Anxiety and computations of uncertainty during reward-based learning in volatile environments

Thomas Peter Hein

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Goldsmiths College, University of London

November 2021

Declaration of Authorship I Thomas Hein hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed: Thomas Hein **Date**: 22/11/2021

Acknowledgements

First, I wholeheartedly thank my primary supervisor, Dr María Herrojo Ruiz, for being consistently supportive and generous with her time. I could not have asked for a more ideal supervisor; her accommodating and patient help, diligence, optimism, encouragement, and incisive wit somehow pulled me through. Maria helped stimulate and develop the ideas and experiments in this thesis while providing sufficient freedom to learn and stumble on independently. Maria is an extraordinary scientist, and her input to this PhD (and my gratitude for this unforgettable and at times even enjoyable experience) is beyond measure. It has been a pleasure and a privilege to learn from and work with you.

Professor Jan de Fockert deserves special thanks for his supervisory support. Jan helped me to develop a sense of experimental psychology, guiding my research and devoting considerable time to enriching my interest in science and the psychology of attention and working memory.

I would also like to thank Dr Lilian Weber, who provided us with code and truly substantial and invaluable discussions, insights, and improvements. Lilian is an exceptional scientist whose compassionate and encouraging words inspired me at each moment throughout this journey.

All my friends deserve thanks for their attempts to just let off steam by coaxing me to social events I would otherwise weasel out of—especially Alex P., George and Jack, B., Luke S., Woody, Simon V, Dave L, and Chris E.

And thanks to my family. My Mum, who, in unwavering kindness, always supported me despite the unfavourable odds. My Dad, whose discussions on psychology and unparalleled self-confessed proclivity to remember negative experiences inspired my work. And my sister Stefanie and her husband Gerald, who kept me smiling and provided countless hours of strategic board game downtime, even if Edith and Henry were troubled by the uproar.

And finally, but by no means least, to my wife and best friend 王郁文 and her sweet and caring parents (王世明 and 梅麗娟), whose love and generosity and support have kept me going over these challenging years. Without you all, none of this could have happened.

Abstract

Uncertainty plays a core mechanistic role in computational psychiatric accounts of clinical disorders. This thesis tests the hypothesis that subclinical anxiety interferes with uncertainty processing and reward learning in unstable environments. In a series of four experiments using a combination of computational modelling and reward-based learning tasks with electrophysiological and neuromagnetic data, we detail the relationships between anxiety, decision making, and uncertainty (inverse, precision) estimates. Using EEG and a hierarchical Bayesian filtering model of behaviour, the first experiment shows how experiencing a temporary state of anxiety can bias uncertainty estimates essential for optimal belief updates in Bayesian theories of learning and impede overall reward learning performance. We also tracked the expression of belief update signals in trial-by-trial EEG amplitudes, revealing precision-weighted prediction errors (pwPE) about stimulus outcomes were represented in the control group only. In the second study, we investigated how the behavioural and modelling findings from the first experiment are associated with the oscillatory representations of pwPEs and predictions hypothesised in predictive coding and Bayesian inference frameworks. Using convolution modelling for EEG oscillatory responses, we reveal the influence of biased uncertainty estimates in state anxiety on the oscillations encoding predictions and pwPEs during reward learning. For the third study, we asked how motivation to reduce anxiety may improve reward-based learning, detailing null results due to unsuccessfully inducing a state of anxiety. Using MEG, the final experimental chapter examined the effects of high levels of trait anxiety on reward learning and the spectral signatures of predictions and pwPEs. We show how high trait anxiety amplifies volatility estimates and increases uncertainty, impairing reward learning and altering the oscillatory responses encoding pwPEs. Together, the results from this thesis suggest that subclinical anxiety impedes reward-based learning by biasing uncertainty estimates and altering the neural encoding of pwPEs and predictions considered essential for optimal belief updating.

Table of Contents

Chapter 1: General Introduction

1.1 Introduction

Uncertainty is immanent in nature. To prevail, organisms develop optimally with their environments—finessing and capitalising on their place within it. The likelihood of survival and reproductive fitness is thus ultimately bound to the responses or decisions each organism makes. Simpler environments impose simpler demands upon their denizens. For some, it is irrelevant and surplus to requirements to learn how to make complex decisions optimally and flexibly. Brains are complicated organs; they evolved to process and solve sophisticated challenges from a unique panoply of intricate and changing environments. [1] As learning agents par excellence, humans have consequently expanded to outperform competitors in choosing actions, reaping the reward of prospering in most of the world's environments. Remarkably, we achieve this using a three pound spongy organ consisting of ~10^{11} neurons, each with around ~10^{3} synaptic connections consuming only ~20 watts of energy—equivalent to that of a regular light bulb (Dalgleish et al., 2020; Merolla et al., 2014; Tsien, 2015).

This truly astonishing organ unlocks actions that are both adapted and adaptive to the challenges of each environmental system we operate in (Friston, 2013; Friston et al., 2014). To function effectively in complex dynamic systems, humans are thought to make predictions based on an internal model of their world, formed primarily through learned experience (Conant & Ross Ashby, 1970; O'Reilly, 2013). The most efficient model would perfectly reconstruct the mapping between sensory input and purposeful actions, affording rapid and reliable predictions about the world and the agents acting within it. To quote Norbert Reiner (1945, p. 320), "...the best material model for a cat is another, or preferably the same cat". Our brains—examples of approximately efficient regulators that prevail across some time—are thought to be engaged in the process of constructing and refining a successful generative model; one that converges on a useful mapping between uncertain sensory input and purposeful actions. This thesis pertains to those processes used by the human brain when attempting to learn from and compute uncertainty in a dynamic environment. However, rather

1 Although there are always exceptions. Surprisingly, a unicellular amoeboid slime mould organism *Physarum polycephalum* has demonstrated its ability, despite being brainless, to solve the multi-arm bandit problem (Reid et al., 2016). In that study, *Physarum polycephalum* compares multiple options and combines data on a reward to make adaptive decisions.

than attempt to address the vast complexity of the brain and plunge into the depths of all human behaviour, the ambitions here are constrained to understanding human learning through the decisions we make, building on years of research where outcomes are used to elicit actions to either obtain a reward or avoid punishments (Thorndike, 1927). We define reward as an outcome that reinforces behaviour; or, as outcomes described by the healthy agent as pleasant or appetitive or hedonic (Moutoussis et al., 2015). In this thesis, we use computational models of learning behaviour (precise mathematical descriptions) to better understand behavioural data during reward learning in an uncertain environment. More specifically, we investigate how healthy individuals might suboptimally represent the current state of the environment by misestimating uncertainty about the probability of rewarded outcomes and their change over time (volatility uncertainty).

Throughout, we draw upon theoretical and empirical efforts in psychology dating back to Helmholtz (1821-94) in understanding human learning and perception (Helmholtz, 1866). The important scion of Helmholtz in modern neuroscience is the notion that the brain seizes upon, refines, and embodies the causal arrangement of the world through the apparatus of statistical inference (Dayan et al., 1995; Hinton & Zemel, 1994). The work contained in this thesis subscribes to the now popular explanation for this process—how messages are encoded and passed in the brain—using a Bayesian predictive coding (PC) process (Aitchison & Lengyel, 2017; Knill & Pouget, 2004; Mumford, 1992; Rao & Ballard, 1999; Srinivasan et al., 1982). Going further, here we examine the interaction between everyday experiences with anxiety and the efficient use of the computational toolkit of hierarchical Bayesian learning. In this way, we connect learning, decision-making, and affective states like anxiety, revealing at a computational level how these interactions can bias information processing and belief formation—as beliefs are essentially (potentially incorrect) representations of the world (Yon et al., 2019). [2] This thesis seeks to capitalise on the benefits of using a computational methodology to study brain and behaviour, leveraging understanding into the central difficulties that subserve behavioural change and mental well-being. The work included here is focused on providing insights into how human anxiety impacts decision making and behaviour change, using a combination of computational modelling and reward-based tasks with electrophysiological and neuromagnetic data. Our hope is that these new measures may also be applied to the study of formal psychiatric illnesses, such as anxiety and major depression, and subsequently inform treatment.

[2] Here and hereafter we mean beliefs in a technical Bayesian sense (viz. probability distributions encoded by neuronal activity/connectivity).

We start with necessary definitions of the interlinked processes detailed above (including decision making and the meaning of value and reward for humans). After, we provide a concise account of the computational operations and current neurobiology that subserves these processes. In doing so, we bring specific attention to the important distinctions between reinforcement learning (RL) and the Bayesian approach, using the uncertainty of complex and changing learning environments to illustrate where each has supplied new information to our knowledge of the brain. Finally, before moving on to new research findings concerning state anxiety (**Chapters 2–3**), motivation and state anxiety (**Chapter 4**), and subclinical trait anxiety (**Chapter 5**), we discuss how these everyday experiences, like with emotion, can influence how the brain computes uncertainty and learns from new information in a Bayesian framework.

