keys ("C" and "M" on a computer keyboard).

8.2.5.4 Distractor trials

In addition to the main task, distractor trials were used to increase the likelihood that participants followed task instructions and were paying attention. The distractor trial began in exactly the same way as a binary preference trial but with one difference. The fixation point within the final pre-target flash was either yellow, red or blue. This change of colour was immediately followed by a response window in which participants were to identify the colour of the last fixation point. An exclusion criterion was associated with these trials in that inattentive participants who scored less than 85% correct would be excluded from the experimental analysis (see results section 8.3.1).

8.2.6 Procedures

All phases of the experiment took place in a sound attenuated recording studio in the Centre for Music and Science, University of Cambridge. Stimuli were presented using Psychophysics-3 Toolbox (Brainard, 1997) on a 22-inch Iiyama Prolite E2202WS screen (vertical refresh rate of 60 Hz) that was positioned 100 cm in front of the participant. Sound was heard through a Beyerdynamic DT 990 Pro headset and responses collected via an Apple keyboard with numeric keypad. The experimental software included custom MATLAB (MathWorks) functions written by the author of this thesis:

`https://github.com/dcgreatrex-phd/experiment_7`

Both tasks began with on-screen instructions followed by 3 practice trials and were administered one after the other within a single experimental session (separated by a 5 minute pause). Task 1 comprised 55 trials which participants did without taking a break. Task 2 comprised 8 blocks of 24 binary preference trials with a unique stimulus pair being used on every trial. In each block, 6 pairs had a liking difference of 1, 6 had a difference of 2 and 12 had a difference of 0 presented in random order. The left right presentational order of the musical items was also controlled within each block whereby the first looked at item appeared on the left hand side of the screen on 12 out of the 24 trials. In addition to these trials, Task 2 included 36 distractor trials that were randomly interleaved throughout the experiment (appearing on average after every 6th binary preference trial). These made it hard for participants to predict what type of task each trial would contain and therefore encouraged them to pay attention to the pre-target flashes on every trial. Unlike in Task 1, participants were given the option to take a short

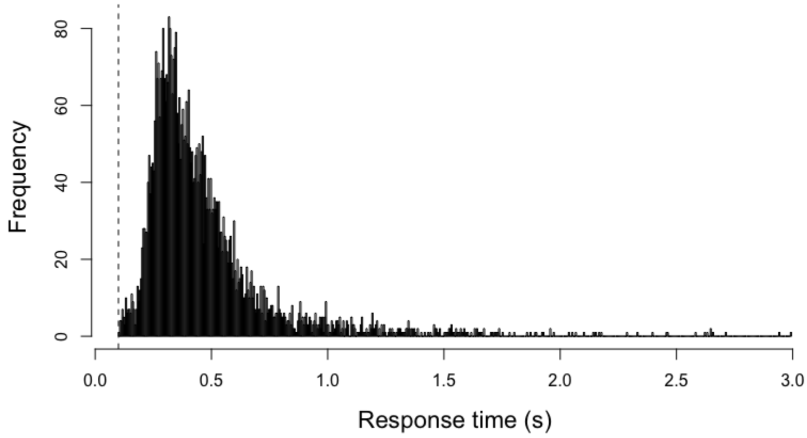


Fig. 8.4 Expt. 7. Histogram of dataset 2 (DS2) showing truncated RT data used in the analysis. The dotted line marks the left truncation point (0.1 s) where very fast responses were removed from the raw data (DS1). Dataset 2 (DS2) is 3.28% smaller than dataset 1 (DS1).

break after every block in Task 2 and were forced to wait 2 seconds after every 6th trial before continuing. These conditions allowed participants sufficient time to rest and refocus their attention during the task.

Finally, an incentive was used to encourage engagement in Task 2. At the start of the task participants were given a £3 iTunes voucher and told that they would need to spend the voucher on music relating to their choice on a randomly selected binary preference trial. This purchase was to be made in front of the experimenter directly after the experiment. The incentive encouraged participants to make honest decisions throughout the experiment as they did not know which trial would be selected at the end.

8.3 Experiment 7: Results

8.3.1 Distractor trials

All participants scored above the 85% accuracy criterion on the distractor trials and were included in the experimental analysis (mean proportion correct = 0.94, SD = 0.04).

8.3.2 Data preparation and analytical procedures

The majority of responses occurred more than 100 ms after the onset of the response period and therefore most participants waited until they had observed the entire stimulus presentation before initiating their response. There was, however, a small number of trials in which responses were much faster than 100 ms. This implies that, on these trials, the decision was made relatively early during the presentation of the second musical item, after which the participant tried to synchronise their response with the onset of the response period. This resulted in a spike in RTs at close to 0 s values. As this response pattern strategy was possible and was allowed in the experiment, such responses were not necessarily outliers, although their clumping close to 0 ms is almost certainty artifactual. Nonetheless, two datasets were used in the analysis. The first dataset (DS1) contained all original trials and the second dataset (DS2, shown in figure 8.4) contained only trials in which RTs were greater than 100 ms. DS2 had 3.28% less trials than DS1 and allowed for standard unimodal statistical tests to be applied to the RT data. Since patterns of statistical significance were identical in both datasets, only the results from DS2 are reported.

The hypotheses were tested using two hierarchical logistic regression analyses that regressed 1) the probability of participants selecting the first item of each pair against the experimental factors and 2) log-transformed RTs against the experimental factors. Both used stepwise selection procedures as well as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Likelihood Ratio Tests (LRTs) for assessing model fit (as recommended by Knoblauch and Maloney (2012)). As there was a high degree of participant variability in the data, choices were also investigated using participant-specific psychometric functions.

8.3.3 Preference reversals

Participants' responses on the binary preference task remained fairly congruent with their choices during the familiarity and liking task. The probability of a preference reversal was 0.153 when the absolute difference score between the paired items was 1, and 0.042 when it was 2. These probabilities are smaller than those recorded by Lebreton et al. (2009) who used a similar binary preference task (hard comparisons: 0.311, easy comparisons: 0.253). As hypothesised, the periodicity of the pre-target flashes did not significantly bias preference reversals when the absolute difference score was greater than zero [$F_{(1,24)} = 0.708$, $p = 0.409$].

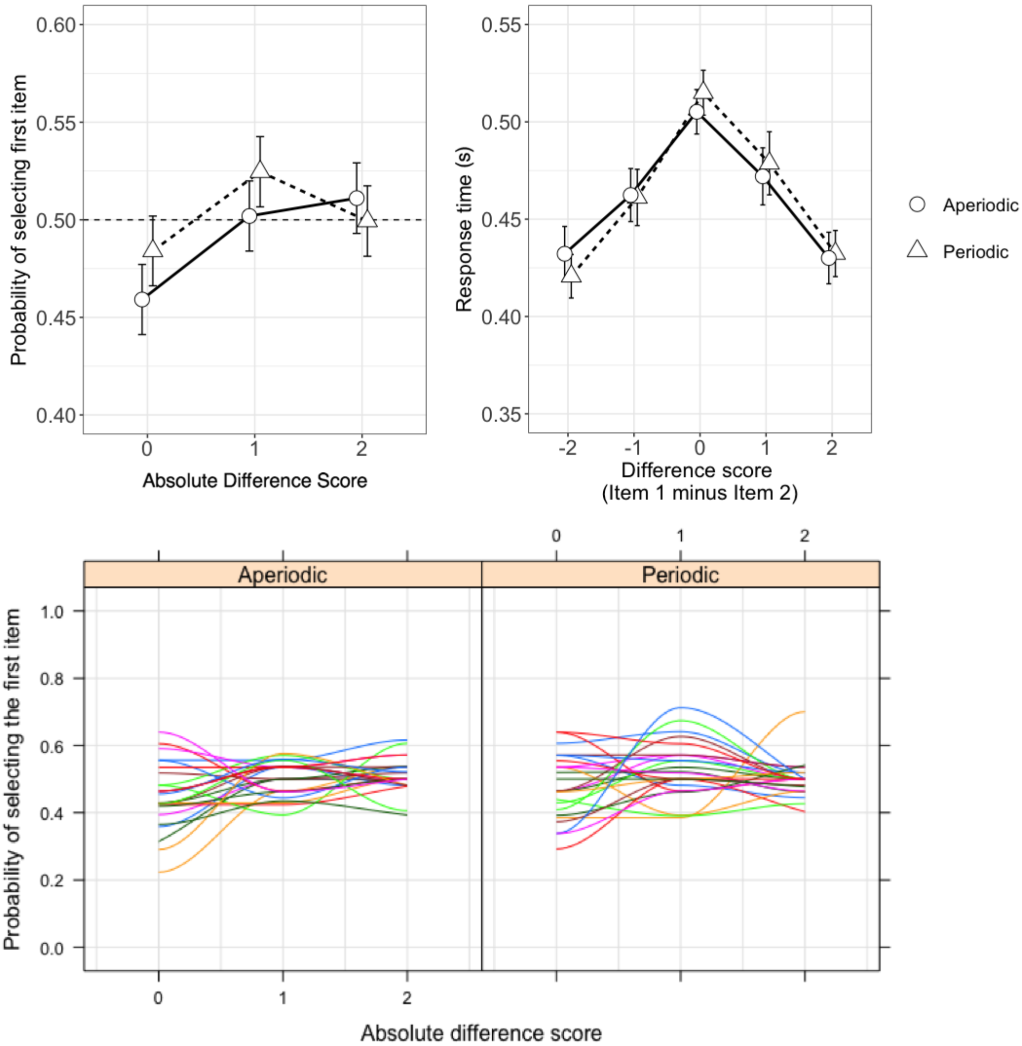


Fig. 8.5 Expt. 7. Choice and RT data. Top-left panel: averaged probability of participants selecting the first musical item by the absolute difference score and periodicity conditions. Top-right panel: averaged RTs by the difference score and periodicity conditions. Bottom panel: participant-specific probabilities of selecting the first musical item by the absolute difference score and periodicity conditions; each curve represents a single participant. Error bars: standard error of the mean.

8.3.4 Probability of selecting the first item

Figure 8.5 shows, for the absolute difference score and corresponding RTs, the probability that a participant selected the first item of the stimulus pair. Figures C.1 and C.2 in appendix C show the same information for individual participants. Contrary to the hypothesis advanced in the introduction, periodicity did not appear to strongly increase the likelihood that participants selected the first item when the difference in liking between the two items was zero. The bottom panel shows that the group averaged results comprised a large degree of between-participant variability.

To make sense of this variability, choice data were submitted to a hierarchical logistic regression analysis. The analysis consisted of two stages: 1. Different generalized linear mixed effect models (GLMM) with the same fixed effects and a variety of random effects were compared to identify the best fitting random factors. 2. Fixed effects were compared using a backwards stepwise procedure to identify the best fitting overall model. The starting model contained fixed effects of Difference score [-2, -1, 0, 1, 2], Periodicity [Periodic, Aperiodic] and Location of the first looked at item [Left, Right]. Four control covariates were also included to account for trial-by-trial carry over effects. The first two covariates, entitled "Recency of item 1" and "Recency of item 2", recorded how many trials earlier in the experiment the corresponding musical item of the pair was last used. The second two covariates, "Repeat of item 1" and "Repeat of item 2", recorded how many times the corresponding musical item had featured in a previous stimulus pair earlier in the experiment. These factors helped to satisfy the independence assumption of the linear model and to reduce the residual error of the model fit (Baayen and Milin, 2015).

Appendix C contains full details of the model selection procedure used to analyse the probability that participants selected the first musical item (see section: "Expt. 7. Hierarchical logistic regression analysis"). In summary, four variants of the starting model with different random effects were compared using a LRT to determine the best fitting random effects. This revealed that the best model to describe variance in the data had random slopes that varied by the Difference score and random intercepts that varied by Participant ID. A backwards stepwise selection procedure then revealed that significant fixed effects were Difference score, Repeat of item 1 and Repeat of item 2 and that Periodicity and Location were not significant. Table 8.1 lists model estimates (intercept and slopes) for the best fitting hierarchical model that excludes non-significant factors. As expected,

	Estimate	Std. Error	z-value
(Intercept)	-0.0447	0.0446	-1.003
Difference score	1.8199	0.1277	14.248
Repeat of item 1	-0.1156	0.0547	-2.116
Repeat of item 2	0.0994	0.0546	1.818

Table 8.1 Expt. 7. Estimated coefficients, standard errors, and z-values for the best fitting generalized linear mixed effect model (Model 3.3) fitted to the probability of participants selecting the first presented item of each musical pair.

there was a strong positive slope associated with the Difference score, which supports the previous finding that participants' preferences remained fairly stable throughout the experiment. Further, the more often an item had been used (in a different stimulus pair) earlier in the experiment, the less likely participants were to select that item in the current trial. As the periodicity of the pre-target flashes did not feature in the best fitting model, any effects of periodicity that may have occurred were not consistent across participants.

8.3.5 Response times

The top-right panel of figure 8.5 shows averaged RT data by the Difference score and Periodicity conditions. RTs increased as the Difference score approached zero but did not appear to be systematically affected by the periodicity of the pre-target flashes. To test for these relationships log-transformed RTs were submitted to a linear mixed-effects regression (LMER) analysis. The starting model contained the same fixed effects and interactions used in the starting model of section 8.3.4, with the addition of Trial ID to capture the autocorrelation of RTs between trials. As with the choice data, the same four combinations of random effects described in appendix C were compared using a LRT. This showed that the best model to explain variance in the data had random intercepts for both Participant ID and Experimental Block. See section "Expt. 7. Linear mixed-effects regression analysis" in appendix C for full details of the model fitting procedure used. A backwards stepwise selection procedure then revealed that only Difference score, Location and Trial ID significantly contributed to the model. Table 8.2 lists model estimates (intercept and slopes) for fixed effects in the best fitting log-transformed RT model. As expected, RTs significantly slowed as the difference score approached zero and were strongly autocorrelated with the RT in the previous trial. There was also a weak spatial bias whereby participants took

	Estimate	Std. Error	t-value
(Intercept)	-0.9769	0.0473	-20.621
Difference score: -1	0.0586	0.0209	2.809
Difference score: 0	0.1468	0.0181	8.134
Difference score: 1	0.0872	0.0208	4.184
Difference score: 2	0.0143	0.0209	0.687
Location: Right	0.0224	0.0120	1.866
Trial ID	-0.1194	0.0166	-7.193

Table 8.2 Expt. 7. Estimated coefficients, standard errors, and t-values for the best fitting linear mixed effects model (Model 2.3) fitted to log-transformed RTs.

slightly longer to respond when the first musical image was presented on the right compared to the left hand side of the screen.

8.3.6 Participant-specific psychometric curves

Although the periodicity of the pre-target flashes did not affect the probability of selecting the first item at the group-level, the rhythm of the pre-target flashes did appear to strongly affect some of the participants' choices (see individual participant plots in appendix C). To explore these cases and to determine whether additional information can be extracted from the data, two sigmoidal cumulative distribution functions were fit to each participants' data by the periodicity condition. This was implemented using the Palamedes toolbox for Matlab (Prins and Kingdom, 2009) and followed procedures described in previous chapters (see sections 5.3.5, 5.4.1, 5.8.1 and 7.3.2). If periodicity affected participants' preferences there should be a difference between the fitted 50% thresholds of each curve (0.5 probability of selecting the first item) with it being smaller on periodic compared with aperiodic trials. The model fitting procedure was applied to all participants' data and Goodness of Fit (GoF) tests used to assess each model fit. Models associated with five out of the twenty five participants [participants 5, 6, 13, 14, 22] failed the GoF test due to high variability in these participants' data. These data were excluded from the analysis.

Figure 8.6 shows data for eight of the remaining twenty participants (see figure C.3 in appendix C for all psychometric curves). The top row shows the four participants who were most affected by periodicity and the bottom row shows those who were least effected by periodicity when the Difference score was zero. Interestingly, although five out of the seven performing musicians (indicated by triangles in figure 8.6) who took part in the experiment are among the eight participants shown in figure 8.6, they are more or less equally distributed between

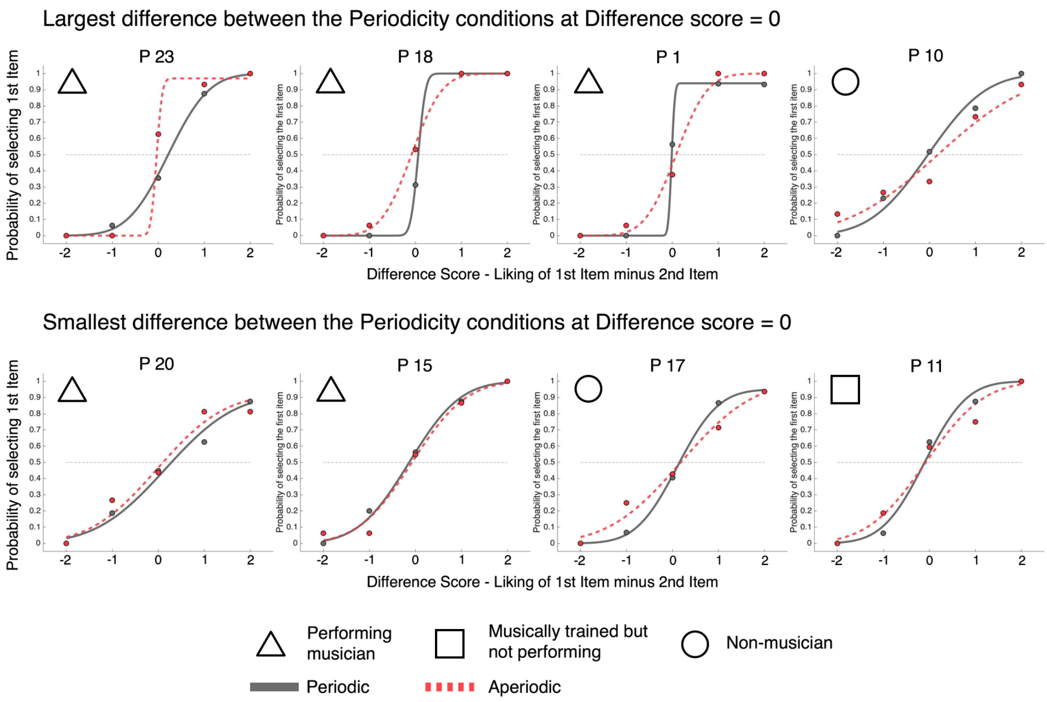


Fig. 8.6 Expt. 7. Psychometric curves for 8 individual participants, showing the probability of selecting the first musical item for a selection of participants by periodicity condition. Top row: participants who were most affected by the periodicity of the pre-target flashes at difference score = 0. Bottom row: participants who were least affected by periodicity at difference score = 0. Shapes: reported musical experience.

the two subgroups: only three were in the group most affected by periodicity, the other two being in the least affected group. This could mean that either the effects of rhythmic temporal expectation on subjective value representations interact with musical experience (such as the type of instrument or music that one plays), or that musical experience simply has no replicable effect on preference choices. As the topic of musical training was not the primary research question for this experiment, further empirical work is needed to clarify this issue.