1.1.1 Anxiety, learning and deciding.

The arrangement of the cognitive processes of interest in this thesis is in the interactions between everyday experiences with anxiety, learning, and the decisions that we take. In line with previous work in anxiety, we define cognition as information processing: the processing of sensory input and the process of assimilating that information to raise overall adaptive strength and reproductive progress (Robinson et al., 2013). The relationship between these three interacting components (anxiety, learning, deciding) is, however, far from trivial, and will to some extent be oversimplified here. A scheme closer to reality would require parallel multi-directional interactions fluctuating across time. In the thesis that follows, the relationship between temporary anxious experiences and learning is tested in **Chapter 2**, which investigates the impact of inducing a state of anxiety in a subclinical population on model-free and model-based learning and brain data using electroencephalography (EEG). In **Chapter 3** we expand upon the findings from **Chapter 2** and ask whether state anxiety alters the neural oscillatory correlates of predicting and learning from rewards. Later in **Chapter 4,** we extend our focus to the potential adaptive motivational component of state anxiety. While in **Chapter 5** we aimed to build on the results from **Chapters 2**–**3** by testing the effect of high levels of (subclinical) trait anxiety on reward learning in a volatile environment using magnetoencephalography (MEG). We approach each chapter from a computational perspective, modelling how the brain may be processing information to improve behaviour.

To measure decision making, we define decisions as taking an action. There exist at least four interacting decision-making systems supported by numerous neural mechanisms that transmit information to ensuing actions (Cisek & Kalaska, 2010; Montague, 2006; Wunderlich et al., 2009). And this only proceeds once specifying what a value is and identifying what we value, unpacked briefly in **Section 1.2.1** (Glimcher & Fehr, 2013; Hull, 1943; Schultz et al., 1997).

While the precise subtleties and variety of these decision making systems are still a matter of debate, typically agreed upon action-selection systems include genetically embedded reflexes and Pavlovian responses (stereotyped and valence dependent anticipatory responses relevant for survival), goal-directed or model-based behaviour, and a procedural action-selection system (Redish, 2013; Van Der Meer et al., 2012). This thesis focuses only on the model-based action-outcome decision system. For clarity, goal-directed or model-based behaviour is a term synonymous with 'value-based' or 'reward-guided' actions, all of which are included in a deliberative decision-making system (Redish, 2013). There are, moreover, numerous learning mechanisms that both work independently and harmonise in the brain, supplying input to decisional systems that develop in tandem under multiple sets of criteria (Cisek & Kalaska, 2010; Clark, 2015). Yet in spite of evidence documenting how dynamic emotional states shape decision-making (Winkielman et al., 2007), our understanding of the diverse connections between emotional and computational learning processes remains unclear (Seth & Friston, 2016).

Understanding multiple and distinct systems in the human brain for both learning and decision-making leads to the question of what function this serves? It is evidently fundamental for organisms to learn to adapt to the changing needs of whichever environment they exist in; even bacteria, fungi, and plants exhibit some ability to use past experience to inform behaviour (Casadesús & D'Ari, 2002; Gagliano et al., 2016; Hilker et al., 2016). And fascinatingly, amoeboid slime moulds can solve simple learning and decision-making tasks (Reid et al., 2016). Propositions of multiple decision-making systems in humans stem back to at least the days of Freud (1961) in psychology and enjoys growing purchase within distributed computing (Minsky, 1988) and behavioural economics (Kahneman, 2011). Further, from neural networks, we can see that a distinction between multiple decision-making systems and learning mechanisms is not necessarily necessary. Neural networks consolidate learning and decision making systems by using a direct mapping from inputs to actions (Krizhevsky et al., 2012; LeCun et al., 2015; Mnih et al., 2015).

Clearly, at the final moment, there is one being who mandates the action. But it is useful to understand that being with regard to the multiple subsystems that comprise it. Thus, for humans, there exist multiple and distinct learning mechanisms that promote specific options for actions to multiple and distinct decision-making systems (Redish, 2013). Based on unique sources of information we can attempt to select the optimal action to obtain a specific outcome

(Daw et al., 2005). In the case of reflexes, this is hardwired but potentially overridden. [3] In the case of Pavlovian responses, learning can occur, calibrating selected actions to predicted cue-outcome pairings (Yerkes & Morgulis, 1909). But model-based and procedural systems are vastly more flexible. They combine together unique forms of learning information. For the deliberative model-based decision-making system, this affords the potential for associating disparate information to arrive at new decisions—such as the combination of my knowledge that my wife loves to eat peaches and that food is served at cinemas.

The above sentence should elicit a strong violation of expectation—as peaches are not typically linked to the foods on offer at cinemas. This serves both as an example for the possibility of incorrect beliefs and for the global currency for learning found in measuring the difference between what you observed and what you expected. As mentioned above, a model-based system can include incorrect beliefs; there is consequently a certain vulnerability to probabilistic learning. Uncertainty is baked into complex environments and can bias beliefs and responses. We embody a generative model of our environment (a formal description of why something happens), as an optimal generative model should include all we need to know to generate the sensory information we receive. But if we misestimate information as it is inherently uncertain, our model will be unreliable. Humans can estimate the uncertainty of our beliefs about how some state of the world functions without those beliefs necessarily corresponding to the true hidden state we are inferring upon. This leaves space for modulation by other systems in the brain (each decision-making system also relies on other processes including motor control, perception, situation recognition, and motivation). Of which, there is still much debate concerning the relationships between them, such as among the Pavlovian action-selection and motivational (Corbit & Balleine, 2005; Corbit et al., 2001; Wassum, Cely, et al., 2011; Wassum, Ostlund, et al., 2011) and emotional systems (Barrett, 2006; LeDoux, 2012).

Our particular focus here is the everyday experiences we pass through like with affective or emotional states (such as anxiety) and motivation (Grupe & Nitschke, 2013; Gutiérrez-García & Contreras, 2013). As the above cinema example illustrates, I am emotionally compelled to surprise and delight my wife with peaches—which perhaps motivates my learning of all the possible contingencies that might lead to maximise her momentary happiness. However, as we will see later, these preferences and outcome contingencies may change across time, leading to wholly different outcomes (**Section 1.4**). Moreover, part of the experience of

[3] One can, with relative ease, maintain grip on a painfully hot cup of tea—overriding the reflex to drop a hot and painful object—in favour of landing the cup somewhere safe first.

momentary happiness is the violation of expectation, thus habituation might demand predicting inconstant preferences (Rutledge et al., 2014). These changes to outcomes require dynamic learning. But do everyday experiences with emotion and motivation influence these learning mechanisms? And to what extent does their effect shape our decisions? These questions concern probabilistic computations in the brain and are ideally tested using computational approaches and conceptually using Bayesian inference. But despite initial work into related fields of 'higher level' cognitive abilities (including planning, cognitive control, social cognition, and language, see Diaconescu et al., 2014; Diaconescu, Mathys, et al., 2017; FitzGerald et al., 2014; Friston et al., 2013, 2015; Harrison et al., 2011; Hohwy, 2013; Moutoussis et al., 2014) sharpening our understanding of those processes, the image concerning the interactions between emotions and cognition remains relatively low resolution (Clark, 2015; Friston et al., 2012; Oudeyer & Kaplan, 2009; Paulus & Yu, 2012; Rutledge et al., 2014; Seth & Friston, 2016).

1.1.2 Computational psychiatry

The work in this thesis is informed by the relatively new field of computational psychiatry. Computational psychiatry seeks to set out a theoretical framework bridging higher-level psychological states (e.g. depression) and neural circuits by mathematically modelling these processes (Huys et al., 2016; Montague et al., 2012; Nair et al., 2020). Computational psychiatry thus uses the tools best suited to describe complex nonlinear learning and decision-making mechanisms. In clinical disorders, these mechanisms can be an intricate combination of working, inoperative, and interacting parts (Corlett & Fletcher, 2014; Redish & Gordon, 2016). In this thesis, as in the field of computational psychiatry, computational means a broad application of sophisticated mathematical and theoretical tools to complex biological systems (Redish & Gordon, 2016). Computational psychiatry thus expands into the gaps between computational neuroscience and psychiatry to provide model-based assays utilised to analyse and identify mental disorders (Huys et al., 2011; Montague et al., 2012; Stephan & Mathys, 2014). It is this seizing upon both function and dysfunction, optimal and suboptimal, adaptive and maladaptive, that makes the computational approach also ideal for understanding the dynamic role of affective states and moods in shaping moment-by-moment learning and decision-making in healthy humans (Aylward et al., 2019; Clark et al., 2018). With the computational approach, we can also reach a more nuanced understanding of the underlying neural substrates involved in this process. Put simply, computational models provide a means to characterise both individuals and groups and to test competing models to best capture the differences between learning styles.

Given a good model fit to observed behaviour, a link can be made to differences in behaviour or neural responses (Corcoran & Cecchi, 2018; Montague et al., 2012; Redish & Gordon, 2016; Wiecki et al., 2015). Recent work has demonstrated just how successful this has been in characterising clinical populations (Parr, Rees, et al., 2018; Smith et al., 2021). There is an intriguing focus on the role of uncertainty (or its inverse, precision) in shaping learning dynamics, with previous studies documenting a whole suite of clinical expressions linked to disruptions in uncertainty estimates (Friston, 2017). These include complex visual hallucinations (e.g. Charles Bonnet syndrome) and issues with movements in Parkinson's disease being linked to increases in uncertainty (lower precision) estimates (Friston et al., 2013; Reichert et al., 2013), while decreases in uncertainty (higher precision) estimates are linked with schizophrenia and autism (Friston et al., 2016; Lawson et al., 2014). Later in **Section 1.4** we decipher and elaborate on the role of uncertainty in perception, learning, and action, detailing the relevance of these 'precision pathologies' in relation to anxiety.

The experimental research in this thesis expands upon prior empirical and theoretical work that demonstrates the importance of uncertainty estimates in optimal Bayesian learning by investigating how healthy individuals are also subject to deviations from optimal performance. If the computational psychiatry framework states that a psychiatric disorder implies suboptimality in decision making (Moutoussis et al., 2015), then we hypothesise in this thesis that this suboptimality is on a spectrum. In doing so, we link experiences (including anxious states and high levels of trait anxiety) in healthy individuals to biases and modifications to how we learn from complex and dynamic reward environments. By modelling behavioural responses using a Bayesian approach—where uncertainty is explicitly quantified—we can provide a model-based computational account of the learning mechanisms involved when decisions are sanctioned. We thus investigate how difficulties learning from incomplete information and misestimation of uncertainty are crucial to understanding how temporary affective states shape learning in healthy individuals—aligning with recent theoretical accounts concerning affective disorders (Pulcu & Browning, 2019; Williams, 2016).

The above sections covering the preliminaries of this work should now form a précis of the essential components examined in the thesis below: learning, decision-making, and healthy subclinical anxiety. Through the conceptual lens of Bayesian learning and the computational toolkit of modelling behavioural responses, we investigate how anxiety modulates learning and estimates of uncertainty, further associating these changes to neural activity using electrophysiological (EEG) and neuromagnetic (MEG) signals.

1.2 Specifying value and deciding how to act

1.2.1 Values

Before we unpack how we learn to adapt our actions through the decisions we make, we must first specify a working definition of value (Lehrer, 2010). Value is something that can be calculated dynamically; it can change with each new evaluation, measuring the amount we are willing to transact for a given outcome (how much we are willing to work, trade, or pay for reward or to bypass punishment). Pain, for instance, articulates a negative scenario demanding palliation for survival. Inanition demands nourishment, and so food is valued. The organism's value of survival can motivate what we find rewarding to maintain homeostasis. Typically this is communicated through internal bodily states like water, fat, sugar, and oxygen levels that generate reward-oriented behaviours (Cannon, 1932; Morville et al., 2018; Pezzulo et al., 2015). Of course, the mechanistic complexity of homeostasis (both operational and computational) extends beyond the mere negative feedback control to a hierarchy of physiological control structures, such as motivational and emotion systems (Berridge, 2004; Carpenter, 2004; Dickinson & Balleine, 1994; Morville et al., 2018; Pezzulo et al., 2015).

Part of the problem for understanding what people value, in order to study how we take actions to obtain value, is that we do not have an efficient means to measure value. In early work on animals, this was tested by measuring the degree of effort exerted to obtain some reward (see: Ahmed, 2010; Salamone, Pardo, et al., 2016). This could be the decision between two different options for food with asymmetric payoffs (one grape or twenty pieces of carrot) to find a threshold of value (Padoa-Schioppa & Assad, 2006). Alternatively, by varying the amount of some negative experience (like an electric shock) we can understand how much an animal will tolerate for reward (Padoa-Schioppa & Assad, 2006). These methods also work equally well on humans. However, economists (who are also concerned with how humans value things) prefer to simply ask participants their preferences between choices. The issue with this technique is that humans are typically inconsistent in their evaluations (Kahneman et al., 1982; Kahneman & Tversky, 2013; Tversky & Kahneman, 1979). [4] Decisions are thought of as the process of dynamically calculating value at each time point, which leads to inconsistencies in 'rational' human decision-making. This has now been widely documented, but perhaps most

[4] It is important to note that others have attempted to replicate Kahneman and Tversky's work, with several studies showing a lack of replicability, see https://replicationindex.com/2020/12/30/a-meta-scientific-perspective-on-thinking-fast-and-slow/. The framing effects reported in Kahneman (1981) have received mixed replications, see Druckman (2001).

famously (and what has, since originally writing this section, become acutely apt, as we battle against a global pandemic and misinformation) in Amos Tversky and Daniel Kahneman's (1981) disease example of how phrasing decisions in terms of wins and losses alter choices:

Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

Problem 1

• If Program A is adopted, 200 people will be saved.

• If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3
probability that no people will be saved.

Which of the two programs would you favor?

Problem 2

• If Program C is adopted 400 people will die.

• If Program D is adopted there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die.

Which of the two programs would you favor?

A significantly higher proportion of participants select options A and D, despite options A and C and options B and D being identical. The explanation offered by Kahneman and Tversky (1981) is one of our greater sensitivity to losses than gains. Consequently, humans are not as willing to risk losing something when compared to their devil-may-care approach to winning. This bias toward safety is exacerbated in anxiety, as we discuss later in **Section 1.5**. Intriguingly, despite being told about the nature of the decision involved in the disease example, and perhaps as these decisions are informed by emotions, the choices when represented to us persist in sounding like the right choice (Dawes et al., 2007).

Along with numerous other examples of the seemingly illogical calculations of value from the psychology and economics literature (for example 'extremeness aversion' in framing, see Ariely, 2008; Glimcher, 2004; Plous, 1993; Shapiro et al., 1998; Simonson & Tversky, 1992), Kahneman and Tversky also documented how humans commonly make decisions discordant with economic expectations (Kahneman et al., 1982). These include a host of social predilections like cooperation, fairness, equality, and egalitarian motives, but also critically

include emotional states (Bechara & Damasio, 2005; Dawes et al., 2007, 2012; Fehr & Schmidt, 1999; Gigerenzer & Goldstein, 1996; Kahneman, 2011; Kahneman et al., 1982, 1986; Lin et al., 2006; Nishi et al., 2015; Plous, 1993; Simon, 1955; Van Der Meer et al., 2012). To summarise this section on value, from the exceptional research over the past several decades, it appears humans do not really estimate the true value of things, and that value is dynamic and not a quality that can be derived directly from an object's nature. Rather, we approximate the current value of something on a moment-by-moment basis using *heuristics*, algorithms that (for the most part) serve us well (Gilovich et al., 2002).

Value-guided choice is therefore driven by information pertaining to the receipt of a specified target (Schultz et al., 1997). Money affords us a shared system for value—an agreed upon fiction in the language of economists (Simmel, 2004; Sims, 1980). However, as will be discussed further in **Chapter 4**, this can also refer to an internal state, like intrinsic motivation (Cameron & Pierce, 1994; Friston, Lin, et al., 2017; Gottlieb et al., 2013; Oudeyer & Kaplan, 2009). Connected to Pavlovian responses, as introduced above, a reward is something usually approached, while punishment is usually avoided. For these reasons, it is easy to conclude that what is valued or rewarding is pleasurable, and humans seek out pleasure and avoid punishment. But value and pleasure are not one and the same. A more accurate description is that we target things we identify as having high value (Rangel et al., 2008). The amount an organism is willing to exert in either effort to approach or work for an outcome is also not proportional to the euphoria (pleasure) experienced in the brain when receiving an outcome (Berridge, 1996). [5] This becomes important to distinguish when we operationalise rewards in psychology and investigate their neural underpinnings as reviewed in **Section 1.2.3** and **Section 1.3** below.

[5] Euphoria: from the Ancient Greek phoria *φορία* meaning 'to bear' or 'feeling' and the root meaning of *εὖ* meaning 'good' has been used as an alternative to 'pleasure' for its antonym in *dysphoria* (Redish et al., 2008; Redish, 2013). Euphoria is signalled by endorphins linked to μ-opioid receptors, whereas dysphoria is signalled by dynorphin associated with κ-opioid receptors (Land et al., 2008; Laurent et al., 2012; Le Merrer et al., 2009).

1.2.1 Decisions

Deciding on which action to take among alternatives is a core issue studied in many fields, from psychology (Dickinson & Balleine, 1994) to economics (von Neumann et al., 2007), cognitive neuroscience (Hunt et al., 2012; Kable & Glimcher, 2009; Padoa-Schioppa, 2011; Rangel et al., 2008; Rushworth & Behrens, 2008), and machine learning (Shachter & Peot, 1992; Sutton et al., 1998). In many human decision-making experiments, value is typically operationalised as a monetary reward (Berridge & Robinson, 1998; Berridge, 2000; Daw & Doya, 2006; Dayan & Daw, 2008). For deliberative model-based decision-making, rewards are dependent on specific responses and can incorporate changes to the calculated value. The first component of this process then involves a kind of episodic 'future thinking' (Atance & O'Neill, 2001; Buckner & Carroll, 2007; Schacter et al., 2007, 2008). Based on prior experience, we must predict the consequences of selecting certain actions.

Deliberative decisions demand a causal understanding of the statistical structure of the world (Niv et al., 2006; Rangel et al., 2008; Van Der Meer et al., 2012) and a network of complex abilities: from imagination (Buckner & Carroll, 2007; Schacter et al., 2007) to working memory (Gnadt & Andersen, 1988; Hill, 2008; Johnson et al., 2007) and motor control and sensorimotor learning (Harris & Wolpert, 1998; Körding & Wolpert, 2004; Wolpert & Landy, 2012). Together, these afford the active calculation of multiple options and the likelihood of the outcomes associated with choosing those options.