Two further characteristics of the choices made by participants in the top row of figure 8.6 are also worth discussing. Firstly, participants 23, 18 and 1 all had highly consistent preferences compared with their familiarity and liking task data. On the binary preference task they made on average 3.7% of preference reversals at difference score = 1 and 0.5% of preference reversals at difference score = 2. This is very low compared with the group average of 15.32% and 4.17% respectively. It may indicate that there is a relationship between preference consistency and increased sensitivity towards the periodicity conditions. To test for this, all participants' data were used to correlate the average proportion of preference reversals for each participant at difference scores 1 and 2 with their sensitivity towards the periodicity conditions. Sensitivity to periodicity was quantified as the absolute difference in the probability of choosing the first item between the periodicity conditions at difference score 0. Figure 8.7 plots this data. As indicated by the blue regression line, there was a mild negative correlation between the two measurements but it fell short of statistical significant (rho = -0.359, p = 0.078).

The second characteristic is that participants in the top row, whose choices were most effected by periodicity, differed in how it affected them. Participants 23 and 18 were more likely to select the first item when the pre-target flashes were aperiodic, whereas, participants 1 and 10 were more likely to select the first item when the pre-target flashes were periodic. These differences between participants presumably at least partly explain why periodicity was not a significant predictor in the hierarchical regression analysis, even though it did affect some individuals. Further empirical work is again required to clarify this problem and to determine whether there are specific reasons for these reversals.

8.4 Experiment 7: Discussion

Experiment 7 aimed to determine whether rhythmic temporal expectations bias subjective value representations during a preference task in which participants

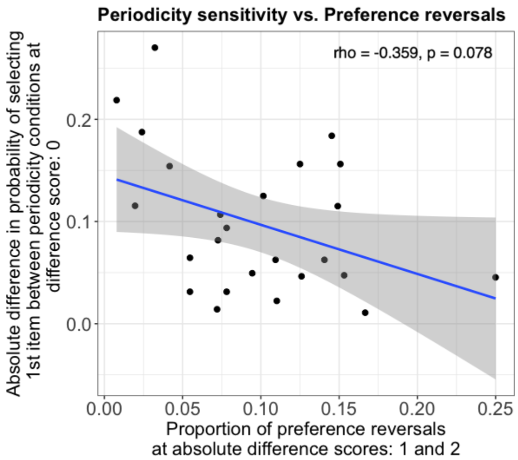


Fig. 8.7 Expt. 7. Proportion of preference reversals at difference scores 1 and 2 (x-axis) against participants' sensitivity towards the periodicity conditions (y-axis). Sensitivity was quantified as the absolute difference in the probability of selecting the first item between the periodicity conditions at difference score 0. The blue line represents the best fitting linear regression line and shaded areas represent predicted confident intervals.

chose between pairs of musical items. The main findings were: 1. Choices were most affected by the difference in liking between the two musical items as well as the frequency of each item. Items with a higher liking rating were more likely to be selected and participants became less likely to select an item the more times it had featured in previous trials. 2. Preferences in the binary preference task were largely consistent with those recorded in the familiarity and liking task and the average number of preference reversals was smaller than that recorded in a similar subjective value experiment that used images of faces, houses and paintings (Lebreton et al., 2009). 3. The periodicity of the pre-target flashes did not bias choices or RTs for the group as a whole but did strongly affect a small subset of participants when the difference score was zero. For these participants the effect of periodicity was not consistent, however, with some participants being more likely to select the first item on periodic trials and others on aperiodic trials. 4. There was a very weak negative correlation between participants' sensitivity towards the periodicity conditions and the number of preference reversals for each participant.

The group results do not support the idea that rhythmic temporal expectations bias subjective value representations. Whilst this finding might generalise to a broader range of tasks and situations, a number of behavioural observations and features of the experimental design encourage consideration of a more nuanced interpretation. For example, some of the participants were strongly influenced by the periodicity of the pre-target flashes when the difference score was zero. This could mean that there were participant-specific factors that were not controlled for that interacted with effects of temporal expectation on subjective value representations. Henry and Obleser (2012) demonstrate that the effects of rhythmic temporal expectation on choice can be participant-specific. They showed that the optimal phase within a periodic stimulus for the correct detection of near threshold targets differs across participants. Specifically, the entrained phase of each participant's neural delta oscillations differed (by varying degrees) in phase-lag relative to the stimulus, and that it was the peak of the delta oscillation and not the stimulus that determined optimal target detection. As the current study tested a relatively small number of participants, further work is required to find out if the reported results replicate under different tasks that test the same research question and whether performance differs across different participant groups. This could involve comparing participant groups associated with specific rhythmic expertise, such as professional drummers, dancers or poets against a control group who have little or no such specific rhythmic training. Alternatively, electroencephalography data could be recorded, as in Henry and Obleser (2012), to determine whether the phase of neural delta oscillations biases the probability of participants selecting the first item.

In addition to the differences between individual participants there were a number of features of the experimental design that may have reduced the interpretability of the data. Firstly, although each trial contained a unique stimulus pair, there was no upper limit to the number of times a single musical item could feature in the experiment. As already noted, this biased responses, with participants being less likely to select an item the more times it was used throughout the experiment. One result of this issue is that item replication will have likely caused forced preference reversals. A simple solution is to present each musical item a fixed number of times in each of the Difference score conditions.

Secondly, a number of participants reported after the experiment that the duration of the pre-target flashes (6.8 seconds) was too long and as a result they lost interest or became frustrated with the task. The long duration was used to ensure that the pre-target flashes induced sensory entrainment in observers. As demonstrated in experiments 1 to 6, however, the effects of timing on complex choice occur after exposure to short rhythmic sequences and therefore this duration could have been much shorter, but was not known when the experiment was conducted. Further piloting is required to determine more appropriate presentation durations so that they remain comparable with natural and unrestricted decision latencies between musical options.

Thirdly, the design of experiment 7 does not allow us to explicitly measure whether the musical judgements were compatible to those about tangible substances such as food. The experimental design could therefore be replicated, except with the musical items replaced by images of food or a range of stimulus types.

Fourthly, the fact that the pre-target flashes were restricted to the visual domain may have reduced the effects of periodicity on choice. This is because it is largely untested whether visual sensory entrainment triggers auditory entrainment and whether a cross-modal connection is made during complex decision making. This means that for participants who based their decision primarily on information contained in the musical clip, rather than in the CD cover image, the effects of visual sensory entrainment would have had little effect on their decision. A simple way to test this would be to include a condition in which the pre-target flashes are synchronised with clicks or noise bursts to ensure that both the visual and auditory cortices are entrained to the periodic pulse prior to the presentation of the musical items.

Finally, the methods used in experiment 7 are at odds with one of the principles outlined in the new experimental approach of chapter 3 (section 3.3.2) as well as a general aim of the thesis. As the pre-target flashes did not contain goal-related information they violated the suggestion of using time as an inherent dimension of targets (section 3.2.3). Experiment 7 also lacked ecological validity as choices made between musical items do not usually occur within rhythmic contexts. These limitations exist, of course, because the experiment was designed and run before the main argument of the thesis was fully formed. There are, however, a number of real world contexts in which value-based decision information is embedded within dynamic rhythmic streams. For example, it is common to form social judgements about the trustworthiness or approachability of people whilst navigating through a busy town. An experiment could be run to test whether the temporal variability of peoples gait, or of one's own movement, biases these judgements. Knight et al. (2016) asked a similar question by having participants indirectly judge the trustworthiness of a person who was filmed walking in-time versus out-of-time with an auditory pulse. They showed that people judged the person to be implicitly engaging in a trustworthy activity when the auditory pulse aligned with their footsteps, but explicitly engaging in an untrustworthy activity when it did not. Alternatively, one could investigate market situations, such as an auction, when participants are required to make value decisions based on rapidly changing rhythmic information. Armel et al. (2008) investigates a similar scenario in the context of dynamic object selection but did not focus explicitly on rhythmic temporal expectations.

8.5 Experiment 7: Summary

To conclude, this chapter documents the first attempt at investigating whether rhythmic temporal expectations bias subjective value representations. To avoid appetitive stimuli and increase the generalisability of value-based decision making research, participants made preference decisions between periodically or aperiodically presented musical items. The results showed that choices were highly sensitive towards the prior liking of each musical item and that participants were less likely to select an option the more times it was used throughout the experiment. The number of preference reversals was also on average smaller than in a similar visual study that used pictures of faces, houses and paintings as stimuli (Lebreton et al., 2009). This finding validates the use of musical stimuli in the experiment and suggests that musical preferences are the same or more reliable than commonly investigated preferences in the decision making literature. Importantly, periodicity did not affect choices for the group as a whole, but did affect decisions made by a subset of the participants, though only when they had no initial difference in preference between a particular pair of items. This implies that participant abilities, which are presumably at least partly affected by their experience, interact with the effects of periodicity on choice. Whilst the current experimental design has a number of limitations, it forms a baseline of results and acts as a starting point from which further research can be made.

Chapter 9

Towards a predictive framework

9.1 Key findings using the new experimental framework

This thesis proposes a new approach for studying temporal expectations and decision making. It argues that to better understand the effects of temporal expectation on perception and action, theories of timing must be tested under experimental conditions that more closely align with everyday goal-directed behaviour. This requires moving beyond simple decisions to consider the effects of temporal expectations on complex decision making, whilst ensuring that stimulus timing is contextualised and modelled as part of, and not separate from, goal-relevant information. The experiments reported in the thesis investigated this interdependence between temporal expectations and complex decision making by conforming to a new experimental framework outlined in chapter 3. The framework addresses five areas of existing temporal expectation research that are in need of further development in order to expand the generalisability and relevance of experimental data to everyday situations and offers guidelines on how to achieve these improvements in experimental design. Specifically, the framework recommends: 1. increasing the complexity of experimental tasks by using multiple targets, 2. requiring participants to make perceptual averaging and subjective valuation decisions, 3. incorporating sequence timing as an inherent dimension of targets, 4. testing degrees of aperiodicity and 5. exploring the effects that prior knowledge about the temporal structure of a stimulus has on choice.

Table 9.1 summarises the main independent and dependent variables and results of each experiment in the thesis.

Table 9.1 Summary of the independent variables (IVs), dependent variables (DVs) and results of the seven experiments reported in the thesis. "C" represents control covariates.

Expt. No.	***IVs***	***DVs***	***Results***
1. Chp. 4	Periodicity Spatial category IOI variance Azimuth variance	Response times Proportion errors Decision weights	- Main effect of periodicity and spatial category on response times and proportion of errors. Faster responses and fewer errors on periodic versus aperiodic trials. Faster responses and more errors on cardinal versus diagonal trials. - Main effect of IOI variance on response times but not proportion of errors (slower responses with increasing IOI variance). - Interaction between azimuth variance and spatial category (higher errors on cardinal trials with low azimuth variance). - Larger decision weights on periodic versus aperiodic trials but no recency or inlying/outlying decision evidence bias.
2. Chp. 4	Same as Expt. 1	Same as Expt. 1	- Main effect of periodicity and spatial category on response times. Faster responses on periodic versus aperiodic trials and on cardinal versus diagonal trials. - No effect of periodicity nor spatial category on proportion of errors. This is different to Expt. 1. - Main effect of IOI variance on response times but not proportion of errors (slower responses with increasing IOI variance). - Main effect of azimuth variance on decision accuracy (more errors with increasing azimuth variance). - Smaller decision weights than in Expt.1 and fewer differences between the periodicity conditions. No recency or inlying/outlying decision evidence bias but an unexpected decision weight interaction between the periodicity and spatial category conditions.

Table 9.1 Continued.

Expt. No.	***IVs***	***DVs***	***Results***
3. Chp. 5	Periodicity Probe tone ILD IOI variance Average noise ILD (C)	Response times % Left responses	- Main effect of periodicity and probe tone ILD on response times. - No effect of periodicity but a main effect of probe tone ILD on decision accuracy. - Main effect of IOI variance whereby responses slowed with increased IOI variance. - Main effect of Average noise ILD on response accuracy and response times highlighting a spatial bias analogous to the "spatial aftereffect" (Phillips and Hall, 2005).
4. Chp. 5	Same as Expt. 3	Same as Expt. 3	- Main effect of periodicity and probe tone ILD on response times (same as Expt. 3). - No main effect of periodicity but a main effect of probe tone ILD on decision accuracy. High ability participants were significantly more accurate on periodic versus aperiodic trials compared with low ability participants. - Main effect of IOI variance whereby responses slowed with increased IOI variance (same as Expt. 3). - Main effect of Average noise ILD on response accuracy highlighting a spatial bias analogous to the "spatial aftereffect" (Phillips and Hall, 2005) (similar to Expt. 3).

Table 9.1 Continued.

Expt. No.	***IVs***	***DVs***	***Results***
5. Chp. 6	Foreperiod IOI variance Session Lagged RT (C)	Response times	- Main effect of foreperiod. Response times were longer on short compared with long duration trials. - Significant interaction between IOI variance and foreperiod. Response times were affected by IOI variance on long but not short foreperiod trials (IOI variance slowed responses). - Main effect of experimental session. Responses were faster in the second half of the experiment. - Main effect of lagged inverse response time. RTs positively correlated with the RT of the previous trial.
6. Chp. 7	Periodicity Average noise ILD	Response times Proportion errors % Right responses HDDM (v, a, T_{er})	- Main effect of periodicity and average noise ILD on response times. RTs were faster on periodic versus aperiodic trials and slowed as the average noise ILD approached zero. - No effect of periodicity on proportion of errors but a main effect of average noise ILD. Participants made more errors under harder task conditions. - Main effect of task difficulty (low versus high average noise ILD) on the drift rate v of the HDDM. v was higher under easier task conditions. - Main effect of periodicity on the decision boundary a of the HDDM. a was smaller on periodic versus aperiodic trials. - No effect of direction (positive versus negative average noise ILD - control covariate) on any of the HDDM parameters.

Table 9.1 Continued.

Expt. No.	***IVs***	***DVs***	***Results***
7. Chp. 8	Periodicity Difference score Location (C) Item rep (C)	Response times % Select item 1 Preference reverse	- No effect of periodicity nor spatial location on response times but a main effect of difference score. Response times increased as the difference score between items decreased. - No effect of periodicity nor spatial location on the probability of selecting the 1st musical item but a main effect of difference score. - Preference reversals decreased with increased difference score and were generally smaller that those reported in similar studies (Lebreton et al., 2009). - Participants were less likely to select the 1st musical item the more times it appeared in the experiment and response times negatively correlated with that of the previous trial.

Experiments 1 and 2 used a complex auditory averaging task to investigate complex decision making and whether prior knowledge about the periodicity of the stimulus, as well as degrees of aperiodicity, bias choice. As this was the first attempt at implementing the new experimental framework, the experiments addressed as many of the above five points as possible. The results showed that, under conditions of temporal uncertainty (experiment 1), the average location of periodic lateralized sequences was categorised faster and more accurately than that of aperiodic sequences. Under situations of temporal certainty (experiment 2) the effect of periodicity on accuracy disappeared and decision weights were more evenly distributed. In both experiments, decision latencies became proportionally longer as the degree of IOI variance in the stimulus increased. Together, these findings imply that the rhythmic presentation of information can bias the accuracy of complex averaging decisions when the task contains temporal uncertainty and that prior knowledge about the timing of the stimulus counteracts this bias, presumably by increasing the likelihood of information being sampled uniformly during the decision. It also shows that the degree of aperiodicity in the stimulus systematically regulates the time it takes to reach a decision in a way that is not reliant on prior knowledge, decision accuracy, nor, presumably, conscious control.

Experiments 3 and 4 developed the approach by having participants make relative rather than absolute auditory spatial averaging decisions and by varying the rate of the stimulus (experiment 3 - fast rate, experiment 4 - slow rate). Unnecessary complexities were also removed from the experimental design or controlled for. These included reducing the number of sounds in the rhythmic sequence, and controlling the order in which spatial targets were presented. The purpose of both experiments was to test whether key findings of experiments 1 and 2 replicate under a different experimental paradigm that tests the same fundamental research question and whether the rate of rhythmic stimuli interacts with the effects of rhythmic temporal expectations during complex averaging. Unlike experiment 1, periodicity did not result in participants responding more accurately in experiment 3 (fast rate) and experiment 4 showed that enhanced decision accuracy on periodic trials only occurred for a subset of participants who were able to do the task under the harder conditions of the slow rate. This suggests that the slower presentational rate increased the likelihood that periodicity would affect choice accuracy for high ability participants, and that effects of temporal expectation on the accuracy of complex decisions is dependent on several factors that include, but are not limited to, decision complexity, the rate of the stimulus, knowledge about the timing of the stimulus and participant ability. The effect of IOI variance on decision latencies was replicated in both experiments 3 and 4 and therefore sensitivity towards IOI variance in a stimulus stream has a more robust effect on decision latencies than the other factors listed here.