By contrast, artificial systems based on search processes do not rely on the above concert of abilities (McCorduck & Cfe, 2004; Newell et al., 1959; Pearl, 1984). In computer science, algorithms permit learning and the comparison of rewarding actions, achieving impressive action selection capabilities based on computational power, like with IBM's Deep Blue and Google's AlphaGo (Campbell et al., 2002; Silver et al., 2017). Importantly, unlike human learning agents, Deep Blue did not learn anything about chess or discover "patterns" for intuition. It simply searched for the optimal move using brute computational force. A large proportion of the computational study of human decision-making thus far has—due to the robotics/machine learning research using sequential update steps—concerned itself with the prediction step of reward-based decision-making (Barto et al., 1989; Botvinick & An, 2009: see also Section 1.3 below). But humans do not typically allocate time and effort to computing all potential paths to an outcome when deciding (Botvinick et al., 2009). As discussed in due course, we use alternative model-based strategies (Daw et al., 2005; Doll et al., 2012; Niv et al., 2006; Simon & Daw, 2011).

From animal studies where neural responses are decoded, we can understand the processes involved in deliberative decision making. These experiments provide evidence that rats use a similar deliberative system when choosing between options in a Tolman maze (T-maze: a maze where a T-intersection requires a choice, but recent changes to the outcomes at the end of each path demand a type of predictive 'future thinking'). For rats in T-mazes, this involves *place cells* (the neural activity of which can provide locations represented at any given time) in the hippocampal structure, just as with humans and other species (Redish, 1999; O'keefe & Nadel, 1978; Squire, 1992). By combining neural decoding and neurophysiological techniques, researchers have demonstrated that rats indeed do pause to deliberate when deciding on a path to take in the T-maze. Evidence of which comes from lesion studies to hippocampal regions (where memory based prediction is no longer available, just as with humans, Hassabis et al., 2007; Maguire & Hassabis, 2011; Schacter & Addis, 2009), and decoding studies showing the neural activity of path option representations preceding actions—searched sequentially and in a meaningful manner using upcoming places (Gupta et al., 2012; Johnson & Redish, 2007; Pfeiffer & Foster, 2013; van der Meer et al., 2010).

For humans, in economics and neuroeconomics, it has been reasonable and plausible to assume that different values are summed and weighed up before we decide upon an action to obtain some reward (Rangel et al., 2008). Yet after further study in neuroscience, this is now thought inaccurate (Dayan & Daw, 2008). The discrepancy concerns the fusing of multiple and distinct sources of information during human decision-making. Consider recent research where humans compare different outcomes with varying probabilities and magnitudes: both the behavioural and neuroimaging data support the notion that humans contrast potential outcomes based on the independent attributes of the options (Hunt et al., 2014; Hunt & Hayden, 2017; Kolling et al., 2016). Importantly, the participants' choices were then biased toward the option with a higher discriminability. To illustrate, we typically make choices (like which apartment to live in) not by summing across all defining characteristics of apartments and then performing one final evaluative process, but by comparing over each token characteristic type (noise level, neighbourhood, size, and so on) independently (Fellows, 2006). This implies that the elements of a choice are vital to understanding value-guided decisions (Vlaev et al., 2011). For neuroscience, this is bolstered by neural network models detailing a kind of competition of choices through mutual inhibition or evidence accumulation and race-to-bound/threshold schemes (FitzGerald, Moran, et al., 2015; Friston et al., 2017; Hunt et al., 2012; Kolling et al., 2016; Latimer et al., 2015; Solway & Botvinick, 2012; Wang, 2008). The nature of value guided decisions then seems more accurately treated in line with the above consideration of moment-by-moment calculations to value: that is to say, value-

guided choices do not follow a serial processing model but a continuous policy (Cisek & Kalaska, 2010).

The importance of continuous moment-by-moment updates to a reward-based quantity while learning and decision-making are explored in each subsequent chapter of this thesis. While we constrain value to one pre-defined target in each experiment, and decisions are constrained to a binary choice, the element under more detailed scrutiny is the continual process of combining multiple sources of probabilistic information and how anxiety modulates this learning and decision-making process—discussed in further detail in **Section 1.4** and **Section 1.5**.

1.2.3 Neural underpinnings of value and decisions

Reward-guided learning and decision-making research in the brain have shown the ventral and medial prefrontal cortex encoding values (Rushworth et al., 2011) and the basal ganglia encoding reward-based learning prediction errors (see **Glossary**, **Section 1.3** below, and Seymour et al., 2004). For reinforcement learning and value-based decision-making, there are, in fact, vital differences between the outcome provided (reward, punishment), the evaluation (euphoria, dysphoria), the effect this has on an animals behaviour (reinforcement, aversion), and the consequences to a lack of delivery from a predicted outcome (disappointment, relief, Redish et al., 2008; Land et al., 2008; Treadway & Zald, 2011). Reinforcement is something that increases the likelihood of a behaviour (Arvanitogiannis & Shizgal, 2008; Bielajew & Harris, 1991; Olds & Milner, 1954) and is linked to medial forebrain bundles and the neurotransmitter dopamine (see **Section 1.3**, Olds & Fobes, 1981; Shizgal, 1997; Wise, 2004). By contrast, aversion decreases responding via some penalising outcome and is less well understood at the neural level (Domjan, 2014; Ferster & Skinner, 1957; Hull, 1943; Mackintosh, 1974; Munn, 1950; Pavlov, 1927; Thorndike, 1932; Watson, 1907). Dopamine alters how much an animal will work, or how much vigour is dispensed when searching or approaching reward (Berridge, 2007, 2012; Berridge & Robinson, 2003; Niv et al., 2007). While euphoria (the evaluation of how good some outcome was) is coordinated by the opioid neurotransmitter system (of which there are three types of receptors and signals in the NA, hypothalamus, amygdala, and several connected regions: see Broom et al., 2002; Kieffer, 1999; Kieffer & Gavériaux-Ruff, 2002; Laurent et al., 2012; Le Merrer et al., 2009; Levine & Billington, 2004).

Neurophysiologically, two key areas involved are the ventral striatum (VS) and orbitofrontal cortex (OFC, Bray et al., 2010; McDannald et al., 2011). Each of these structures contributes

to motivation and are problematic for addictive behaviours (Koob & Volkow, 2010) and motivation-based disorders such as avolition, aboulia, and anhedonia (Der-Avakian & Markou, 2012; Dowd et al., 2016; Flagel et al., 2009; Mucci et al., 2015; Robinson & Flagel, 2009). Both structures are modulated by hippocampal projections, showing increased neural responses at the moment where the predicted possibilities of choices occur (Steiner & Redish, 2012; van der Meer & Redish, 2009). When the VS and OFC become damaged, in both animals and people, a resultant difficulty evaluating potential outcomes is observed (Bechara, Tranel, et al., 2000; Damasio, 2003; Fellows, 2006; McDannald et al., 2011; Rolls & Grabenhorst, 2008). And importantly, both structures include cells that represent both expected value and predicted reward, discussed below in further detail (Padoa-Schioppa, 2009; Padoa-Schioppa & Assad, 2006; Padoa-Schioppa & Conen, 2017; Schultz et al., 1998, 2000; Tremblay et al., 1998; Tremblay & Schultz, 1999). Importantly, the role the VS (and more specifically the nucleus accumbens [NA]) has on value and motivation learning from reward scales up to include (somewhat) healthy human endeavours such as science and its proxy for reward—*publications in high impact journals* (see Paulus et al. [2015] for an interesting treatment of this subject).

Consequently, Reddish (2013) has speculated that one reason humans developed opioid systems is to identify the outcomes of value—and consequently pleasure. Whereas we may have developed a dopamine system to orient toward reward-motivated behaviour. As such, one function of dopamine is to modulate wanting, while one function of opioids is to modulate liking (Redish, 2013). But how is value as discussed above represented? Subjective utility and value, according to economists, exhibits a concave relationship where higher levels of wealth bring lower increases to utility (Kahneman, 2011). This subjective utility is associated with different circuits of the brain when participants are informed to select the choice yielding the greatest long term reward given alternatives (Bartra et al., 2013; De Martino et al., 2006; Glimcher & Fehr, 2013; Kable & Glimcher, 2007; Rangel et al., 2008; Tom et al., 2007). Further, this subjective value appears linked to the biological values we mentioned concerning homeostasis (water, fat, sugar, and oxygen levels). However, a more domain-general scaling between values of *incommensurable objects* is needed so that information about the distinct kinds of rewards can be consolidated (FitzGerald et al., 2009; Plassmann et al., 2007). The likely neural candidates for this are in the ventromedial prefrontal cortex (vmPFC) and VS (Bartra et al., 2013; FitzGerald et al., 2009; Plassmann et al., 2007), while the OFC might realise a more *identity-specific* value function (Enel et al., 2020; Howard et al., 2015).

The same circuits process value when concerned with value from other domains, such as emotional value (Bechara, Damasio, et al., 2000; Winecoff et al., 2013), humour (Azim et al.,

2005), and the appreciation of a beautiful face (Aharon et al., 2001). Work with animals like non-human primates corroborates similar responses in line with subjective utility from prefrontal cortex (PFC) single cell recordings that might represent the coordination of probability, magnitude, and effort inputs (Hosokawa et al., 2013; Kennerley et al., 2011; Padoa-Schioppa, 2011; Padoa-Schioppa & Assad, 2006). The emerging consensus is that the brain estimates value for alternative choices using a common metric (Montague & Berns, 2002). In the account of reinforcement learning and Bayesian PC that follows, we shall lay out how this neatly lines up with how learning agents need to maintain calculations of value linked to specific states and take these to inform decisions to maximise reward.

1.2.4 Summary

We have concisely probed the meaning of value-guided choice and how subserving systems in the brain realise the specification of a target. By comparing value from an economics viewpoint to a psychological one, we have laid the groundwork for an internal value estimate that needs to be tracked to obtain some desired outcome. This is useful for the following sections and chapters based on value-guided decision-making and learning from rewards. Value is linked to a hierarchy of physiological mechanisms such as the motor system, memory, but also important informers like motivation and affective states. Together, these drive action toward value through learning from reward-based signals. We commission purposeful action toward things we identify as having high value, and these values influence several decision making systems, from Pavlovian to procedural and through to the model-based deliberate system. We presented evidence above that humans use dynamic learning from different sources of information to optimally decide upon actions to obtain reward, and the details of this dynamic process will be expanded on in later pages in **Section 1.4**. For now, we take on the question of which computational processes undergird how organisms link value to decisions and responses that move them closer to rewarding outcomes.

1.3 Surprising events: learning from experience

Computational tools like the above mentioned IBM's Deep Blue are simple systems that solve an exclusive complex problem. They are rigid and likely to be of little use when the problems around them change. Deep Blue, as an exemplar of inflexible exhaustive searching (rather than deep learning which appeared on the artificial intelligence scene a decade later) would fail ruinously at playing a game of backgammon instead of chess. The reason is that Deep Blue displays no general intelligence or ability to learn, but is instead a reflection of

computation individually tailored by the designer. Biological intelligence, however, is characterised by a flexible capacity to adapt and solve whichever problem is presented to the learner. For instance, non-human primates such as Gibbons (*Hylobates*, meaning 'forest walker') are territorial animals who continually guard boundaries from opponent conspecifics; they need to update their representation of their territory to track food supplies from different fruiting trees and appeal to potential mates. In the section to come, we detail how agents solve these learning problems using multiple distinct algorithms, later turning to how these processes are fit to changes in the environment.

1.3.1 Reward prediction

A core feature of animal behaviour is learning to value cues that help to anticipate future rewards. Reinforcement learning algorithms have thus given us useful models describing how anticipated rewards are updated given newly experienced outcomes from the environment (Inglis et al., 2020). From a behavioural ecology perspective, this makes sense. For example, predatory animals lacking a prediction mechanism might be trapped in a never-ending cycle of being surprised at the sudden appearance of prey—missing their fleeting chance of dinner. Humans go to great lengths to hunt things that are not necessarily of primary biological value, but instead have elaborated upon a constellation of proxy secondary rewards—such as photography, board games, reading, and of course, money.

In the work of Ivan Pavlov (1849–1936) and other *classical conditioning* studies, it was discovered that animals could not only predict future reward but link a conditioned stimulus (CS), like the sound of a bell, to a subsequent unconditioned stimulus (US), like food (Gormezano & Moore, 1966; Pavlov, 1927; Yerkes & Morgulis, 1909). It is thought of as a prototypical example of prediction learning (Niv, 2009). By contrast, the *operant conditioning* approach formalised by Edward Thorndike (1874–1949) and many others showed how animals are able to learn from the causal impact of their actions, given *contiguous* temporal and if possible spatial distribution (Staddon & Cerutti, 2003; Thorndike, 1898). Further, it was discovered that the *contingency* of the response-dependent effect was also of vital importance to operant conditioning. In the event the effect occurred from time to time spontaneously or occurred at a level approaching chance, learning slowed. Consequently, the rate of behavioural adaptation varies as a function of contingency strength and temporal and spatial contiguity of the outcome. Together, these shape how animals approach and avoid unconditional stimuli, allocate attentional resources toward events, or even on occasion mistakenly attempt to consume paired conditional stimuli during experiments. Observing these results and qualitative interpretations led Robert Rescorla and Allan Wagner to develop the

influential Rescorla-Wagner model, which provides an elegant description of how an agent learns the associations between co-occurring stimuli (Glautier, 2013; Rescorla & Wagner, 1972).

1.3.1.1 Rescorla-Wagner-Model

Put forward originally to describe quantitatively a normative model of classical conditioning, the Rescorla-Wagner model (RW) has since been applied to operant conditioning and even predicts numerous other behavioural patterns such as the blocking of new associations (Kamin, 1969), the extinction of previously learned behaviour (Kimble & Kimble, 1970; Quirk & Mueller, 2008; Rescorla & Wagner, 1972), and conditioned inhibition (where an unconditioned stimulus that blocks anticipated positive reward becomes negatively conditioned: see Rescorla, 1969). The satisfying simplicity of the model is that it describes how surprising events are vital to motivate adaptations to future expectations, directly inspired by Kamin (1969) and Bush and Mosteller (1951). More simply put, animals learn through the discrepancy between experience and their expectations, concisely articulated by the below equation:

$$\Delta V_t^A = V_{t-1}^A + \rho \cdot (\lambda^{US} - V_{t-1}^A) \quad (1)$$

In words, the RW equation quantifies learning by the incremental change to an assumed associative strength V_t^A from a stimulus or behaviour *A* on a single event, choice or trial *t.* The update ΔV_t^A proceeds as the previous trials associate strength V_{t-1}^A plus the difference between an outcome or unconditioned stimulus λ^{US} and V_{t-1}^A multiplied by a learning rate ρ. Increasing experiences with that type of event brings the agent's associations closer to the maximum associative strength λ^{US} of the unconditioned stimulus (positive or negative reinforcement). Further, the saliency of the term ρ (bounded by 0 and 1) regulates the adaption of the association strength, which is fixed. The learning rate scales the update V_t^A, with lower ρ producing slower updates.

Essentially, the RW model emphasises that learning occurs through difference signals—the violation between the expected associative value V_{t-1}^A and the outcome λ^{US}—resulting in a learning curve that decreases in slope as learning proceeds (Glautier, 2013). It succinctly describes how animals construct a basic understanding of their environment through the linking of their anticipations to stimuli. However, the RW algorithm suffers from limitations. Firstly, it does not distinguish between associative strengths, which may denote state or action

values (Averbeck & Costa, 2017). Further it considers the *CS* and *US* as qualitatively distinct, and thus fails to incorporate the second-order objects humans learn to value, like money (as touched on above). To illustrate, if cue *Y* predicts an outcome of affective value (like food or pain) and cue *Z* predicts cue *Y*, a consequence is that cue *Z* also gains reward predictive value (Niv, 2009). Lastly, the RW algorithm fails to predict learning in advance at smaller timescales.

1.3.1.2 Temporal difference learning

Around a decade later, a newer refinement was proposed by Richard Sutton and Andy Barto that emphasised the importance of predictions across time—an aspect considered missing from the RW model (Sutton & Barto, 1981). Sutton and Barto's (1989; 1998) temporal difference (TD) reinforcement learning connected the RW psychology model with control theory, and also to work by Richard Bellman (1958) on operations research (to which we will turn shortly when discussing behavioural policies: **Section 1.3.1.4**). In doing so, they derived a powerful and general algorithm subsequently applied throughout robotics, psychology, machine learning and beyond (Kurth-Nelson & Redish, 2009; O'Doherty et al., 2003; Sutton & Barto, 2018).

TD is a method for learning to predict. It is widely used in RL to predict future reward or value functions. An everyday example of which is learning to predict that the distinct redolence of cooking typically precedes a delicious meal. TD learning thus mathematically describes how humans can learn a prediction from another later prediction. It is the learning of some value function without a model. This seizes upon vital temporal parts of learning that the RW model is incapable of, like the value of sequential timing for conditioned and unconditioned stimulus pairings (Sutton & Barto, 1981).

There are two critical differences between the RW and TD models. First, the RW model assumes predictions concerning reward are obtainable from single individual cues. Whereas TD is founded on categorisations of the world known as *states*; or, s_t (in reward learning, the information that predicts upcoming rewards, Daw, 2003; Sutton et al., 1998). Secondly, the TD model predicts *value*. In doing so, it addresses the shortcoming of distinguishing between states and action values and also the second-order conditioning in humans and the temporal relationships within trials during experiments (Averbeck & Costa, 2017; Niv, 2009). Therefore, TD learning shows how we can use the next prediction to form a target, learning the value function simultaneously with executing actions. All the agent's predictions of future reward are incorporated into the state value, which expresses all the (time-discounted) rewards that the

agent anticipates to obtain in following states. Prediction errors (PEs) are thus calculated in TD learning as the discrepancy between rewards and the value of the next state relative to previous states' values:

$$\Delta V^{A}_{(s_t)} = V^{A}_{(s_t)} + \alpha \cdot \left[R_{t+1} + \gamma \cdot V^{A}_{(s_{t+1})} - V^{A}_{(s_t)}\right] \quad (1)$$

Above, a state-value function $\Delta V^{A}_{(s_t)}$ is derived from the current state the agent is in (s_t), where $\alpha \in [0,1]$ is the learning rate and γ is a discount function that determines how much the agent prefers reward now compared to later. An action from a behavioural policy π in a state (s_t) $\left(A_t = \pi(s_t)\right)$ leads to state (s_{t+1}) where the agent encounters a reward shown as R_{t+1}. The term in the brackets is the temporal difference error: the currently expected future reward when executing an action in a state. Namely, the total reward R_{t+1} minus $V^{\pi}(s_t)$ plus the expected discounted future reward $V^{\pi}(s_{t+1})$ in the new state s_{t+1} (Butz & Kutter, 2016).