Experiment 5 diverged from complex decision making and the rules outlined in chapter 3 to investigate the cause and characteristics of the IOI variance finding in experiments 1 to 4. Experiment 5 comprised a simple detection task making minimal demands on memory and cognitive processing. If the same effects of IOI variance on RTs was observed in this task it would provide evidence that the findings of experiments 1 to 4 were caused by cognitive processes that are relatively general to perception-action tasks. The results showed that RTs were sensitive to the IOI variance in the precursor sequence but that the effects of IOI variance were limited to trials that had long precursors. This suggests that sensitivity towards IOI variance is not specific to complex decision making although prolonged exposure to a stimulus is needed for responses to be affected.

Experiment 6 returned to complex decision making and the experimental framework to investigate how rhythmic temporal expectations bias complex averaging when participants choose how much of a stimulus to listen to before making a decision. Unlike experiments 1 to 5, participants were allowed to respond as soon as they had an answer rather than being required to wait until the end of the stimulus before responding. This had the benefit of affording greater insight into the process of evidence accumulation, whilst simultaneously increasing the ecological validity of the task. The results showed that participants made faster decisions based on less evidence on periodic trials without making significantly more errors. Whilst this meant that there was no difference in accuracy between the periodicity conditions, periodic decisions were more efficient as they required less information and time. A hierarchical implementation of the Drift Diffusion Model (HDDM) using the data from experiment 6 revealed that the periodicity of the noise burst sequence affected the boundary separation component of the model on medium and high difficulty trials. This means that rather than periodicity enhancing the quality of accumulating decision information, on aperiodic trials the decision process simply required more accumulated evidence before terminating. One hypothesis that explains this behaviour, but needs testing, is that aperiodicity made participants less confident and hence unnecessarily conservative in their deliberation time.

Experiment 7 addressed the last unexplored feature of the experimental framework by testing whether rhythmic temporal expectations bias subjective value decisions. This topic remains at the forefront of the investigation and is key to understanding interdependence between temporal expectation and complex decision making. On each trial participants made a preference decision between pairs of musical items that were preceded by a rhythmic sequence of on-screen flashes. This use of dynamic audio-visual stimuli and music aimed to increase the generalisability and ecological validity of value-based decision research, which traditionally uses static images of appetitive food items as stimuli. It was hypothesised that if rhythmic temporal expectations boosted subjective value, participants would be more likely to select the first item of the pair on periodic trials in which both musical items were equally liked. The results showed that contrary to the hypothesis, the periodicity of the on-screen flashes strongly biased neither the preference decisions nor the response times. Decisions were, however, highly sensitive towards previously recorded preferences for each musical item; furthermore, participants were less likely to select an option the more times it was used throughout the session. These findings, when paired with the number of observed preference reversals, validates the use of musical stimuli in the experiment and suggests that musical preferences are the same as or more reliable than more commonly investigated preferences in the decision making literature such as those associated with food and faces. The fact that rhythmic temporal expectations did not influence subjective value decisions may have been due to features of the experimental design, such as the use of rhythmic precursors, which is argued against in chapter 3. For this reason, further research needs to be conducted into the area before this conclusion concerning the effects of temporal expectations on subjective value can be supported.

In summary, the seven experiments produced four key findings that have theoretical implications for future research:

1. The degree of IOI variance in the stimulus proportionally affects decision latencies but not decision accuracy during both complex averaging and simple response time tasks. As this was observed in experiment 5, it is likely that the bias is caused by cognitive processes that are relatively general to perception-action tasks rather than just to complex decision making.

2. Apart from experiment 1 and a subset of participants in experiment 4, rhythmic temporal expectations did not bias the accuracy of complex averaging or subjective value decisions. This finding is contrary to key theories in the temporal expectation literature which are validated using simply perceptual decision making tasks. During complex decision making, however, it appears that the degree to which rhythmic temporal expectations will bias the accuracy of the decision is dependent on contextual factors such as the complexity of the decision, the rate of the stimulus, knowledge about the timing of the stimulus and participant ability at the task.

3. Participants required less decision information in order to respond faster and with the same accuracy on periodic versus aperiodic trials. This suggests that whilst periodicity does not consistently increase the accuracy of complex decisions, it does make them more efficient. This behaviour was associated with a change in the boundary separation of a HDDM and therefore may have occurred due to aperiodicity making participants less confident and unnecessarily conservative in their deliberation time.

4. Periodicity did not appear to influence subjective value representations of musical items when rhythmic precursors were used to generate temporal expectations. Musical preferences are, however, the same or more reliable than commonly investigated preferences in the decision making literature and therefore can be used in future experimentation.

9.2 Theoretical implications

9.2.1 Dynamic inhibition and boosting

Dynamic Attending Theory (DAT), described in section 2.2.1, provides an elegant account of how attending neural oscillations entrain to periodic regularities in a sensory signal. To recap, as entrainment increases, due to fewer expectancy violations, attentional energy concentrates towards the peak of the oscillatory cycle resulting in attentional fluctuations (Large and Jones, 1999). As a result, information becomes easier to detect, memorise and respond to if it falls near to the peak of an entrained cycle, and it is suppressed if it falls far away from the peak. The evidence supporting this theory shows that targets that are presented in time with a preceding periodic pulse are responded to more accurately than those that are presented slightly out of time. The regularity of the pulse is assumed to entrain neural oscillations which then influence how the response targets are processed. An explanation for this behaviour is that entrainment enhances the quality of decision information when it aligns with the period and phase of the entrainment cycle. This notion of perceptual enhancement is commonly used throughout the temporal expectation literature (Klein and Jones, 1996; Escoffier et al., 2010; Walker and King, 2011; Cravo et al., 2013; Lawrance et al., 2014).

A second idea associated with DAT and theories of neural oscillatory entrainment is that in the absence of periodic stimulation, neural oscillations revert from a rhythmic to a continuous processing mode in which information is sampled uniformly (Schroeder and Lakatos, 2009; Henry and Herrmann, 2014). This is caused by the attentional energy that is concentrated at the peaks of an entrained oscillatory cycle becoming more evenly distributed across all moments in time. It therefore becomes more likely that the observer will detect a greater number of irregularly spaced events, but less likely that the quality of the perceptual information will be enhanced. Whilst this idea implies that it should be easier to recall features of periodic versus aperiodic information due to the enhanced quality of information, few studies have explicitly asked this question. For the few that did, the results are inconclusive. Jones et al.'s (2002) experiment 3 showed that whilst aperiodicity resulted in a flatter expectancy profile, it did not reduce response accuracy compared with periodic stimuli. Similarly, Mathewson et al. (2012) showed that accuracy was not significantly different between periodic and aperiodic stimulus conditions. Rohenkohl et al. (2012) and Cravo et al. (2013) on the other hand showed that periodic stimuli were responded to faster and more accurately than aperiodic stimuli. Finally, in this thesis, periodicity had no

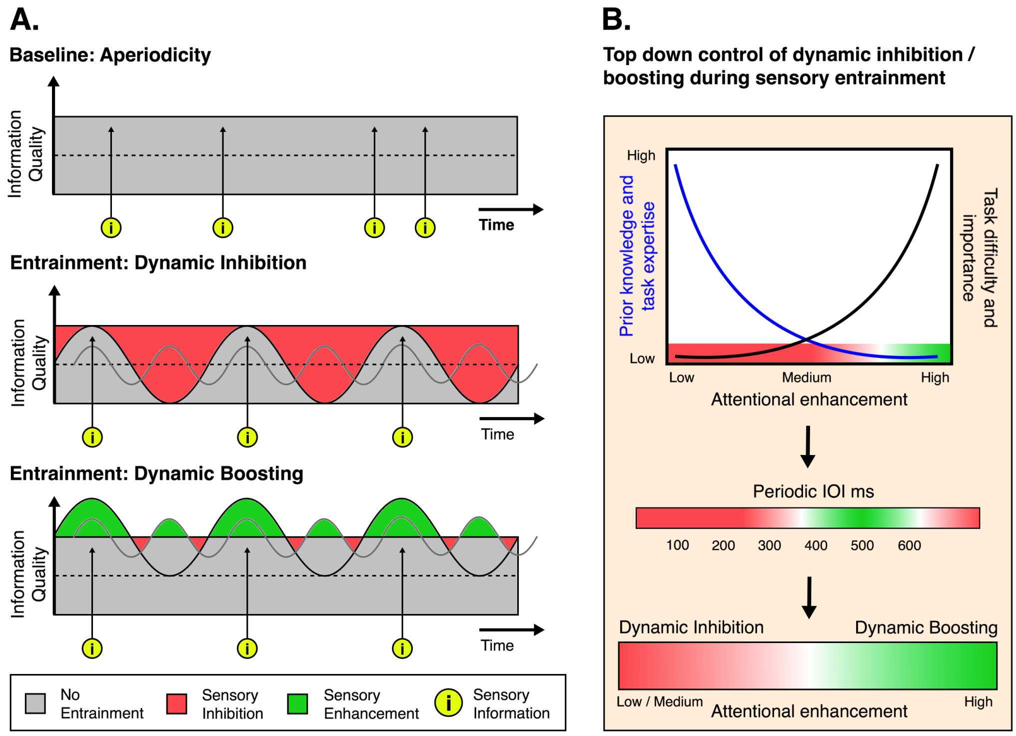


Fig. 9.1 A. Schematic illustration of dynamic inhibition and boosting: A. top panel: continuous processing mode. A. middle panel: dynamic inhibition of out-of-time information. A. bottom panel: dynamic boosting of in-time information. Red shading = inhibited sensory information. Green shading = enhanced sensory information. Grey shading = baseline processing. Yellow circles = time of onset of sensory information. B. Schematic illustration of top-down control mechanism for dynamic boosting: B. top panel: examples of attentional enhancement filters that account for prior knowledge, expertise, task difficulty, task importance. B. middle panel: periodic rate filter. B. bottom panel: classification filter that determines degree of dynamic boosting. Red shading = dynamic inhibition. Green shading = dynamic boosting

effect on the accuracy of complex decisions on four out of the six classification experiments.

What is the reason for this inconsistency and why did periodicity not guarantee more accurate decision making? A theoretical explanation to this question is that rather than enhancing decision information, the default behaviour of periodically entrained neural oscillations is to inhibit. I will call this behaviour "dynamic inhibition". It means that under situations of periodic stimulation, entrained oscillations inhibit the processing of information that does not align with the peak of the oscillation, but importantly, does not simultaneously enhance information processing that does coincide with a peak in entrained oscillations. This is illustrated in the top two panels of figure 9.1A. The top panel shows that when sensory information is aperiodic, a continuous processing mode applies in which all information is processed equally and with the same degree of perceived quality. When the sensory information is periodic (middle panel), oscillatory activity systematically inhibits the quality of perceived sensory information the further the onset of information is from the oscillatory peak. This activity is computationally more efficient than the continuous processing mode and produces the same transduction results as long as the sensory information remains periodic. This model better accounts for the behavioural results in the literature and is compatible with DAT; not only will targets that do not temporally align with a periodic pulse remain harder to detect than targets that do align during sensory entrainment, but periodic stimuli will not necessarily be classified more accurately than aperiodic stimuli.

As demonstrated in experiment 1 and for a subset of participants in experiment 4, there are instances in which periodic stimuli lead to more accurate decision making. I suggest that this is the result of "dynamic boosting" as illustrated in the bottom panel of figure 9.1A. Dynamic boosting is a context specific effect that occurs when the conditions of the decision require attentional enhancement. Its function is to ensure that the quality of sensory information is enhanced when that information aligns with peaks of an entrained oscillatory response but is inhibited when it does not. Dynamic boosting is thus a more computationally expensive procedure than the more general and low-level process of dynamic inhibition and is most likely controlled via a combination of factors.

Figure 9.1B shows an example of what the control mechanism for dynamic boosting might look like. I suggest that during complex decision making at least four filters are used to determine the degree to which attentional enhancement is required for the task (two are illustrated in the top panel of figure 9.1B). The first two filters modify attentional enhancement depending on the degree of prior knowledge and expertise one has about the task (blue curve). The second two filters modify attentional enhancement depending on the difficulty and importance of the task (black curve). The resulting value for attentional enhancement from this stage is most likely a weighted average of all four. The second stage (middle panel) indicates that the rate of the periodic stimulus may have an influencing effect on attentional enhancement, whereby IOIs in the range of 500 $\pm$100 ms promote dynamic boosting. The evidence for this come from experiment 4, in which the slower periodic rate increased the likelihood that participants were more accurate on periodic trials. As with the previous filters, the degree to which the input signal is affected by this filter will most likely be participant-specific, perhaps dependent on preferred beat and tapping rates. Finally, because dynamic inhibition is a computationally more efficient mechanism, only control signals that indicate high attentional enhancement will trigger dynamic boosting (bottom panel).

Dynamic inhibition and boosting are testable hypotheses that together make the following predictions: 1. Sensory information that falls out-of-time with an entrained periodic pulse will be harder to detect and respond to compared with information that aligns with the pulse. 2. Under situations that do not require high attentional enhancement, periodic and aperiodic stimuli will be classed equally accurately. 3. The effect of periodicity on task accuracy during complex decision making can be increased by increasing task difficulty, task importance and by using slower periodic rates within the estimated IOI range of 500 $\pm$100 ms. Features of this theory needing investigation most urgently are the specific characteristics of the attentional enhancement filters, the specific range of IOIs that affect dynamic boosting and the generalisability of the theory across a range of decision types, rates and contexts.

9.2.2 Message passing using temporal probability matching

The most consistent finding throughout this thesis is that decision latencies vary proportionally with IOI variance in a stimulus sequence in a way that does not affect decision accuracy. The larger the IOI variance in a stimulus sequence, the longer the RTs. The fact that this finding was observed in experiment 5 suggests that it is a general characteristic of perception-action tasks and is not specific to complex decision making. It should therefore be modelled as a separate component of the response process.

A possible explanation for why RTs were affected by IOI variance in the stimulus sequence is that responses were regulated by message passing functions using a form of probability matching. This seems plausible because the bias was specific to a statistical quality in the sensory signal (IOI variance) and did not seem to be influenced by the signals content or prior knowledge. To explain, imagine the challenge that an adaptive network of neurons face when trying to transmit a signal as quickly and efficiently as possible: if individual neurons fire too early or too late and there will be a delay in transmission, if they fire on time the information will propagate quickly. Perhaps the reason why RTs were consistently sensitive to IOI variance, even when participants had prior knowledge about the timing of the stimulus (experiment 2), is that IOI variance is a measure of randomness. As IOI variation decreases, the accuracy of neural prediction increases, systematically reducing the mismatch between the predictions and input and proportionally increasing the speed of transmission. This could also explain why effects of IOI variance were not immediate and required prolonged exposure to the rhythmic signal: without sufficient evidence it is impossible to determine whether a signal is random or structured. Whilst the idea of temporal probability matching has strong connections with the topics of reinforcement learning, adaptive resonance theory and predictive coding (Sutton and Barto, 1998; Grossberg, 2000; Summerfield and de Lange, 2014), it is hard to test and thus remains speculative in nature. This is emphasised by the fact that it could occur at any point within the cognitive circuit that transmits the sensory signal, such as during pre-processing, evaluation or memorisation.

9.2.3 Confidence and risk

Experiment 6 demonstrates that stimulus periodicity increases the efficiency of complex averaging decisions by reducing the amount of time and information needed to respond. Temporal probability matching can account for the reduced RTs, but it cannot account for why participants needed less decision information on periodic trials, nor why periodicity affected the boundary separation and not the drift rate of the HDDM. In chapter 7 (section 7.4) I proposed the explanation that stimulus aperiodicity influenced participants' confidence and made them unnecessarily conservative in their deliberation time. A different way to frame this proposal is that aperiodicity influenced perceived risk during the decision process. The more aperiodic the sequence became, the more risk participants associated with responding early. What is interesting about this observation is that the change towards a more delayed decision did not afford the obvious benefit of increasing decision accuracy, which implies that aperiodicity made participants more cautious, but unnecessarily so i.e. without any obvious benefit of the caution. This implies that were participants to have forced themselves to respond as quickly on aperiodic trials as they did on periodic trials, there would have likely been no difference in the outcome of the decision.

The connection of temporal expectations with confidence and risk is novel yet can be contextualised somewhat from the perspective of the explore / exploit dilemma (Stephens and Krebs, 1986; Cohen et al., 2007; Wittmann et al., 2016). The explore / exploit dilemma describes the trade-off that foraging animals face between exploiting current knowledge about their environment or exploring new untested areas in the hope that they may contain a better source of food. A common strategy when confronted with this dilemma is to exploit when the environment is stable, going directly to places where good food is known to be, and explore when the environment is volatile, because volatility increases the likelihood that knowledge about the environment is out-of-date. Sequence aperiodicity may have had a similar effect, with aperiodicity representing volatility. Whilst this logic describes the behaviours observed in experiment 6, it does not necessarily mean that periodicity will always result in exploitation. As discussed by Yu and Dayan (2005) and recently by Geana et al. (2016), if the task becomes overly predictable, such as during periods of prolonged periodicity, participants are likely to revert to exploration to avoid boredom. In terms of task performance this will lead to increased errors due to reduced attention to the task.

The effects of temporal expectations on confidence and risk have untested implications for a range of interactive activities that involve fast adaptive decision making based on rhythmically structured information. For example, a fighter jet pilot should be quicker to decide whether an approaching dot on a radar screen represents a friend or foe if the dot flashes periodically versus aperiodically and may weight possible preparatory responses differently depending on the (a)periodicity of this information. During group musical improvisation, musicians may assign more value to exploring untested and risky ideas when the structure of the interaction is temporally predictable (i.e. when co-performers play in a prolonged predictable manner), whereas they should weight safe and well-practiced ideas more positively when the structure of the interaction is temporally unpredictable (i.e. when mistakes are made or co-performers play out of time). Financial traders interacting in the financial markets may be faster to predict short term market movements if the structure of the on-screen order flow is periodic. They may also be more likely to take additional risk when the structure of market orders is temporally predictable over a long period of time, but less likely to take risks during short bursts of temporal uncertainty. This last point could partially explain why it is common for market participants to report boredom as the reason for engaging in excessive trading (Willman et al., 2006). Finally, similar predictions could be made for gambling situations, with obvious implications for both the auditory and visual stimuli used in fruit machines and the like, and the regularity with which, for example, a croupier deals out cards and the like, and perhaps even how they speak.

9.3 Applications and future work

The experiments reported in this thesis have relevance for both theoretical research and technology. Firstly, the investigation highlighted the fact that not all popular theories in the timing literature, such as the ability of periodicity to enhance perception and action, generalise directly to complex decision making. This observation should encourage psychophysicists, for example, to develop new theories and paradigms for testing the effects of temporal expectations in situations that are more akin to everyday decision making. The new experimental framework described in chapter 3 outlines one approach, but significant progress will only be likely to be made when the topic is investigated using a wide range of varied and independent methods. One application of this thesis is to provide researchers with an example of how this can be done.

A second application relates to information prioritisation and the field of human-machine interface design. In addition to formatting and spatial layout, stimulus timing can be used to either increase or decrease the likelihood that key information is detected, memorised and responded to quickly when users interact with an interface. For example, during tasks where the importance of certain events and information streams are known, goal relevant information could be presented via an interface in time with a medium to slow periodic pulse and goal irrelevant information slightly out of time with a fast periodic pulse. As the importance of information streams change, based on the demands of the task or changing goals, so to can the timed delivery of information. This example of dynamic interface design would be particularly useful for high stress scenarios in which users make fast decisions whilst dealing with multiple sources of dynamically changing information.

Lastly, if future experimentation validates the idea that temporal expectations interact with confidence and risk, bias warning systems can be designed to help people make less biased decisions. Again, a good application for this could be within the financial services industry with the aim of helping traders make better decisions whilst monitoring price change and order flow on computer screens. Whereas the rate of orders can inform about changes in supply and demand, the rhythmic structure of the order flow is random and should be ignored. A warning system could be designed that monitors the variance in the order flow and signals to users when the temporal structure of the market is likely to negatively influence their willingness to expose themselves to risk.

There are three avenues that seem worth investigating in order to support the development of the above applications:

1. Develop a formal computational model of dynamic inhibition and boosting: Whilst the theory outlined in section 9.2.1 provides testable predictions, the next step is to build a mathematical model that can be used to predict trial-by-trial responses. This will require further investigation into the weighting functions that underlie dynamic boosting as well as the formal integration of the theory with the models that define DAT.

2. Investigate whether the rhythmic presentation of time series data on a computer screen biases valuation and directional judgements: The rhythmic presentation of numerical time series data offers a novel and ecologically valid way of investigating the effects of temporal expectations on subjective value, confidence and risk. A simple experiment, that has parallels with the financial markets, is for participants to watch a dynamic stream of numerical data (representing prices) whilst predicting whether the price series will increase or decrease in value at a future point in time. By systematically manipulating the timed presentation of this information, it should be possible to determine how rhythmic temporal expectations affect positive and negative directional judgements. Variants of this experimental design can be used to investigate the effects of stimulus timing on risk by having participants place monetary bets on their directional predictions.

3. Investigate the effects of temporal expectation on cross-modal decision making: Although experiments 1 - 6 investigated auditory decision making, their congruence with many other results in the literature lead me to expect that the present findings are unlikely to be exclusive to auditory processing. For this reason, further work should test whether the findings replicate during complex decisions that rely on visual and tactile information, as well as combinations of different sensory streams.

9.4 Closing remarks

Throughout this thesis I have attempted to bring the temporal expectation and decision making literatures together to investigate the effects of temporal expectation on complex decision making. This involved arguing against common experimental approaches and designing hybrid solutions that better account for a range of interactive decisions. The results demonstrated that existing theories of temporal expectation do not necessarily generalise to complex decision making and that the effects of timing on choice interact with a number of factors associated with prior knowledge, stimulus rate, variance, decision type and task complexity. This does not mean that existing theories, such as DAT, are necessarily wrong, but simply that they lack detail—specifically, they are too narrow—to fully account for the role of timing in complex choice.

As we strive towards developing a unified theory of how our interaction with the environment affects choice it is crucial that timing is not artificially separated from the activities and judgements that comprise the initial source of interest in the topic. This will require developing an application led research agenda and designing new experimental scenarios that more closely reflect real life, and in doing so they should replicate and probe a wide variety of contextualised goal-relevant behaviours. This type of research programme should not only advance scientific knowledge about how the brain operates, leading to more complete computational and biologically plausible models, but also produce a range of applications that can be used to help humans as they attempt to successfully navigate and interact with the world. Whilst timing, expectation and decisions are all fundamental aspects of cognition, it is their interdependence, and the questions that this brings, that makes the idea of future research in this area so appealing.

Bibliography

Ahveninen, J., Kopčo, N., and Jääskeläinen, I. P. (2014). Psychophysics and neuronal bases of sound localization in humans. *Hearing Research*, 307, 86–97.

Albrecht, A. R. and Scholl, B. J. (2010). Perceptually averaging in a continuous visual world: extracting statistical summary representations over time. *Psychological Science*, 21(4), 560–567.

Albrecht, A. R., Scholl, B. J., and Chun, M. M. (2012). Perceptual averaging by eye and ear: computing summary statistics from multimodal stimuli. *Attention, Perception & Psychophysics*, 74(5), 810–815.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, 15(3), 122–131.

Alvarez, G. A. and Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science*, 19(4), 392–398.

Ariely, D. (2001). Seeing sets: representation by statistical properties. *Psychological Science*, 12(2), 157–162.

Armel, K. C., Beaumel, A., and Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making*, 3(5), 396–403.

Arnal, L. H., Doelling, K. B., and Poeppel, D. (2014). Delta–beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex*, 25(9), 3077–3085.

Arnal, L. H. and Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, 16(7), 390–398.

Baayen, R. H. (2013). *languageR: data sets and functions with "Analyzing linguistic data: A practical introduction to statistics"*. R package version 1.4.1.

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.

Baayen, R. H. and Milin, P. (2015). Analyzing reaction times. *International Journal of Psychological Research*, 3(2), 12–28.

Barnes, R. and Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology*, 41(3), 254–311.

Basten, U., Biele, G., Heekeren, H. R., and Fiebach, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences*, 107(50), 21767–21772.

Bech, M. L. and Garratt, R. (2003). The intraday liquidity management game. *Journal of Economic Theory*, 109(2), 198–219.

Beer, R. D. (2014). Dynamical systems and embedded cognition. In Frankish, K. and Ransey, W., editors, *The Cambridge Handbook of Artificial Intelligence*, chapter 6, pages 128–148. Cambridge University Press.

Behrens, T. E., Woolrich, M. W., Walton, M. E., and Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9), 1214–1221.

Bernieri, F. J., Reznick, J. S., and Rosenthal, R. (1988). Synchrony, pseudosynchrony, and dissynchrony: Measuring the entrainment process in mother-infant interactions. *Journal of Personality and Social Psychology*, 54(2), 243–253.

Bolger, D., Trost, W., and Schön, D. (2013). Rhythm implicitly affects temporal orienting of attention across modalities. *Acta Psychologica*, 142(2), 238–244.

Bolton, T. (1894). Rhythm. *The American Journal of Psychology*, 6(2), 145–238.

Bosman, C. A., Womelsdorf, T., Desimone, R., and Fries, P. (2009). A microsaccadic rhythm modulates gamma-band synchronization and behavior. *Journal of Neuroscience*, 29(30), 9471–9480.

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

Breska, A. and 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.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., and Drake, C. (2003). The "ticktock" of our internal clock. direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*, 14(4), 362–366.

Brown, A. D. and Stecker, G. C. (2011). Temporal weighting functions for interaural time and level differences. ii. the effect of binaurally synchronous temporal jitter. *The Journal of the Acoustical Society of America*, 129(1), 293–300.

Busch, N. A. and VanRullen, R. (2010). Spontaneous eeg oscillations reveal periodic sampling of visual attention. *Proceedings of the National Academy of Sciences*, 107(37), 16048–16053.

Buzsáki, G. and Chrobak, J. J. (1995). Temporal structure in spatially organized neuronal ensembles: a role for interneuronal networks. *Current Opinion in Neurobiology*, 5(4), 504–510.

Buzsáki, G. and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, 304(5679), 1926–1929.

Calderone, D. J., Lakatos, P., Butler, P. D., and Castellanos, F. X. (2014). Entrainment of neural oscillations as a modifiable substrate of attention. *Trends in Cognitive Sciences*, 18(6), 300–309.

Carlile, S., Hyams, S., and Delaney, S. (2001). Systematic distortions of auditory space perception following prolonged exposure to broadband noise. *The Journal of the Acoustical Society of America*, 110(1), 416–424.

Carpenter, R., Reddi, B., and Anderson, A. (2009). A simple two-stage model predicts response time distributions. *The Journal of Physiology*, 587(16), 4051–4062.

Cavanagh, J. F., Wiecki, T. V., Cohen, M. X., Figueroa, C. M., Samanta, J., Sherman, S. J., and Frank, M. J. (2011). Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nature Neuroscience*, 14(11), 1462–1467.

Cavanagh, J. F., Wiecki, T. V., Kochar, A., and Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology: General*, 143(4), 1476–1488.

Cheadle, S., Wyart, V., Tsetsos, K., Myers, N., De Gardelle, V., Castañón, S. H., and Summerfield, C. (2014). Adaptive gain control during human perceptual choice. *Neuron*, 81(6), 1429–1441.

Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2008). Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience*, 20(2), 226–239.

Chib, V. S., Rangel, A., Shimojo, S., and O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience*, 29(39), 12315–12320.

Chong, S. C. and Treisman, A. (2003). Representation of statistical properties. *Vision Research*, 43(4), 393–404.

Chong, S. C. and Treisman, A. (2005a). Attentional spread in the statistical processing of visual displays. *Attention, Perception & Psychophysics*, 67(1), 1–13.

Chong, S. C. and Treisman, A. (2005b). Statistical processing: computing the average size in perceptual groups. *Vision Research*, 45(7), 891–900.

Christopoulos, V., Bonaiuto, J., and Andersen, R. A. (2015). A biologically plausible computational theory for value integration and action selection in decisions with competing alternatives. *PLoS Comput Biol*, 11(3), e1004104.

Clark, H. H. (2005). Coordinating with each other in a material world. *Discourse Studies*, 7(4-5), 507–525.

Clayton, M., Sager, R., and Will, U. (2005). In time with the music: the concept of entrainment and its significance for ethnomusicology. In *European meetings in ethnomusicology.*, volume 11, pages 1–82. Romanian Society for Ethnomusicology.

Cohen, J. D., McClure, S. M., and Angela, J. Y. (2007). Should i stay or should i go? how the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362(1481), 933–942.

Correa, A. (2010). Enhancing behavioural performance by visual temporal orienting. In Nobre, A. C. and Coull, J. T., editors, *Attention and time*, chapter 26, pages 357–370. Oxford University Press, Oxford.

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

Correa, A. and Nobre, A. C. (2008). Neural modulation by regularity and passage of time. *Journal of Neurophysiology*, 100(3), 1649–1655.

Coull, J. T. and Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both pet and fmri. *Journal of Neuroscience*, 18(18), 7426–7435.

Cravo, A. M., Rohenkohl, G., Wyart, V., and 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.

Csibra, G. and Gergely, G. (2006). Social learning and social cognition: the case for pedagogy. In Munakata, Y. and Johnson, M. H., editors, *Processes of Change in Brain and Cognitive Development. Attention and Performance XXI*, pages 249–274. Oxford University Press.

Curtanda, D., Correa, A., and Sanabria, D. (2015). Auditory temporal preparation induced by rhythmic cues during concurrent auditory working memory tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 41(3), 790–797.

Dahmen, J. C., Keating, P., Nodal, F. R., Schulz, A. L., and King, A. J. (2010). Adaptation to stimulus statistics in the perception and neural representation of auditory space. *Neuron*, 66(6), 937–948.

Davis, R. (1965). Expectancy and intermittency. *Quarterly Journal of Experimental Psychology*, 17(1), 75–78.

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In Delgado, M. R., Phelps, E., and Robbins, T., editors, *Decision making, affect, and learning: Attention and performance XXIII*, chapter 1, pages 3–38. Oxford University Press, Oxford.

De Gardelle, V. and Summerfield, C. (2011). Robust averaging during perceptual judgment. *Proceedings of the National Academy of Sciences*, 108(32), 13341–13346.

de la Rosa, M. D., Sanabria, D., Mariagrazia, C., and Correa, A. (2012). Temporal preparation driven by rhythms is resisitant to working memory interference. *Frontiers in Psychology*, 3(308), 1–9.

Dingle, R. N., Hall, S. E., and Phillips, D. P. (2010). A midline azimuthal channel in human spatial hearing. *Hearing Research*, 268(1), 67–74.

Doherty, J. R., Rao, A., Mesulam, M. M., and Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *Journal of Neuroscience*, 25(36), 8259–8266.

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

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., and Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *Journal of Neuroscience*, 32(11), 3612–3628.

Dunovan, K. E., Tremel, J. J., and Wheeler, M. E. (2014). Prior probability and feature predictability interactively bias perceptual decisions. *Neuropsychologia*, 61, 210–221.

Ellis, R. J. and Jones, M. R. (2009). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance*, 35(1), 264–280.

Engel, A. K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience*, 2(10), 704–716.

Escoffier, N., Sheng, D. Y. J., and Schirmer, A. (2010). Unattended musical beats enhance visual processing. *Acta Psychologica*, 135(1), 12–16.

Filipic, S., Tillmann, B., and Bigand, E. (2010). Judging familiarity and emotion from very brief musical excerpts. *Psychonomic Bulletin & Review*, 17(3), 335–341.

Forstmann, B., Ratcliff, R., and Wagenmakers, E.-J. (2016). Sequential sampling models in cognitive neuroscience: advantages, applications, and extensions. *Annual Review of Psychology*, 67(1), 641–666.

Frank, M. J., Gagne, C., Nyhus, E., Masters, S., Wiecki, T. V., Cavanagh, J. F., and Badre, D. (2015). fmri and eeg predictors of dynamic decision parameters during human reinforcement learning. *Journal of Neuroscience*, 35(2), 485–494.

Frederick, S., Loewenstein, G., and O'Donoghue, T. (2002). Time discounting and time preference: a critical review. *Journal of Economic Literature*, 40(2), 351–401.

Frith, C. D. (2007). The social brain? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362(1480), 671–678.

Frith, C. D. and Frith, U. (2007). Social cognition in humans. *Current Biology*, 17(16), R724–R732.

Frith, U. and Frith, C. (2010). The social brain: allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1537), 165–176.

Fritts, L. (2012). *Electronic music studio.* University of Iowa, http://theremin.music.uiowa.edu/MIS.html.

Gallagher, S. (2009). Philosophical antecendents to situated cognition. In Robbins, P. and Aydede, M., editors, *The Cambridge Handbook of Situated Cognition*, pages 35–51. Cambridge University Press.

Geana, A., Wilson, R., Daw, N., and Cohen, J. (2016). Boredom, information-seeking and exploration. In *Proc. 38th Annu. Conf. Cogn. Sci. Soc.*, number http://www.princeton.edu/ ndaw/gwdc16.pdf, pages 1751–56, Philadelphia.

Geiser, E., Notter, M., and Gabrieli, J. D. (2012). A corticostriatal neural system enhances auditory perception through temporal context processing. *Journal of Neuroscience*, 32(18), 6177–6182.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457–472.

Gelman, A, C. J. S. H. and Rubin, D. (1995). *Bayesian data analysis*, volume 1. CRC Press, New York.

Gjerdingen, R. O. and Perrott, D. (2008). Scanning the dial: the rapid recognition of music genres. *Journal of New Music Research*, 37(2), 93–100.

Glass, L. (2001). Synchronization and rhythmic processes in physiology. *Nature*, 410(6825), 277–284.

Glimcher, P. W. (2004). *Decisions, uncertainty, and the brain: the science of neuroeconomics*. MIT press, Cambridge MA.

Gold, J. and Heekeren, H. (2014). Neural mechanisms for perceptual decision making. In Glimcher, P. W. and Fehr, E., editors, *Neuroeconomics: decision making and the brain*, volume 2, chapter 19, pages 355–372. Academic Press.

Gold, J. I. and Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, 5(1), 10–16.

Gold, J. I. and Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30, 535–574.

Goupell, M. J., Laback, B., and Majdak, P. (2009). Enhancing sensitivity to interaural time differences at high modulation rates by introducing temporal jitter. *The Journal of the Acoustical Society of America*, 126(5), 2511–2521.

Greatrex, D. (2011). Attention in multitask musical performance: a cross modal investigation into dynamic attending. Master's thesis.

Green, D. and Swets, J. A. (1966). *Signal detection theory and psychophysics.* Wiley, New York.

Green, L., Fristoe, N., and Myerson, J. (1994). Temporal discounting and preference reversals in choice between delayed outcomes. *Psychonomic Bulletin & Review*, 1(3), 383–389.