The TD error supplies the agent with a value PE term, with 'more reward experienced than predicted' constituting a positive TD error, and 'less reward experienced than predicted' being a negative TD error. What falls out from this is the indication that the action just sanctioned was either better or worse than expected, or exactly as anticipated, with zero learning required. The influence of current reward on updates to the value function is modulated by the learning rate, which varies between 0 and 1, with values closer to 1 meaning the reward value is taken on rapidly, and values closer to 0 determining slower updates.

The only distinction between the PE terms of the RW and TD models is the presence of $V^{\pi}(s_{t+1})$, which indicates both value from the state transition and the future reward linked with the new state. To take one out of many everyday examples, an upcoming talent in sports who wins some form of national prize might then find themselves in line for a lucrative team signing, perhaps even in position to represent the nation at the Olympics, but also to acquire profitable sponsorship deals, endorsements, PR opportunities, celebrity appearances, and presumably the list goes on.

Getting back to reinforcement learning, the indispensable message connecting both the RW and TD models—and the later discussed Bayesian generative model-based learning schemes—is the global currency for learning encapsulated within the reward PE (RPE) difference term:

$$\delta_t = R_t - V_t \quad (2)$$

Where the PE delta (δ_t) can inform future predictions by measuring the discrepancy between reward that is predicted V_t and the reward that is received R_t.

1.3.1.3 Dopaminergic reward prediction error signals

Following what is perhaps one of the more exhilarating and phenomenal series of findings in neuroscience, RPEs are encoded in phasic dopaminergic bursts in the brainstem (Ljungberg et al., 1992; Montague et al., 1996; O'Doherty et al., 2003; Schultz et al., 1997, see **Figure 1**). The theory asserts that RPEs advance synaptic plasticity in the striatum, turning the links between actions and their outcomes into optimised behavioural policies using transient bursts of firing in dopamine neurons in the Ventral Tegmental Area (VTA) and Substantia Nigra pars compacta (Daw & Tobler, 2014; Hollerman & Schultz, 1998; Wang et al., 2018). These structures take input from diverse afferents (Watabe-Uchida et al., 2012) and distribute widely to cortical and subcortical regions, most strikingly in the striatum and PFC (Haber & Fudge, 1997). Importantly, Schultz and colleagues found phasic dopamine bursts increase as a function of increases to the predicted value, decrease with unpredicted reductions to value, and more generally represent the PEs outlined in the RW and TD learning models (Hollerman & Schultz, 1998; Padoa-Schioppa, 2009; Padoa-Schioppa & Assad, 2006; Padoa-Schioppa & Conen, 2017; Schultz, 1998; Schultz et al., 1998, 2000; Schultz & Dickinson, 2000; Tremblay et al., 1998; Tremblay & Schultz, 1999). In the last twenty years of research, growing evidence for this RPE driven theory has earned it the place as a standard learning model in neuroscience (Wang et al., 2018).

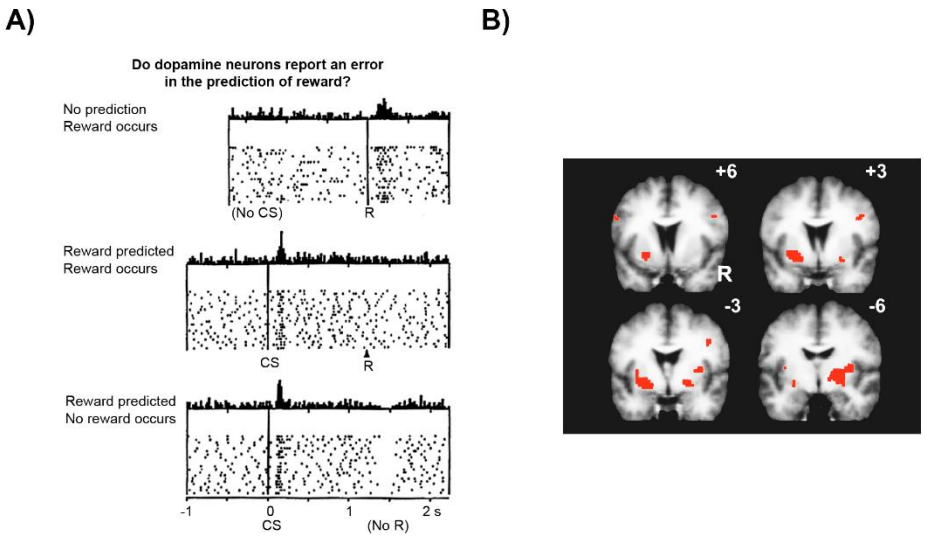


Figure 1. Dopaminergic Reward Prediction Errors. A) In panel A is the single cell recording of the macaque monkey VTA while performing a classical conditioning task. The top panel shows the dopaminergic neural response in a no-conditioning phase (CS), with increased dopamine firing after the reward is provided. The middle panel demonstrates the firing of dopamine neurons occurring after a predictive conditioned stimulus (CS) is provided—but not at the time of reward. And the bottom panel gives the same positive PE response at the CS with a later negative PE response when a reward is predicted but absent. (Figure adapted with permission from Schultz et al. [1997]. Copyright © 1997, © 1997 American Association for the Advancement of Science). **B)** This figure depicts the RPE signal in a fMRI experiment supporting the dopaminergic response from Schultz et a. (1997) revealing BOLD activation in the VS. (Figure adapted with permission from O'Doherty et al. [2003]. Copyright © 2003 Cell Press. All rights reserved).

In these experiments, recordings were taken while monkeys were engaged with simple instrumental or Pavlovian conditioning. Dopamine was originally thought to represent the value of the rewarding stimuli (Wise, Spindler, DeWit et al., 1978; Wise, Spindler, Legault et al., 1978). Yet these studies demonstrated that phasic dopamine bursts were only initially associated with the rewarding stimuli (as they were initially surprising). Later, they responded only to the paired predictive cue of the rewarding stimuli—as TD errors only occur to *unpredicted* stimuli. The biological underpinnings of this RPE signal in the dopaminergic system have since been studied extensively (Niv, 2009). To outline, dopamine neurons baseline fire at a few spikes per second; positive responses involve bursts of these spikes and negative a cessation; this transmits a 'better-than' and 'worse-than' expected signal confirming the RPE hypothesis of dopamine originally propounded by Montague et al. (1996) and Schultz et al. (1997). These results have since been replicated by a welter of neuroimaging work in rodents (Cohen et al., 2012; Steinberg et al., 2013), non-human primates (Bayer & Glimcher, 2005; Fiorillo, 2013; Fiorillo et al., 2003; Tobler et al., 2003; Waelti et al., 2001), and humans (Bang et al., 2020; D'Ardenne et al., 2008; Iglesias et al., 2013; O'Doherty et al., 2003; Pessiglione et al., 2006; Rutledge et al., 2010; Seymour et al., 2004).

Beyond these core findings followed further intriguing discoveries. Notably, prior reward contributes to the brain's RPE signal corresponding precisely to the TD model scheme; that is, exponentially weighted according to the average of past experience (Bayer et al., 2007; Bayer & Glimcher, 2005). Also, varying probabilistic rewards are represented by a dopamine firing response which back-propagates the error signal (Niv, Duff, et al., 2005) proportional to the degree (Tobler et al., 2005) and contingency of the expected reward (Fiorillo et al., 2003; Morris et al., 2004; Niv, 2009). Moreover, meta-analyses distinguishing between absolute and signed PEs discovered that absolute PEs were specifically associated with the dopaminergic midbrain while signed PEs were expressed particularly in the ventral part of the striatum (Fouragnan et al., 2018; Garrison et al., 2013).

1.3.1.4 Electrophysiological evidence of RPEs

Additional studies have since associated RPE signals to other areas in the brain, such as activity from the anterior cingulate cortex (ACC) which is thought essential for encoding information concerning beliefs (for example probabilistic outcomes and values) and posterior medial frontal cortex (pMFC, Debener et al., 2005; Holroyd et al., 2003; Kennerley et al., 2009; Montague et al., 2004; Yeung et al., 2005). Extracting event related potentials (ERPs) from scalp EEG signal in humans shows the RPE across frontal medial electrodes, a negative deflection reflecting feedback (Cohen et al., 2007; Holroyd & Coles, 2002) pinpointed using trial-wise analysis of EEG data with fMRI to the ACC (Debener et al., 2005). One example of RPEs in EEG is the error related negativity (ERN), an ERP in response to explicit errors (approximately 100 ms post error across frontocentral sensors) and the feedback ERN (fERN), a negative deflection approximately 250 ms subsequent to negative feedback in central and parietal regions (Holroyd et al., 2003; Montague et al., 2004; Nieuwenhuis et al., 2004; Yeung et al., 2005). Of particular relevance to the work in this thesis, the fERN is more specifically thought to express the degree of surprise (PE) and has been estimated in response to rewards using reinforcement learning models (Gehring & Willoughby, 2004; Holroyd & Coles, 2008; Holroyd & Krigolson, 2007). An additional ERP component related to RPEs is the P300, which is proposed related to valence and surprise, exhibiting a parietal topography (approximately 250–500 ms post error, Hajcak et al., 2005, 2007; Polich, 2007; Wu & Zhou, 2009).