Grossberg, S. (2000). Adaptive resonance theory. In *Encyclopedia of Cognitive Science*, London. Macmillan Reference Ltd.

Haberman, J., Harp, T., and Whitney, D. (2009). Averaging facial expression over time. *Journal of Vision*, 9(11), 1–1.

Hafter, E. and Dye Jr, R. (1983). Detection of interaural differences of time in trains of high–frequency clicks as a function of interclick interval and number. *The Journal of the Acoustical Society of America*, 73(2), 644–651.

Hare, T. A., Camerer, C. F., Knoepfle, D. T., O'Doherty, J. P., and Rangel, A. (2010). Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *Journal of Neuroscience*, 30(2), 583–590.

Hawkins, S. (2010). Phonological features, auditory objects, and illusions. *Journal of Phonetics*, 38(1), 60–89.

Henry, M. J. and Herrmann, B. (2014). Low-frequency neural oscillations support dynamic attending in temporal context. *Timing & Time Perception*, 2(1), 62–86.

Henry, M. J. and Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences*, 109(49), 20095–20100.

Herrmann, B., Henry, M. J., Haegens, S., and Obleser, J. (2016). Temporal expectations and neural amplitude fluctuations in auditory cortex interactively influence perception. *NeuroImage*, 124(Part A), 487–497.

Hickok, G., Farahbod, H., and Saberi, K. (2015). The rhythm of perception. entrainment to acoustic rhythms induces subsequent perceptual oscillation. *Psychological science*, 26(7), 1006–1013.

Hopkins, K. and Moore, B. C. (2010). Development of a fast method for measuring sensitivity to temporal fine structure information at low frequencies. *International Journal of Audiology*, 49(12), 940–946.

Howard, D. M. and Angus, J. (2009). *Acoustics and psychoacoustics*. Taylor & Francis, New York.

Iversen, J. R., Repp, B. H., and Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences*, 1169(1), 58–73.

Iyengar, S. S. and Lepper, M. R. (2000). When choice is demotivating: can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79(6), 995–1006.

Jahfari, S., Ridderinkhof, K. R., and Scholte, H. S. (2013). Spatial frequency information modulates response inhibition and decision-making processes. *PLoS ONE*, 8(10), e76467.

Jepma, M., Wagenmakers, E.-J., and Nieuwenhuis, S. (2012). Temporal expectation and information processing: a model-based analysis. *Cognition*, 122(3), 426–441.

Johnson, B. W. and Hautus, M. J. (2010). Processing of binaural spatial information in human auditory cortex: neuromagnetic responses to interaural timing and level differences. *Neuropsychologia*, 48(9), 2610–2619.

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

Jones, M. R. (1990). Musical events and models of musical time. In Block, R. A., editor, *Cognitive models of psychological time*, pages 207–240. Psychology Press, New York.

Jones, M. R. (2004). Attention and timing. In Neufoff, J., editor, *Ecological psychoacoustics*, pages 49–85. Elsevier Academic Press, New York.

Jones, M. R. (2009). Musical time. In Hallam, S., Cross, I., and Thaut, M., editors, *The Oxford handbook of music psychology*, volume 1, chapter 8, pages 81–92. Oxford University Press, Oxford.

Jones, M. R. (2010). Attending to sound patterns and the role of entrainment. In Nobre, A. C. and Coull, J. T., editors, *Attention and time*, chapter 23, pages 317–330. Oxford University Press, Oxford.

Jones, M. R. and Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 96(3), 459–491.

Jones, M. R., Johnston, H. M., and Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive Psychology*, 53(1), 59–96.

Jones, M. R., Moynihan, H., MacKenzie, N., and Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, 13(4), 313–319.

Kable, J. W. and Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, 10(12), 1625–1633.

Kable, J. W. and Glimcher, P. W. (2009). The neurobiology of decision: consensus and controversy. *Neuron*, 63(6), 733–745.

Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Attention, Perception, & Psychophysics*, 49(3), 227–229.

Kanizsa, G. (1985). Seeing and thinking. *Acta Psychologica*, 59(1), 23–33.

Kashino, M. and Nishida, S. (1998). Adaptation in the processing of interaural time differences revealed by the auditory localization aftereffect. *The Journal of the Acoustical Society of America*, 103(6), 3597–3604.

Keating, P. and King, A. J. (2015). Sound localization in a changing world. *Current Opinion in Neurobiology*, 35, 35–43.

Kiani, R., Hanks, T. D., and Shadlen, M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *Journal of Neuroscience*, 28(12), 3017–3029.

Kilner, J. M., Friston, K. J., and Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, 8(3), 159–166.

Klein, J. M. and Jones, M. R. (1996). Effects of attentional set and rhythmic complexity on attending. *Attention, Perception & Psychophysics*, 58(1), 34–46.

Klemmer, E. T. (1956). Time uncertainty in simple reaction time. *Journal of Experimental Psychology*, 51(3), 179.

Knight, S., Spiro, N., and Cross, I. (2016). Look, listen and learn: exploring effects of passive entrainment on social judgements of observed others. *Psychology of Music*, 45(1), 99–115.

Knoblauch, K. and Maloney, L. T. (2012). *Modeling psychophysical data in R*. Springer, New York.

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13(10), 1292–1298.

Krajbich, I., Lu, D., Camerer, C., and Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology*, 3.

Krajbich, I. and Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, 108(33), 13852–13857.

Krumhansl, C. L. (2010). Plink: "thin slices" of music. *Music Perception: An Interdisciplinary Journal*, 27(5), 337–354.

Kruschke, J. (2014). *Doing Bayesian data analysis: a tutorial with R, JAGS, and Stan*. Academic Press, London.

Kumpik, D. P., Kacelnik, O., and King, A. J. (2010). Adaptive reweighting of auditory localization cues in response to chronic unilateral earplugging in humans. *The Journal of Neuroscience*, 30(14), 4883–4894.

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

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, 320(5872), 110–113.

Laming, D. R. J. (1968). *Information theory of choice-reaction times*. Academic Press, Oxford.

Lange, K. (2009). Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain and Cognition*, 69(1), 127–137.

Lange, K. and Röder, B. (2006). Orienting attention to points in time improves stimulus processing both within and across modalities. *Journal of Cognitive Neuroscience*, 18(5), 715–729.

Large, E. W. (1994). *Dynamic representation of musical structure*. PhD thesis, The Ohio State University, Columbus.

Large, E. W. (2008). Resonating to musical rhythm: theory and experiment. In Grondin, S., editor, *The psychology of time*, chapter 6, pages 189–232. Emerald, West Yorkshire.

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

Large, E. W. and Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, 1169(1), 46–57.

Lawrance, E. L., Harper, N. S., Cooke, J. E., and Schnupp, J. W. (2014). Temporal predictability enhances auditory detection. *The Journal of the Acoustical Society of America*, 135(6), EL357–EL363.

Lebreton, M., Jorge, S., Michel, V., Thirion, B., and Pessiglione, M. (2009). An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron*, 64(3), 431–439.

Lim, S.-L., O'Doherty, J. P., and Rangel, A. (2011). The decision value computations in the vmpfc and striatum use a relative value code that is guided by visual attention. *Journal of Neuroscience*, 31(37), 13214–13223.

Link, S. and Heath, R. (1975). A sequential theory of psychological discrimination. *Psychometrika*, 40(1), 77–105.

Lo, S. and Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, 6.

London, J. (2012). *Hearing in time: psychological aspects of musical meter*. Oxford University Press, Oxford, 2 edition.

Louie, K. and De Martino, B. (2013). The neurobiology of context-dependent valuation and choice. In Glimcher, P. W. and Fehr, E., editors, *Neuroeconomics: Decision Making and the Brain*, chapter 24, pages 455–476. Academic Press, London, 2 edition.

Maier, J. K., Hehrmann, P., Harper, N. S., Klump, G. M., Pressnitzer, D., and McAlpine, D. (2012). Adaptive coding is constrained to midline locations in a spatial listening task. *Journal of Neurophysiology*, 108(7), 1856–1868.

Maier, J. K., McAlpine, D., Klump, G. M., and Pressnitzer, D. (2010). Context effects in the discriminability of spatial cues. *Journal of the Association for Research in Otolaryngology*, 11(2), 319–328.

Marchant, J. L. and Driver, J. (2012). Visual and audiovisual effects of isochronous timing on visual perception and brain activity. *Cerebral Cortex*, 23(6), 1290–1298.

Marsh, K. L., Richardson, M. J., and Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, 1(2), 320–339.

Mathewson, K. E., Fabiani, M., Gratton, G., Beck, D. M., and Lleras, A. (2010). Rescuing stimuli from invisibility: inducing a momentary release from visual masking with pre-target entrainment. *Cognition*, 115(1), 186–191.

Mathewson, K. E., Prudhomme, C., Fabiani, M., Beck, D. M., Lleras, A., and Gratton, G. (2012). Making waves in the stream of consciousness: entraining oscillations in eeg alpha and fluctuations in visual awareness with rhythmic visual stimulation. *Journal of Cognitive Neuroscience*, 24(12), 2321–2333.

McAuley, J. D. (1995). *Perception of time as phase: toward an adaptive-oscillator model of rhythmic pattern processing.* PhD thesis, Indiana University, Bloomington, IN.

McAuley, J. D. (2010). Tempo and rhythm. In Jones, M. R., Fay, R., and Popper, A., editors, *Music perception*, chapter 6, pages 165–199. Springer, New York.

McGrath, J. E. and Kelly, J. R. (1986). *The Guilford social psychology series. Time and human interaction: toward a social psychology of time.* Guilford Press, New York.

Michael, E., De Gardelle, V., and Summerfield, C. (2014). Priming by the variability of visual information. *Proceedings of the National Academy of Sciences*, 111(21), 7873–7878.

Miller, J. E., Carlson, L. A., and McAuley, J. D. (2012). When what you hear influences when you see. listening to an auditory rhythm influences the temporal allocation of visual attention. *Psychological Science*, 24(1), 11–18.

Mills, A. W. (1958). On the minimum audible angle. *The Journal of the Acoustical Society of America*, 30(4), 237–246.

Moelants, D. (2002). Preferred tempo reconsidered. In Stevens, C., Burnham, D., McPherson, E., Schubert, E., and Renwick, J., editors, *Proceedings of the 7th international conference on music perception and cognition*, volume 2002, pages 580–583, Sydney.

Montague, P. R. and Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, 36(2), 265–284.

Moore, B. (2012). *An introduction to the psychology of hearing.* Emerald Group, Bingley, 6th edition.

Morillon, B., Schroeder, C. E., and Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nature Communications*, 5(5255).

Morillon, B., Schroeder, C. E., Wyart, V., and Arnal, L. H. (2016). Temporal prediction in lieu of periodic stimulation. *Journal of Neuroscience*, 36(8), 2342–2347.

Nelken, I. and de Cheveigné, A. (2013). An ear for statistics. *Nature Neuroscience*, 16(4), 381–382.

Newhall, S. (1923). Effects of attention on the intensity of cutaneous pressure and on visual brightness. In Woodworth, R. S., editor, *Archives of Psychology*, number 61. New York.

Ng, B. S. W., Schroeder, T., and Kayser, C. (2012). A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. *Journal of Neuroscience*, 32(35), 12268–12276.

Nie, Q.-Y., Maurer, M., Müller, H. J., and Conci, M. (2016). Inhibition drives configural superiority of illusory gestalt: Combined behavioral and drift–diffusion model evidence. *Cognition*, 150, 150–162.

Niemi, P. and Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin*, 89(1), 133.

Nobre, A. C., Correa, A., and Coull, J. T. (2007). The hazards of time. *Current Opinion in Neurobiology*, 17(4), 465–470.

Nobre, A. C. and Rohenkohl, G. (2014). Time for the fourth dimension in attention. In Nobre, A. C. and Kastner, S., editors, *The Oxford handbook of attention*, chapter 24, pages 676–721. Oxford University Press, Oxford.

OED (2016). *"goal, n." OED Online.*, volume Web. 18 November 2016. Oxford University Press.

Osman, A., Moore, C. M., and Ulrich, R. (1995). Bisecting rt with lateralized readiness potentials: precue effects after lrp onset. *Acta Psychologica*, 90(1–3), 111–127.

Padoa-Schioppa, C. and Assad, J. A. (2006). Neurons in orbitofrontal cortex encode economic value. *Nature*, 441(7090), 223–226.

Phillips, D. P. and Hall, S. E. (2005). Psychophysical evidence for adaptation of central auditory processors for interaural differences in time and level. *Hearing Research*, 202(1), 188–199.

Piazza, E. A., Sweeny, T. D., Wessel, D., Silver, M. A., and Whitney, D. (2013). Humans use summary statistics to perceive auditory sequences. *Psychological Science*, 24(8), 1389–1397.

Piazza, M. and Veronique, L. (2009). How humans count: Numerosity and the parietal cortex. *The Neuroscientist*, 15, 262–273.

Pikovsky, A., Rosenblum, M., and Kurths, J. (2001). *Synchronization: a universal concept in nonlinear sciences*. Cambridge University Press, Cambridge.

Plassmann, H., O'Doherty, J., and Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of Neuroscience*, 27(37), 9984–9988.

Posner, M. I., Snyder, C. R., and Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109(2), 160–174.

Prins, N. and Kingdom, F. (2009). Palamedes: Matlab routines for analyzing psychophysical data. www.palamedestoolbox.org.

Rangel, A., Camerer, C., and Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, 9(7), 545–556.

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

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114(3), 510–532.

Ratcliff, R. and Childers, R. (2015). Individual differences and fitting methods for the two-choice diffusion model of decision making. *Decision*, 2(4), 237–279.

Ratcliff, R. and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922.

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

Rimmele, J., Jolsvai, H., and Sussman, E. (2011). Auditory target detection is affected by implicit temporal and spatial expectations. *Journal of Cognitive Neuroscience*, 23(5), 1136–1147.

Rohenkohl, G., Cravo, A. M., Wyart, V., and Nobre, A. C. (2012). Temporal expectation improves the quality of sensory information. *Journal of Neuroscience*, 32(24), 8424–8428.

Rohenkohl, G. and Nobre, A. C. (2011). Alpha oscillations related to anticipatory attention follow temporal expectations. *Journal of Neuroscience*, 31(40), 14076–14084.

Rushworth, M. F., Mars, R. B., and Summerfield, C. (2009). General mechanisms for making decisions? *Current Opinion in Neurobiology*, 19(1), 75–83.

Sameiro-Barbosa, C. M. and Geiser, E. (2016). Sensory entrainment mechanisms in auditory perception: Neural synchronization cortico-striatal activation. *Frontiers in Neuroscience*, 10(361).

Sanabria, D., Capizzi, M., and Correa, A. (2011). Rhythms that speed you up. *Journal of Experimental Psychology: Human Perception and Performance*, 37(1), 236–244.

Sanabria, D. and Correa, A. (2013). Electrophysiological evidence of temporal preparation driven by rhythms in audition. *Biological Psychology*, 92(2), 98–105.

Schroeder, C. E. and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, 32(1), 9–18.

Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, 10(2), 70–76.

Simons, D. J. and Chabris, C. F. (1999). Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception*, 28(9), 1059–1074.

Skyrms, B. (2001). The stag hunt. In *Proceedings and Addresses of the American Philosophical Association*, volume 75, pages 31–41.

Smith, P. L. and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, 27(3), 161–168.

Smith, P. L. and Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology*, 32(2), 135–168.

Snyder, J. S. and Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, 24(1), 117–126.

Spitzer, B., Blankenburg, F., and Summerfield, C. (2016). Rhythmic gain control during supramodal integration of approximate number. *NeuroImage*, 129, 470–479.

Starns, J. J. and Ratcliff, R. (2010). The effects of aging on the speed–accuracy compromise: boundary optimality in the diffusion model. *Psychology and Aging*, 25(2), 377–390.

Stefanics, G., Hangya, B., Hernádi, I., Winkler, I., Lakatos, P., and Ulbert, I. (2010). Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *Journal of Neuroscience*, 30(41), 13578–13585.

Stephens, D. W. and Krebs, J. R. (1986). *Foraging theory.* Princeton University Press, Princeton.

Strogatz, S. H., Stewart, I., et al. (1993). Coupled oscillators and biological synchronization. *Scientific American–American Edition*, 269, 68–68.

Summerfield, C. and de Lange, F. P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. *Nature Reviews Neuroscience*, 15(11), 745–756.

Summerfield, C. and Egner, T. (2014). Attention and decision-making. In Nobre, A. C. and Kastner, S., editors, *The Oxford handbook of attention*, chapter 29. Oxford University Press, Oxford.

Summerfield, C. and Tsetsos, K. (2012). Building bridges between perceptual and economic decision-making: neural and computational mechanisms. *Frontiers in Neuroscience*, 6(70).

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement learning: an introduction.* MIT press, Cambridge MA.

Taylor, T. E. and Lupker, S. J. (2001). Sequential effects in naming: a time-criterion account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1), 117–138.

Townsend, J. T. and Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge University Press, Cambridge.

Tsetsos, K., Chater, N., and Usher, M. (2012). Salience driven value integration explains decision biases and preference reversal. *Proceedings of the National Academy of Sciences*, 109(24), 9659–9664.

Tsujimoto, S. and Sawaguchi, T. (2005). Neuronal activity representing temporal prediction of reward in the primate prefrontal cortex. *Journal of Neurophysiology*, 93(6), 3687–3692.

Tversky, A. and Simonson, I. (1993). Context-dependent preferences. *Management Science*, 39(10), 1179–1189.

Usher, M. and McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological Review*, 108(3), 550–592.

van Noorden, L. and Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research*, 28(1), 43–66.

Vandekerckhove, J. and Tuerlinckx, F. (2008). Diffusion model analysis with matlab: a dmat primer. *Behavior Research Methods*, 40(1), 61–72.

Velasco, M. J. and Large, E. W. (2011). Pulse detection in syncopated rhythms using neural oscillators. *Pulse*, 1(2), 3–4.

Venables, W. N. and Ripley, B. D. (2002). *Modern applied statistics with S*. Springer, New York, fourth edition. ISBN 0-387-95457-0.

Vesper, C., van der Wel, R. P., Knoblich, G., and Sebanz, N. (2011). Making oneself predictable: reduced temporal variability facilitates joint action coordination. *Experimental Brain Research*, 211(3-4), 517–530.

Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, 13(1), 37–58.

Vickers, D. (1979). *Decision processes in visual perception*. Academic Press, London.

Wagenmakers, E.-J. (2009). Methodological and empirical developments for the ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, 21(5), 641–671.

Wald, A. (1947). *Sequential analysis*. Dover Publications, New York.

Walker, K. M. and King, A. J. (2011). Auditory neuroscience: temporal anticipation enhances cortical processing. *Current Biology*, 21(7), R251–R253.

Wang, X.-J. (2008). Decision making in recurrent neuronal circuits. *Neuron*, 60(2), 215–234.

Warren, R. M. (1984). Perceptual restoration of obliterated sounds. *Psychological Bulletin*, 96(2), 371–383.

Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L. (1993). Localization using nonindividualized head-related transfer functions. *The Journal of the Acoustical Society of America*, 94(1), 111–123.

Wiecki, T. V., Sofer, I., and Frank, M. J. (2013). Hddm: hierarchical bayesian estimation of the drift-diffusion model in python. *Frontiers in neuroinformatics*, 7(14).

Willman, P., Fenton-O'Creevy, M., Nicholson, N., and Soane, E. (2006). Noise trading and the management of operational risk; firms, traders and irrationality in financial markets. *Journal of Management Studies*, 43(6), 1357–1374.

Wilson, M. and Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review*, 12(6), 957–968.

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., and Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences*, 106(7), 2468–2471.

Wittmann, M. K., Kolling, N., Akaishi, R., Chau, B. K., Brown, J. W., Nelissen, N., and Rushworth, M. F. (2016). Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nature Communications*, 7(12327).

Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 358(1431), 593–602.

Woodrow, H. (1914). The measurement of attention. *The Psychological Monographs*, 17(5), i–158.

Wright, B. A. and Fitzgerald, M. B. (2001). Different patterns of human discrimination learning for two interaural cues to sound–source location. *Proceedings of the National Academy of Sciences*, 98(21), 12307–12312.

Wundt, W. (1874/1904). *Principles of physiological psychology*, volume 1. Swan Sonnenschein, London.

Wyart, V., De Gardelle, V., Scholl, J., and Summerfield, C. (2012). Rhythmic fluctuations in evidence accumulation during decision making in the human brain. *Neuron*, 76(4), 847–858.

Wyart, V., Myers, N. E., and Summerfield, C. (2015). Neural mechanisms of human perceptual choice under focused and divided attention. *Journal of Neuroscience*, 35(8), 3485–3498.

Yost, W. A. and Dye Jr, R. H. (1988). Discrimination of interaural differences of level as a function of frequency. *The Journal of the Acoustical Society of America*, 83(5), 1846–1851.

Yu, A. J. and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, 46(4), 681–692.

Zanto, T. P., Snyder, J. S., and Large, E. W. (2006). Neural correlates of rhythmic expectancy. *Advances in Cognitive Psychology*, 2(2-3), 221–231.

Zhang, J. and Rowe, J. B. (2014). Dissociable mechanisms of speed-accuracy tradeoff during visual perceptual learning are revealed by a hierarchical drift-diffusion model. *Frontiers in Neuroscience*, 8(69).

Appendix A

Experiments 1 and 2

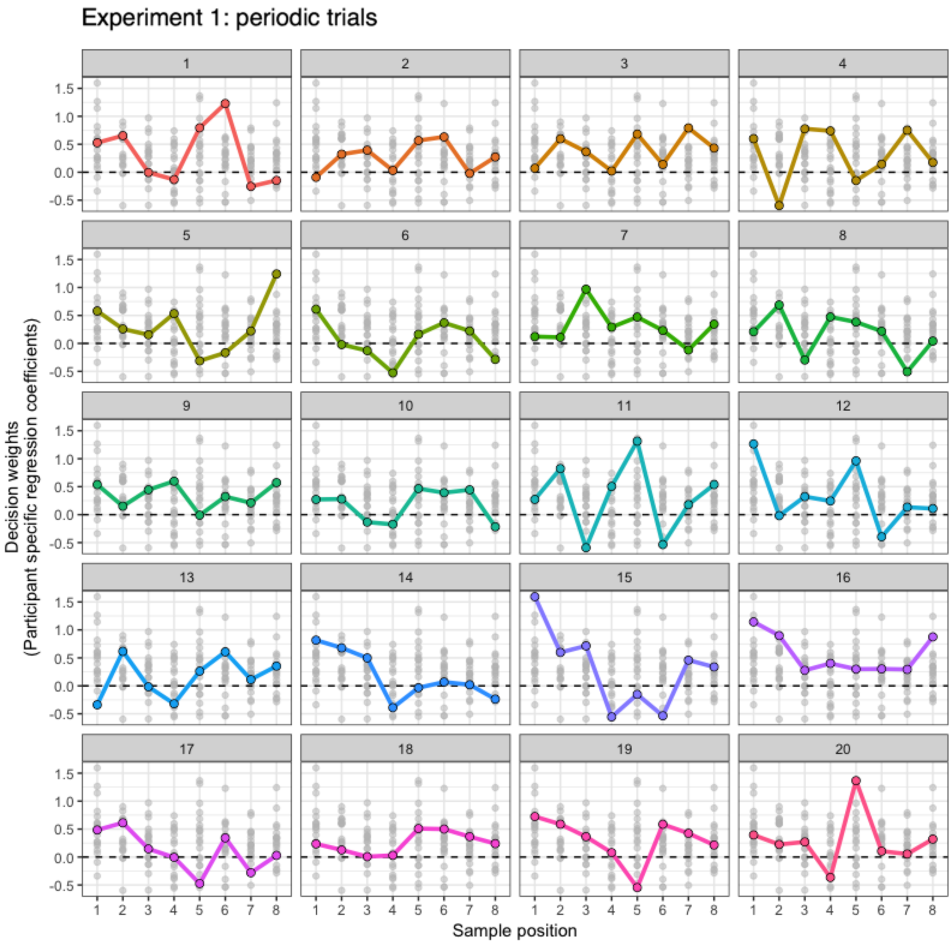


Fig. A.1 Expt. 1. Periodic trials: mean logistic regression coefficients indexing the 8 decision-weights for each position in the noise burst sequence for each participant. Panels: individual participants. Grey dots: All group averaged coefficients.

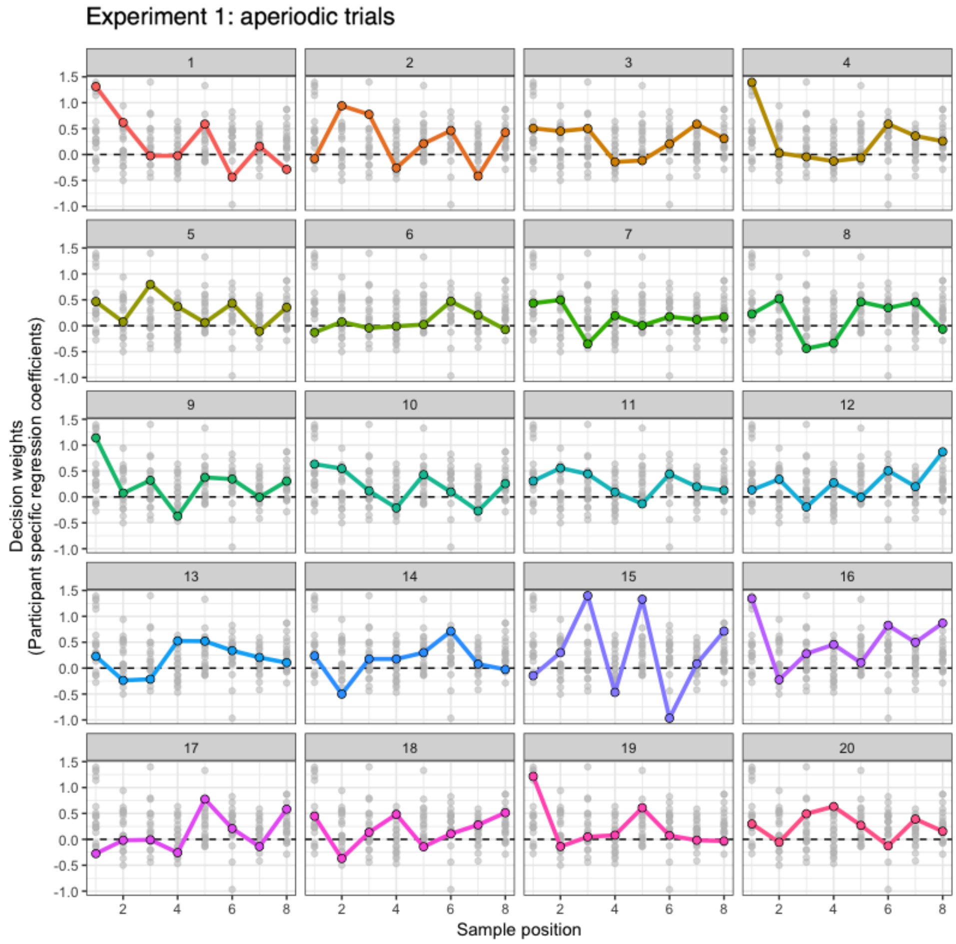


Fig. A.2 Expt. 1. Aperiodic trials: same structure as in figure A.1.

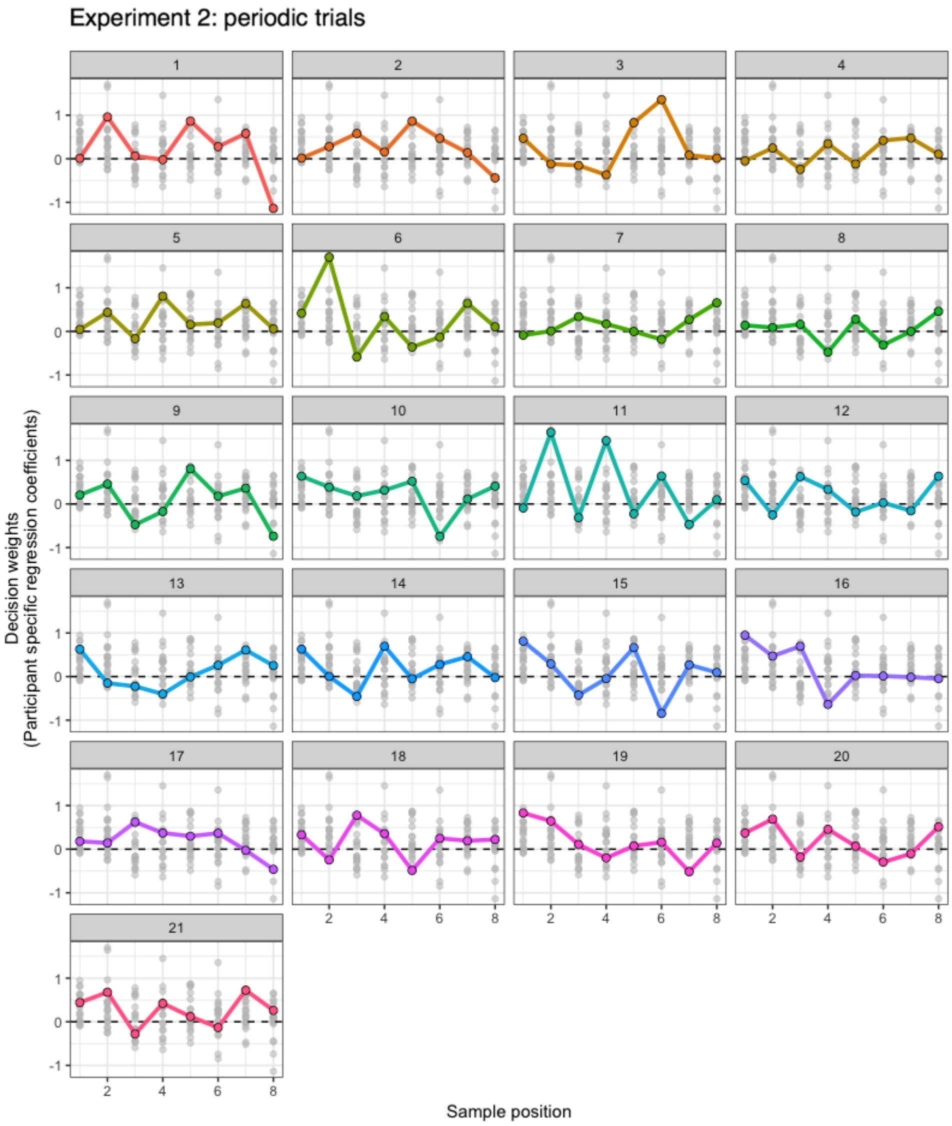


Fig. A.3 Expt. 2. Periodic trials: mean logistic regression coefficients indexing the 8 decision-weights for each position in the noise burst sequence for each participant. Panels: individual participants. Grey dots: All group averaged coefficients.

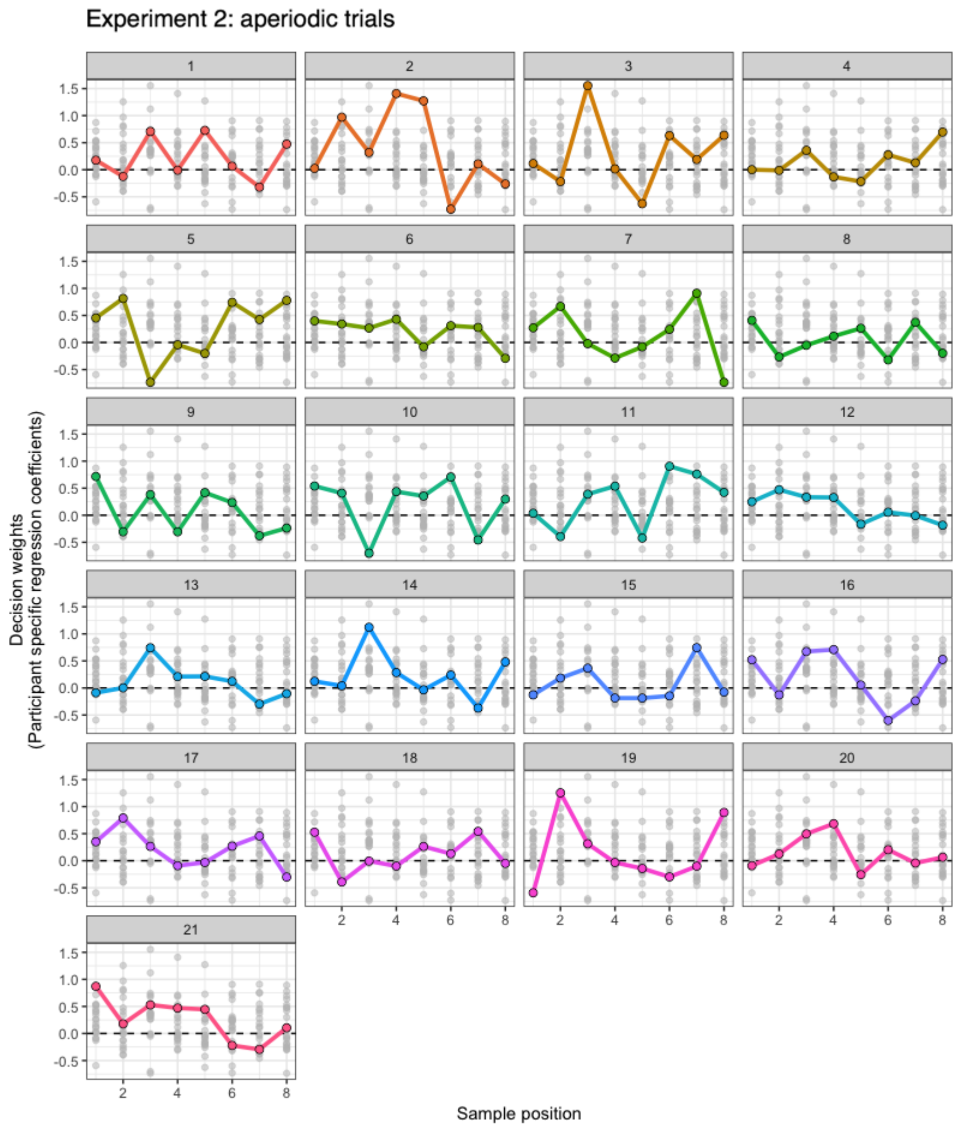


Fig. A.4 Expt. 2. Aperiodic trials: same structure as in figure A.3.