1.3.1.5 RPE Summary

RPE signals are not an exhaustive and conclusive explanation of dopaminergic functioning within the brain. The dopamine cells and projections exhibit a wide display of responses in line with their afferents and efferents and the circumstances governing reward (Lammel et al., 2011, 2012; Matsumoto et al., 2016; Matsumoto & Hikosaka, 2009). Moreover, modulating dopamine without reward outcomes does not lead to learning (Adamantidis et al., 2011), and rewards and reinforcement learning are possible without the involvement of a dopaminergic response (Berridge, 2007; Cannon & Palmiter, 2003; Flagel et al., 2011). Also, while it is assumed that dopamine neurons encode the RPE, it is also widely understood that the RPE signal is an incomplete and insufficient predictor of dopamine response (Inglis et al., 2020). To illustrate, dopaminergic activity is also documented in response to salient information and novel experiences with no connection to reward (Horvitz, 2002), and significant individual differences exist in the magnitude of dopamine RPE responses, for example varying with personality type (Pickering & Pesola, 2014).

Consequently, the picture concerning dopamine function during reward based learning and decision making is far from complete; however, a full review of dopaminergic function is outside the potential scope of this thesis (for further details see Bromberg-Martin, Matsumoto, & Hikosaka, 2010; Schultz, 2016; Wise, 2004). For now, it is sufficient that neural representations of PEs signalling updates to beliefs about rewards in humans can be detected using electrophysiological recording methods, verified by non-human animal and human imaging research, as we take this approach in **Chapters 2 and 3** using EEG and in **Chapter 5** using MEG. In fact, it is thought that superficial pyramidal cells encoding PE signals represent a substantial quantity of the M/EEG signal (Buffalo et al., 2011; Edwards et al., 2012; Friston & Kiebel, 2009a).

1.3.2 From prediction to action

Predictions, however accurate, do not in and of themselves elicit rewards; this only occurs when we mandate actions, interact with our environment, and elicit feedback. Consequently, it might seem plausible to suggest that the highest aim for predictive learning is to stimulate accurate action selection (Niv, 2009). As mentioned above, instrumental conditioning describes the process by which a learning agent takes on a new repertoire of responses that can result in receiving rewards or avoiding punishments (reinforced learning). Below we take on one specific realisation of how state values can be used to inform actions selection using

Q-learning. But before that, we must briefly discuss how behavioural policies are formulated using Q-functions and the Bellman equation.

1.3.2.1 Behavioural Policy

As touched on above, the learning agent's behavioural policy is defined by $\pi: S \rightarrow A$, which denotes that for all possible states of the environment $s \in S$ there is a corresponding action the agent will take $a \in A(s)$. This behavioural policy is then refined toward some value function as specified in the TD model. However, this does not provide information concerning the alternative possible actions for any given state $A(s) \neq \pi(s)$. For that, we need a state-action-value function known as a Q-function: $Q^{\pi}(s, a)$. This gives values for all potential actions $a \in A(s)$ for all possible states $s \in S$ in the learning agent's environment. Given this definition, an optimal policy π^{*}can be mathematically defined as:

$$\pi^{*}(s) := \underset{a}{argmax} Q^{*}(s, a) \tag{3}$$

Yet the optimal values for $Q^{*}(s, a)$ are unknown. These can be derived using the Bellman equation (Bellman, 1958), which is used to calculate how valuable a learner should estimate each state is, given they aim to maximise reward in that environment. While we do not have space here to consider Bellman's principle of optimality and dynamic programming based on recursion (Bellman, 1966), the important part to take away is that the TD learning model used Bellman's concept that an optimal behavioural policy would be one that sums immediate reward with discounted future estimated reward and combined that with PE-based learning.

1.3.2.2 Q-Learning

The final link forming the foundation of RL approaches we will consider here—before moving on to a generative model-based approach built on these foundations—is Q-learning, a type of trial and error technique used to update the value of alternative actions (Watkins & Dayan, 1992). Q-learning proceeds as an iterative process of updating the Q-value using TD learning. It diverges from TD by distinguishing the policy π from the values, which update Q-value estimates using the Bellman equation. To make this clearer, Q-learning is alternatively known as an *off-policy* RL approach. This is because the action performed at any given time does not need to comply with the behavioural policy. Instead, the Q-value is updated using the state s_t, the action executed a_t (that is not necessarily from $\pi(s_t)$), and the reward experienced R_{t+1} in the following state s_{t+1}:

$$\Delta Q(s_t, a_t) = Q(s_t, a_t) + \alpha \cdot \left[\left(R_{t+1} + \gamma max Q(s_{t+1}, a_{t+1})\right) - Q(s_t, a_t)\right] \quad (4)$$
at+1

The maximum operator means the best future discounted reward updates the Q-value function. Again, the learning rate is given by α and the discount function by γ. The goal of Q-learning is to find the optimal policy by learning the optimal Q-values (Q^*) for each state-action pair. One further useful component is that a learning agent can describe how good an action and state is using a single number, which, as mentioned before in **Section 1.2.3**, may serve to update values represented in the OFC in humans (Schuck et al., 2016). The above Q-learning procedure provides a simple update rule whereby a learner can perform optimally in a state it first is wholly ignorant of, simply by gaining sufficient experience.

The simplicity of fitting Q-learning models to behaviour has since furnished a more complete mechanistic description of animal predictive learning using the reinforcement approaches detailed above, and also informed multiple studies testing the influence of pharmacological modulation on behaviour (Marshall et al., 2016; Rutledge et al., 2009).

1.3.2.3 Q-learning in the brain

Neurobiologically, the striatum is thought to be the ideal candidate for value learning and action selection in the brain (Dayan & Daw, 2008; Doya, 2000; Reynolds et al., 2001; Samejima et al., 2005). Interestingly, when the basal ganglia is not functioning properly, the capacity to pair outcomes to actions is inoperative, resembling defective Q-value learning (Mishkin et al., 1984). Our understanding has since been sharpened toward striatal cells, with human fMRI work demonstrating neural correlates of action-values in the striatum (FitzGerald et al., 2012; Pessiglione et al., 2006). More contemporary evidence points to the VS being linked to distinct processes, from information expectation to the anticipation and receipt of reward (Filimon et al., 2020). In a study by Samejima et al. (2005)—where monkeys made voluntary saccades to spatial locations to obtain rewards with varying probability—striatal cells tracked the value of making a saccade in line with the contingencies of reward, to either the left or right. Along with other animal studies (Lau & Glimcher, 2008; Seo et al., 2012), this evidence suggests cells in the striatum code for action values. While the exact nature and subtleties of how these kinds of RL algorithms are organised and enacted in the brain are still under discussion (Vo et al., 2014), robust connections between dopaminergic neurons in the VTA and substantia nigra (realising RPEs) and the striatum (Haber, 2003) form a persuading case for an action selection system (Frank, 2011). This is particularly compelling given akinetic

mutism, an apathetic disorder leading to failed action selection in humans resulting from a lesioned striatum.

In addition, Parkinson's disease is typified by a depletion in dopamine neurons in the substantia nigra (Dauer & Przedborski, 2003; Frank, 2005). These dopaminergic neurons are thought vital to RL through representing the delta PE function (Schultz et al., 1997). However, our understanding of the relationship between phasic and tonic dopamine firing and decision making and Parkinson's is complex and remains incomplete (Moustafa & Gluck, 2011; Rutledge et al., 2009). What we do know is that Parkinson's impairs decision making (Frank et al., 2007; Shohamy et al., 2005) and action (Denny-Brown, 1962; Jankovic & Tolosa, 2007; Koller, 1992; Martin, 1967; Wiecki & Frank, 2010), and that tonic levels of dopamine are linked with situation recognition (Redish et al., 2007) and movement motivation or invigoration (Berridge, 2012; Berridge & Robinson, 2003; Friston et al., 2012; Kurniawan et al., 2011; Montague, 2006; Niv et al., 2007; Salamone et al., 2009; Wise, 2004). As a final consideration, it is also possible to simulate dopaminergic lesions and display dysfunctional behavioural responses like in Parkinson's disease by altering the precision (uncertainty) associated with PE signals encoded by dopaminergic firing (FitzGerald, Dolan, et al., 2015; Friston et al., 2012). Put together, the case above for the motif of reward learning is one where the updated connection is regulated by the potency of the dopaminergic input coming from RPE signals. Actions commissioned toward some rewarding outcome will update the behaviour policy, and this process has been shown by calculating the degree of synaptic potentiation after intracranial self stimulation of the substantia nigra in the basal nuclei (Reynolds et al., 2001).

1.3.2.4 The limits of reinforcement-learning

Q-learning and other RL approaches thus describe how learning agents arrive at a behavioural policy without any knowledge of the environment's reward structure. RL algorithms like these have found many wider applications in artificial (machine) intelligence. Yet their simplicity in reducing complex learning patterns down to the sole aim of tracking a single value via an update term is limited. Human learning in the brain resolves multiple classes of problems simultaneously. Biological learning exhibits far greater levels of dynamic and adaptive learning complexity. For these types of uncertain and changing problems, information theory indicates that RL may be suboptimal (MacKay & Mac, 2003). A better method when working with the kinds of dynamic states that comprise our world, according to probability theory, is to quantify uncertainty and represent beliefs as probability distributions (Jaynes, 2003; O'Reilly et al., 2012). RL does not provide us with the capability of rendering our confidence and using this to inform future learning. Nor does it typically perform well as a model of learning in changing

environments. Thus, within a neurobiological context, RL appears too simple for most real-world learning problems.