Appendix B

Experiment 6

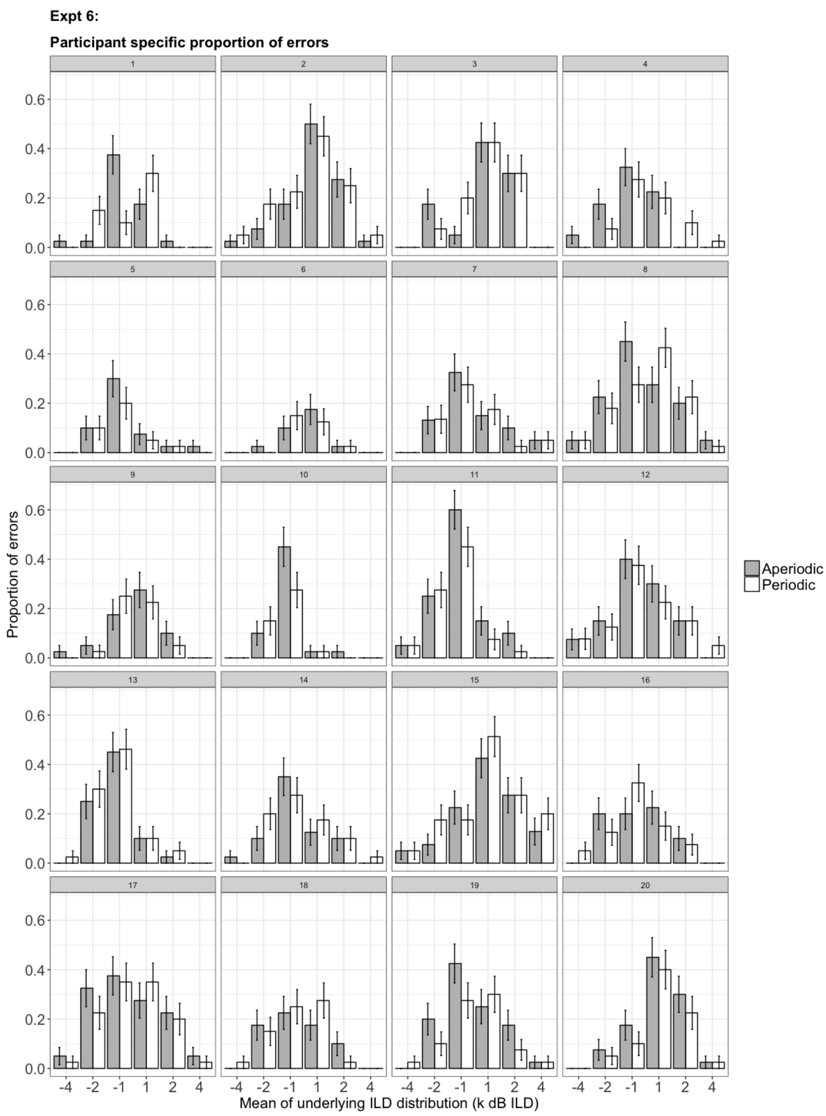


Fig. B.1 Expt. 6. Mean proportion of errors at each of the seven ILD distribution locations relative to mid-point for periodic and aperiodic stimuli. Panels: individual participants. Error bars: standard error of the mean.

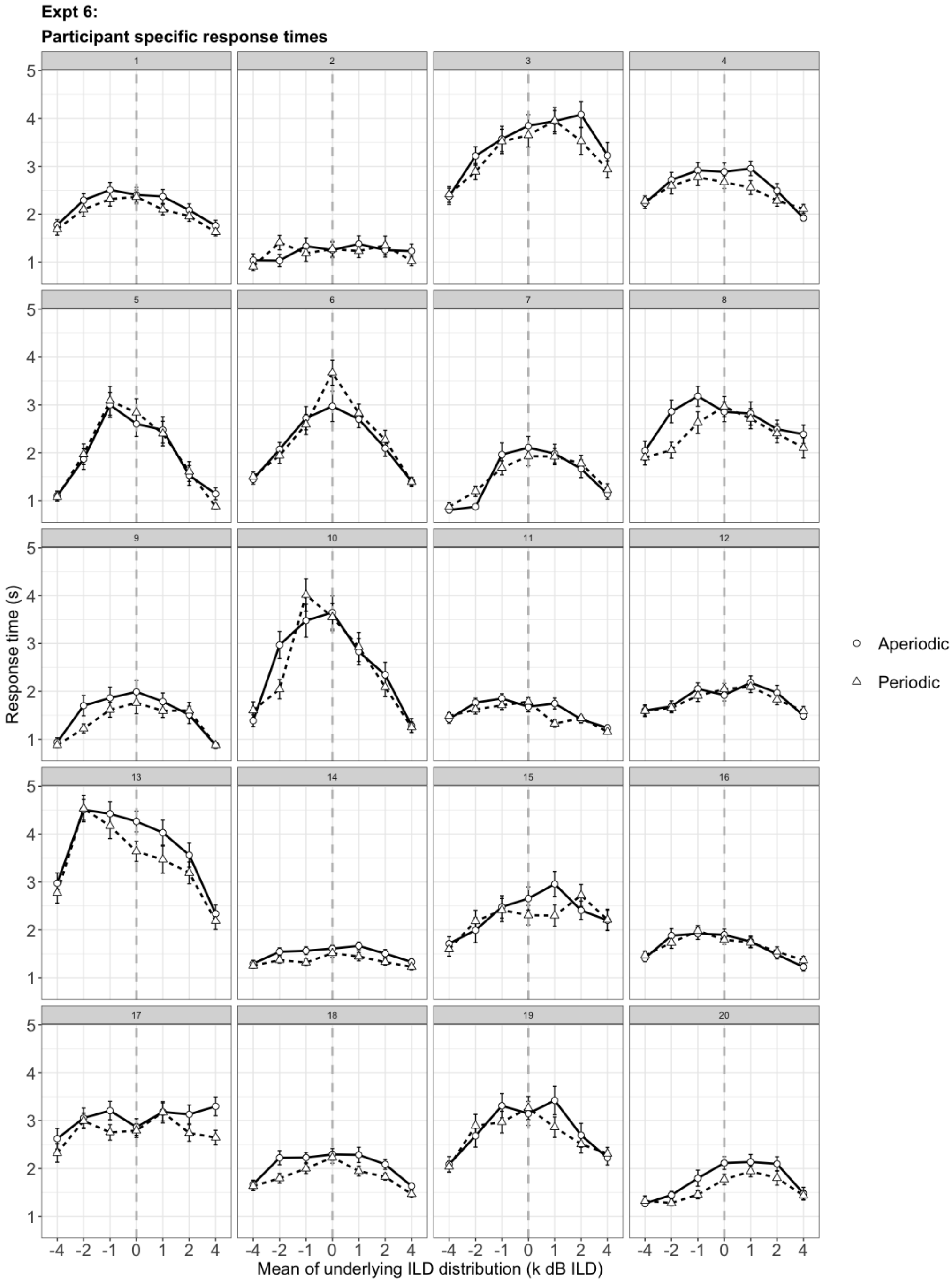


Fig. B.2 Expt. 6. Response time data at each of the seven ILD distribution locations relative to mid-point for periodic and aperiodic stimuli. Panels: individual participants. Error bars: standard error of the mean.

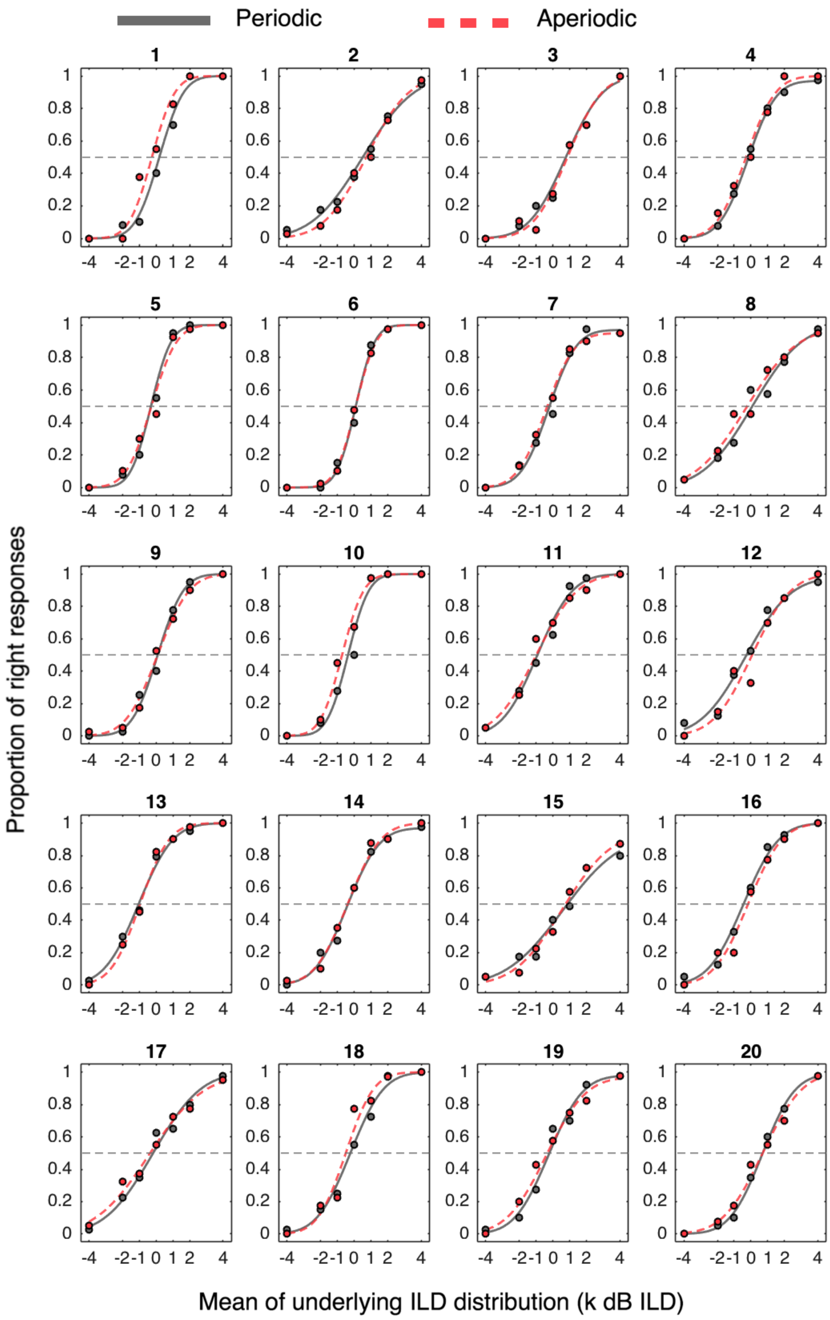


Fig. B.3 Expt. 6. Fitted psychometric curves for each participant showing proportion of Right responses relative to the mean of the underlying ILD distribution. *k* refers to each participant's decision threshold value that was estimated during the calibration session. Grey solid curves: periodic stimuli. Red dashed curves: aperiodic stimuli.

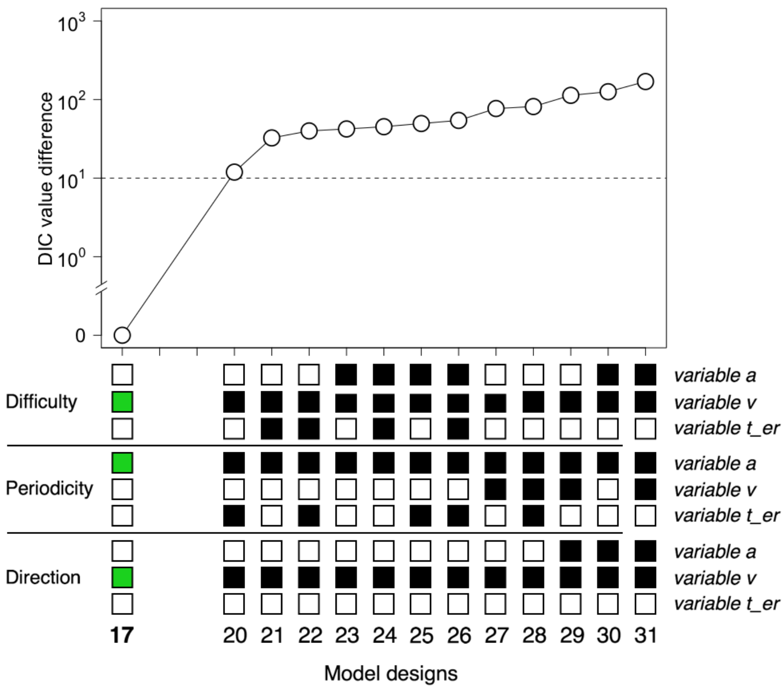


Fig. B.4 Expt. 6. The deviance information criterion (DIC) value differences between 12 additional variants of the drift-diffusion model compared with the base Model 17. The horizontal dashed line represents a significance cut-off point for difference in DIC value. Filled squares indicate that the particular parameter (labelled at the right) varied by the corresponding experimental condition (labelled at the left). Empty squares indicate that the parameter value was constant in the corresponding experimental condition. Green squares indicate the best fitting model (Model 17, DIC = 22066.22).

P ID	Threshold	Slope	Threshold Std Error	Slope Std Error	Model Deviance	GoF p-value	Periodicity
1	0.1951	0.876	0.1303	0.1192	7.8114	0.07	Periodic
1	-0.2465	0.9233	0.1194	0.1242	8.2592	0.0975	Aperiodic
2	0.5181	0.415	0.2538	0.0592	1.8527	0.8675	Periodic
2	0.7249	0.4997	0.218	0.068	2.4134	0.75	Aperiodic
3	0.7752	0.5748	0.198	0.0788	5.6921	0.28	Periodic
3	0.8692	0.6157	0.1664	0.0857	9.2078	0.0825	Aperiodic
4	-0.1875	0.7515	0.1721	0.112	0.5646	0.95	Periodic
4	-0.2896	0.7158	0.1472	0.0942	6.9509	0.145	Aperiodic
5	-0.3049	1.0474	0.1194	0.1396	3.8509	0.26	Periodic
5	-0.277	0.839	0.1369	0.1135	5.6027	0.2025	Aperiodic
6	0.0992	1.1115	0.1176	0.1684	2.3456	0.5525	Periodic
6	0.0995	1.0348	0.1198	0.1398	0.5684	0.94	Aperiodic
7	-0.2531	0.765	0.1738	0.1286	5.1887	0.2175	Periodic
7	-0.4278	0.7423	0.1687	0.1277	1.2279	0.8525	Aperiodic
8	0.0799	0.429	0.2524	0.0607	4.2368	0.5375	Periodic
8	-0.2812	0.4203	0.2487	0.0588	2.5164	0.7525	Aperiodic
9	0.0975	0.8269	0.128	0.0966	2.295	0.585	Periodic
9	0.0952	0.6703	0.1608	0.0803	4.5753	0.3275	Aperiodic
10	-0.3623	1.0175	0.1147	0.1493	7.2974	0.0625	Periodic
10	-0.717	0.9797	0.1202	0.1458	2.867	0.4575	Aperiodic
11	-0.9261	0.603	0.1763	0.0764	3.2968	0.63	Periodic
11	-0.9921	0.5152	0.1919	0.0627	3.5543	0.5825	Aperiodic
12	-0.3155	0.4734	0.23	0.0618	3.8757	0.5575	Periodic
12	0.0724	0.5467	0.1863	0.0867	9.2957	0.075	Aperiodic
13	-1.0388	0.6184	0.1571	0.0777	1.753	0.8325	Periodic
13	-0.9532	0.7319	0.1636	0.0959	3.0094	0.62	Aperiodic
14	-0.4218	0.6584	0.1862	0.0979	2.7821	0.635	Periodic
14	-0.3834	0.6421	0.153	0.0808	3.6085	0.53	Aperiodic
15	0.5827	0.3757	0.3109	0.0497	3.235	0.6825	Periodic
15	0.5774	0.4421	0.2799	0.0618	3.0155	0.69	Aperiodic
16	-0.4403	0.6025	0.1635	0.0715	3.7424	0.555	Periodic
16	-0.17	0.625	0.1626	0.073	5.1422	0.2825	Aperiodic
17	-0.1742	0.4434	0.2291	0.0636	2.7489	0.715	Periodic
17	-0.4169	0.3968	0.2813	0.0635	1.9815	0.775	Aperiodic
18	-0.2211	0.6358	0.155	0.075	4.9003	0.3575	Periodic
18	-0.4962	0.758	0.1462	0.102	7.9217	0.07	Aperiodic
19	-0.2505	0.622	0.1744	0.0933	4.2751	0.3525	Periodic
19	-0.4005	0.5356	0.2152	0.0711	3.8251	0.485	Aperiodic
20	0.6848	0.652	0.1741	0.0922	0.9672	0.895	Periodic
20	0.7003	0.5331	0.1943	0.07	2.2738	0.7625	Aperiodic

Table B.1 Expt. 6. Fitted thresholds and slopes values, estimated standard errors, deviance and goodness of fit parameters for fitted psychometric functions associated with each participant and periodicity condition.

Periodicity	Difficulty	Direction	PE Expt	PE M 17	RT Expt	RT M 17
Aperiodic	High	Left	0.309	0.260	2.584	2.515
Aperiodic	High	Right	0.239	0.226	2.558	2.414
Aperiodic	Medium	Left	0.143	0.110	2.247	2.166
Aperiodic	Medium	Right	0.132	0.106	2.218	2.095
Periodic	High	Left	0.264	0.267	2.435	2.385
Periodic	High	Right	0.248	0.235	2.354	2.300
Periodic	Medium	Left	0.141	0.117	2.117	2.059
Periodic	Medium	Right	0.111	0.114	2.114	2.017

Table B.2 Expt. 6. Simulated mean posterior predictions using the best fitting Model 17 (M 17) compared with experimental data (Expt) for both proportion of errors (PE) and response times (RT) across the periodicity, difficulty and direction conditions.

Appendix C

Experiment 7

Expt. 7. Hierarchical logistic regression analysis

The starting model contained fixed effects of Difference score [-2, -1, 0, 1, 2], Periodicity [Periodic, Aperiodic] and Location [Left, Right]. Four control covariates were also included to account for trial-by-trial carry over effects. The first two covariates, entitled "Recency of item 1" and "Recency of item 2", recorded how many trials earlier in the experiment the corresponding musical item of the pair was last used. The second two covariates, "Repeat of item 1" and "Repeat of item 2", recorded how many times the corresponding musical item had featured in a previous stimulus pair earlier in the experiment. These factors helped to satisfy the independence assumption of the linear model and to reduce the residual error of the model fit (Baayen and Milin, 2015).

Four variants of the starting model with different random effects were compared using a LRT to determine the best fitting random effects. Model 1 contained random intercepts for Participant ID. Model 2 contained random intercepts for both Participant ID and Experimental Block. Model 3 contained random slopes by the Difference score and random intercepts for Participant ID. Model 4 contained random slopes by the Difference score and random intercepts for both Participant ID and Experimental Block. Due to the relatively small number of trials, more complicated design matrices were excluded from the comparison due to lack of model convergence. The test showed that Model 3 (L.R: $X^2 = 65.77$, df $= 1$, p $= < 0.001$, BIC = 4010, AIC = 3913) had the best fit and lowest AIC and BIC estimations compared with the other models (Models 1,2 4: BIC = [4058, 4067, 4018], AIC = [3975, 3977, 3915]).

The backwards stepwise selection procedure was then performed on all fixed effects and interactions of Model 3. Single term deletions (achieved using the dropterm function in the "MASS" package in R (Venables and Ripley, 2002)) showed that neither the three-way interaction between Difference score, Periodicity and Location factors, nor the covariate Recency of item 1 were significant predictors in the model. Both were removed to create Model 3.1 which decreased model complexity and increased model fit (change in BIC = -12.9, change in AIC = -1). The selection process was repeated using Model 3.1 which showed that none of the two-way interactions between the Difference score, Periodicity and Location factors, nor the Recency of item 2, were significant. These were removed to create Model 3.2 which again increased model fit (change in BIC = -28.8, change in AIC = -3.1). Finally, factors Periodicity and Location were removed from Model 3.2 to create Model 3.3. This reduced model complexity and increased fit (change in BIC = -16.1, change in AIC = -3.2). All remaining factors significantly contributed to Model 3.3. The best fitting model therefore contained fixed effects of Difference score, Repeat of item 1 and Repeat of item 2, with a random slope for Difference score and a random intercept for Participant ID.

Expt. 7. Linear mixed-effects regression analysis

The starting model contained the same fixed effects and interactions as described in appendix C "Expt. 7. Hierarchical logistic regression analysis" with the addition of Trial ID to capture the autocorrelation of RTs between trials. Next, the same four combinations of random effects described in appendix C "Expt. 7. Hierarchical logistic regression analysis" were compared using a LRT. This showed that the second model (L.R: $X^2 = 31.05$, df = 1, $p < 0.001$, BIC = 5102, AIC = 4922) with a random intercept for both Participant ID and Experimental Block had the best fit and lowest AIC and BIC estimations compared with the other model variants (Models 1, 3, 4: BIC = [5125, 5171, 5147], AIC = [4951, 4928, 4977]). A backwards stepwise selection procedure was then performed on all fixed effects and interactions of Model 2. Firstly, the three-way interaction between the Difference score, Periodicity and Location factors, as well as all control covariates apart from Trial ID were classified as non-significant. They were removed from the model to create Model 2.1 which decreased model complexity and increased model fit (change in BIC = -61.5; change in AIC = -10.1). The selection process was repeated using Model 2.1 which showed that none of the two-way interactions between the Difference score, Periodicity and Location factors were significant. All three were removed to create a better fitting Model 2.2 (change in BIC = -72.7; change in AIC = -14.8). Finally, the Periodicity factor was removed from

Model 2.2 to create Model 2.3. This reduced model complexity and increased fit (change in BIC = -8.2, change in AIC = -1.8). The remaining fixed effect factors of Difference score, Location and Trial ID significantly contributed to Model 2.3.

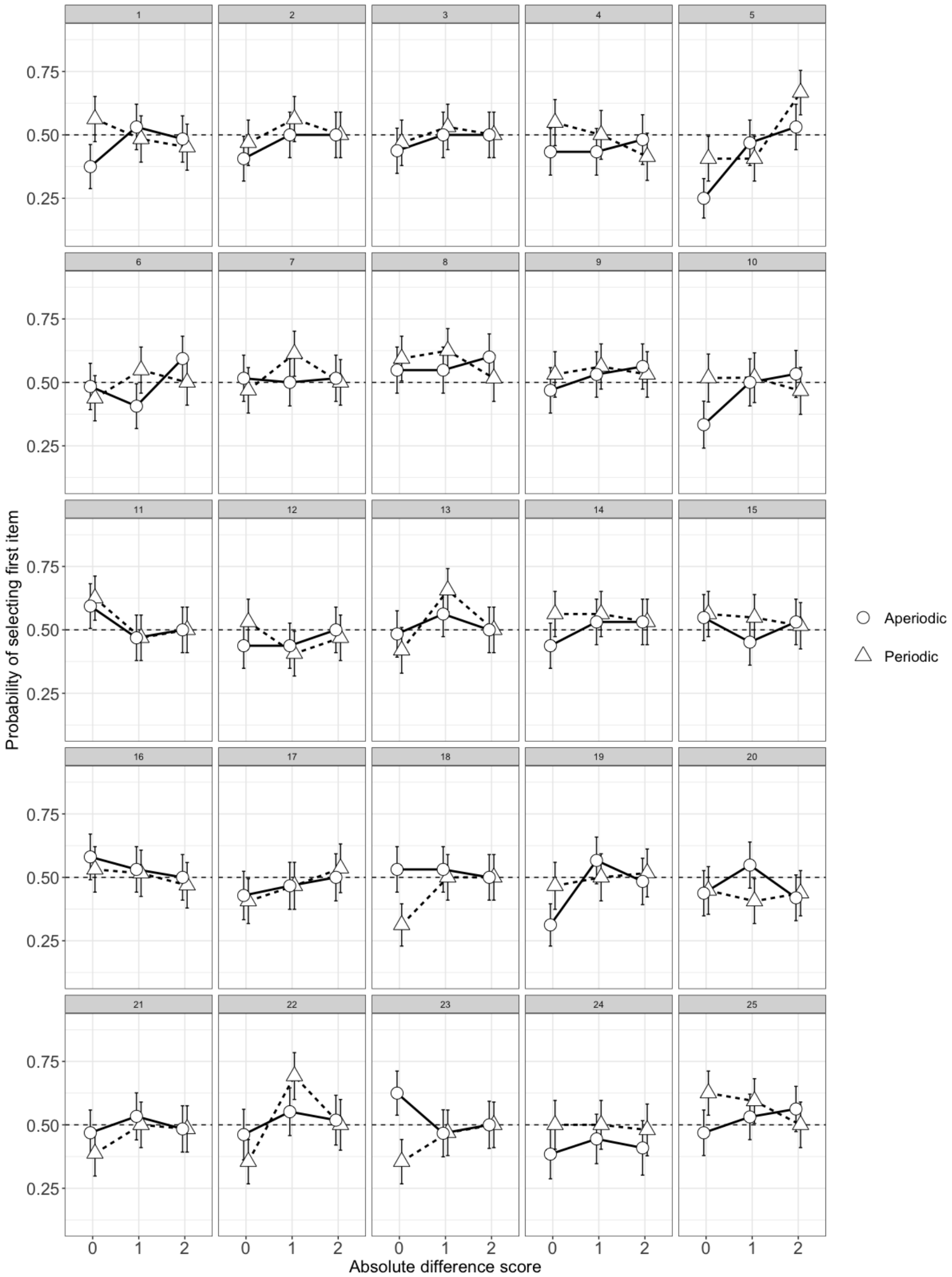


Fig. C.1 Expt. 7. Probability of participants selecting the first musical item by the absolute difference score and periodicity conditions. Panels: individual participants. Error bars: standard error of the mean.

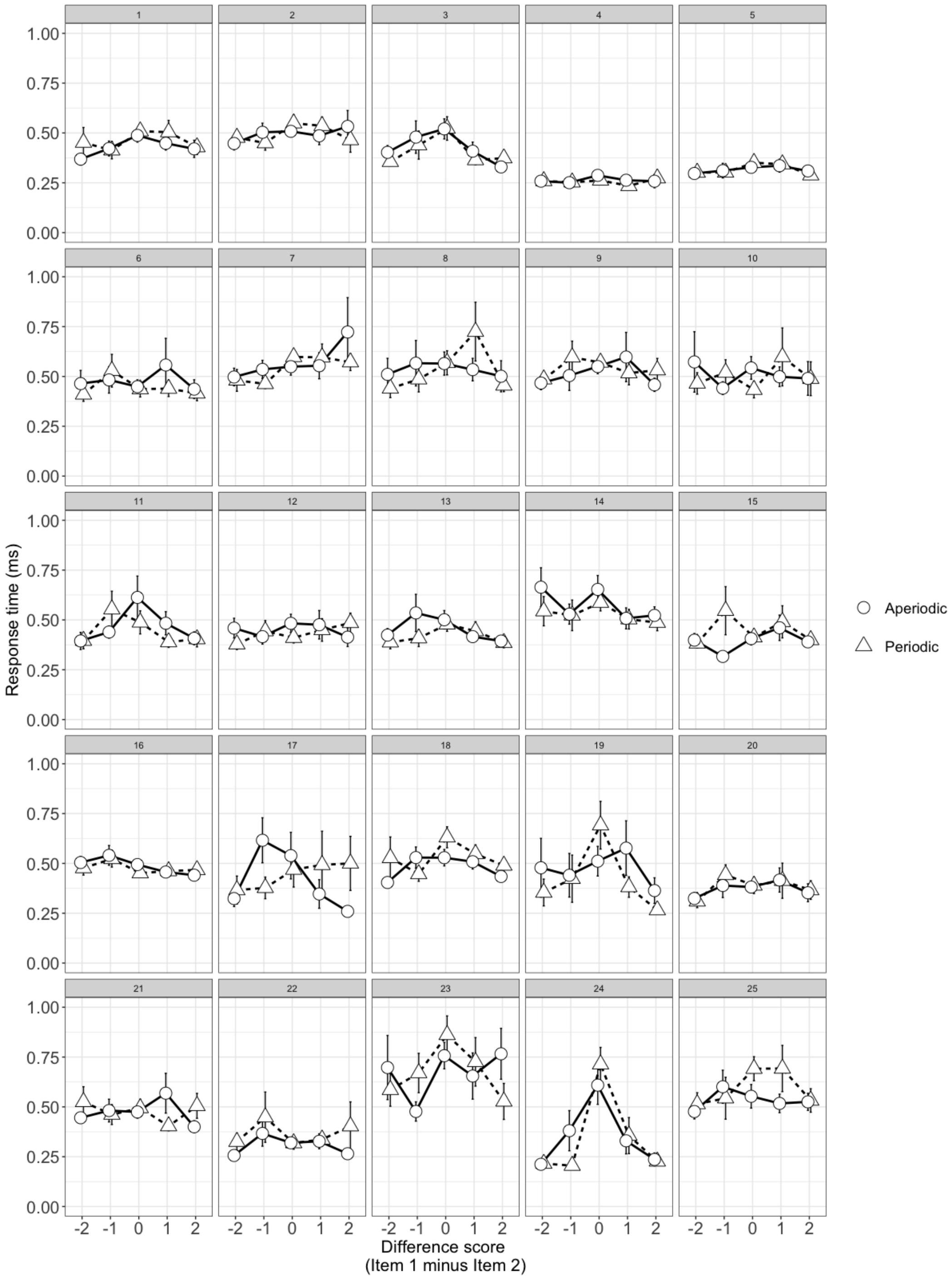


Fig. C.2 Expt. 7. Response times by the difference score and periodicity conditions. Panels: individual participants. Error bars: standard error of the mean.

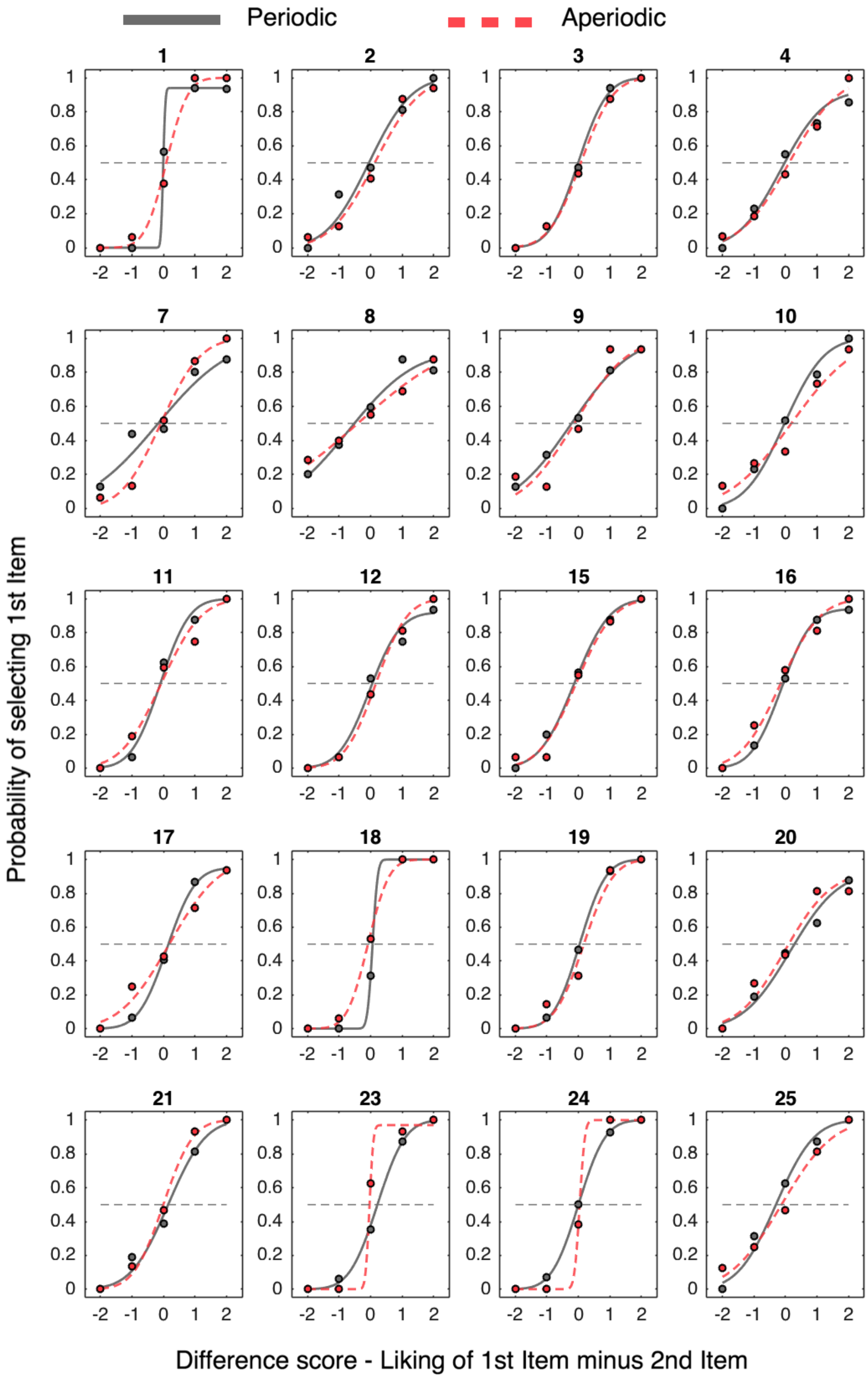


Fig. C.3 Expt. 7. Participant-specific psychometric curves showing the probability of selecting the first musical item by periodicity condition. Grey solid curves: periodic stimuli. Red dashed curves: aperiodic stimuli.

Musical tracks	
Stevie Wonder: Superstitious	Miles Davis: All blues
The Beach Boys: God only knows	Jacqueline Du Pré: Elgar cello concerto
John Williams: Jurassic Park main theme	John Lennon: Imagine
Beastie Boys: Intergalactic	Out Kast: Hey Ya
Whitney Houston: I will always love you	Led Zeppelin: Stairway to heaven
Paul McCartney: Hey Jude	Glen Gould: The Goldberg variations
Donna Summer: I feel love	Simon and Garfunkel: Bridge over troubled water
Jackson 5: ABC	Amy Winehouse: Rehab
U2: Beautiful day	Coldplay: Viva La Vida
Duke Ellington: Take the A train	Louis Armstrong: What a wonderful world
Beyonce: Put a ring on it	Lang Lang: Chopin
Christina Aguilera: Beautiful	Bob Dylan: Mr. Tambourine man
Guns n roses: Sweet child of mine	Will Smith: Getting jiggy wit it
Dave Brubeck: Take 5	Jimi Hendrix: Purple Haze
Nigel Kennedy: Beethoven violin concerto	Madonna: Like a virgin
ABBA: Dancing queen	Eagles: Hotel California
Queen: Bohemian Rhapsody	Michael Jackson: Thriller
Brad Mehldau Trio: Day is done	Nirvana: Smells like teen spirit
Jeff Buckley: Hallelujah	Aretha Franklin: Respect
Bob Marley: No woman no cry	The Police: Every breath you take
Keith Jarrett: Someone to watch over me	Britney Spears: Baby one more time
Debussy: Clair de lune	Red Hot Chili Peppers: Californication
Marvin Gaye: I heard it through the grapevine	Brian Adams: Everything I do
John Coltrane: Giant steps	Elvis Presley: Hound dog
Elgar: Nimrod	Ray Charles: Hit the road Jack
Jodi Mitchell: Both sides now	Run DMC: Walk this way
Katie Perry: I kissed a girl	Trinity choir: Hark the herald angels sing
Rolling Stones: Satisfaction	

Table C.1 Expt. 7. A list of the 55 culturally familiar musical tracks that were used to create the experimental stimuli.