Take for example a simple case of learning from a coin toss where you bet on tails. Using a model-free RL approach $V_{heads} = 0$ and $V_{tails} = 1$. Before the toss, a value of 0.5 would be placed on the outcome. When the fortuitous tails is revealed following the flip, a positive RPE is commissioned (+0.5). Given a heads outcome, this would be negative (−0.5). The unambiguous product of this scenario is the mandating of RPEs when there is precisely nothing to learn. Yet as we will see in the sections to come when including a model of the environment in human learning, one would represent the statistical composition of coin flips as having a probability of 0.5. While a coin flip may elicit a response from observing the outcome, knowing that the coins are completely unbiased would ensure the information does not update beliefs about the statistical structure of the possible outcomes. Numerous studies have, however, demonstrated that the resolution of probabilistic events that drift from 0.5 elicit strong dopaminergic responses in the brain (D'Ardenne et al., 2008; Gläscher et al., 2010; Hart et al., 2014; Rutledge et al., 2010; Wolfram Schultz, 2015).

Thus the primary limitation of reinforcement models is the inability to calculate and use uncertainty to inform beliefs. Irrespective of the number of flips we make, *V* prior to the flip is always precisely the same as though a reward had been encountered. We are left with the poorly rendered characterisation of a world in which we could flip a coin 5000 times and still learn wholly nothing about the reward state. Given a model, we could learn the subtleties of the coin, perhaps slight imbalances in the physical properties that lead to a bias. But a RL model provides us with no means to represent that decreasing uncertainty with each flip. Implications for the neurobiology underlying this incautious process would be the complete saturation of the brain in dopaminergic output. This would consequently lead to defective adaptation in the brain associated with events where no learning should have occurred (Niv, 2009; Yu & Dayan, 2005). While some have attempted to incorporate uncertainty into RL (Preuschoff & Bossaerts, 2007), the majority consider it a critical shortcoming (Clark, 2015; Glimcher & Fehr, 2013).

1.3.2.5 Models of behaviour

As alluded to above, the model-free approach describes a learning agent who does not use an internal model of the learning environment. Intuitively, this seems like a poor resolution description of human learning: applying trial and error as a solution to all scenarios that require learning. Model-free RL methods also perform poorly when reinforcers are not forthcoming

from the world or if they are few and far between. For instance, when humans participate in intricate arrangements of behavioural sequences to obtain a far off goal where reward may not be received in the intermediate junctures or pauses—such as dieting. The distinction between behaviour that is modelled and reinforcement learning descriptions that do not require a model is also bolstered by experimental psychology research on reinforcer devaluation. Independent habitual and goal-directed systems were established by carefully observing the behavioural responses of animals taught to obtain food from lever presses (Everitt & Robbins, 2005; Redish, 2013). After devaluing (by pairing outcomes to undesirable ones using quinine or saline) one would expect continued lever pressing (stimulus-response) if only reinforced habitual responses controlled learning. Alternatively, one would expect extinguished lever pressing if the devaluation of the goal-directed response had been changed (Adams & Dickinson, 1981).

Devaluation thus gave a straightforward and sophisticated means to test the characteristics of learning from behavioural responses. We also know from a welter of research, in the tradition of Helmholtz's (1866)—to be discussed in more detail in the following chapter on Bayesian learning—that the assignment of perception is to reassemble from noisy sensory data the true latent structure of the world (Gregory, 1980). Accordingly, we conceive of perception as an inversion of the generated model of sensory data from our environment. Using predictive learning to infer upon the causes of these sensory inputs can further fine-tune an accurate model of the world, and this is irrespective of reward or explicit reinforcers (Arnal & Giraud, 2012; Dayan et al., 1995; Hohwy, 2013; Llinas, 2002; Summerfield et al., 2006; Yon et al., 2018).

Specific evidence of the brain making predictions about upcoming sensory events comes from (as just one out of numerous possible examples) an fMRI study where BOLD activity from the visual cortex was decoded using a separate localiser to determine the processing pattern when perceiving distinct colours, such as yellow (Bannert & Bartels, 2013). Subsequently, the authors gave grayscale images of bananas and demonstrated that human participants actively predict the properties of visual stimuli (such as the absent yellow from the banana) by recovering fMRI activity previously linked with the perception of yellow (Bannert & Bartels, 2013).

The history of work on human and animal cognition shows that our understanding has not been limited to only reinforced learning, with animals being proposed to use internal models of their immediate surroundings to find reward from as early as Tolman's rats (Tolman, 1948). Recent research using a task designed to test the part played by model-free learning and the

part played by model-based learning shows that the brain uses both, with prefrontal dopamine associated more with model-based learning and striatal dopamine with model-free learning (Doll et al., 2016). Put simply, both model and model-free algorithms are thought to work concurrently (Daw et al., 2011; Wunderlich, Dayan, et al., 2012). Model-based approaches, while more complex, are vastly more flexible. In terms of RL, agents can use a generative model to shepherd learning to different states based on a modelled network to other such states (Dayan, 1993). Models further allow us to plan, conceive of, contemplate, and imagine a newly valued state—and in which way we can obtain it. Continued examination of the relative contribution of model-free and model-based learning algorithms shows that the equilibrium can be modulated by dopamine (Wunderlich, Smittenaar, et al., 2012) and the balance can be interfered with by the functioning of the DLPFC (Smittenaar et al., 2013). Moreover, and of particular relevance to the work presented in this thesis, uncertainty processing in the frontopolar cortex arbitrates between model-based and model-free learning (Lee et al., 2014). The work comprising this thesis will introduce how the brain uses a hierarchical generative model to learn the statistical composition of its reward environment. We also elaborate upon the role emotional states like anxiety play in shaping how we construct, retain or update our internal model.

1.3.3 Summary and relevance

Let us pause for a moment to consider the role of the RL algorithms of the last section for the work in this thesis. The foundation of learning is the difference signal: the surprise elicited when learning the associations between a stimulus, an action, and a reward. The RW model gives us a trial-based detailing of how sensory inputs are paired with rewards. The TD model contributed to our understanding of how a learning agent can seize upon the closest predictor of reward. And the later introduced Q-learning algorithm formally illustrates how learners choose actions to acquire rewards. At all events, these mathematical descriptions have proven practical and of beneficial use for investigating the neural correlates of learning computations and system processes in the brain. However, RL algorithms do not incorporate uncertainty, a compelling wrinkle to be taken up in later models better suited to describe human learning in real-world uncertain environments. Models provide a means to incorporate dynamics and uncertainty into our understanding of human learning, affording adaptable behaviour and planning. Models also aid in the testing of and bridging between the different levels of analysis: the computational, the algorithmic, and the physical. From an evolutionary selection perspective, we might also expect human learning to correspond closely to a statistical optimum; in the section to follow, we explore research detailing how the brain continuously uses and refines a generative model of its sensory input.

1.4 Learning from dynamic and uncertain environments

Everyday environments are replete with uncertainty. While autonomic responses in animals efficiently deal with fixed contexts, most real-world and interesting decisions are made under conditions where environmental contingencies change. To illustrate just how important the brain is for making decisions under uncertain conditions, consider the autonomic response action-selection system of *reflexes*. This describes a genetically hardwired system that affords an organism a direct one to one mapping between incoming sensory data and response. Aside from some reflexes connecting with the brainstem, the brain beyond is not involved with these action-selection responses. The reflex action-selection system is woven into our spinal cords and peripheral nervous systems so as to be effectively without uncertainty. These are the simplest kinds of decisions we can make, learned over the course of an evolutionary time scale—rather than being learned from birth. The opportunity they provide us, among many others, is the kind of rapid response necessary to avoid danger—such as with the retraction of a hand gripping a searing hot baking tin. While we have the ability to cognitively override these reflex responses, the above simply serves to demonstrate how the acutely overwhelming number of situations where uncertainty flourishes are typically responded to using the brain. The intricate and sophisticated manoeuvres of which are concerned with seizing upon which decisions to make given uncertain contexts. In fact, recent theoretical accounts of predictive processing suggest that the principal function of the neocortex is to reduce uncertainty by forming increasingly accurate predictions about upcoming events (Clark, 2013b; Friston, 2005; Heeger, 2017; Hohwy, 2016; Rao & Ballard, 1999; Shipp, 2016).

1.4.1 Bayesian inference

Uncertainty, however, makes optimising predictive behaviour challenging (Behrens et al., 2007; Clark, 2013b; Knill & Pouget, 2004). To operate successfully we need to estimate the hidden states in our world, adapt to their changes, and determine the uncertainty of our estimations (Friston, 2005; Friston & Kiebel, 2009a; Shipp, 2016; Summerfield & de Lange, 2014). While the brain is placed under a heavy demand to process this deluge of uncertainty, current theoretical accounts assert that it does so by optimally quantifying uncertainty using Bayes' Theorem; that is, performing inference on uncertain quantities according to the rules of probability theory (Jaynes, 2003). Bayesian inference thus provides the ideal computational and conceptual tools to update probabilistic beliefs during continuous learning about uncertain states (maximizing Bayesian model evidence, see Clark, 2013b; Friston, 2010; Friston, Parr, et al., 2017; Hohwy, 2016; Mumford, 1992). More simply put, our generative model of the

hidden states in the world is a compromise between what we knew before and any new information. This accommodates both inference at the perceptual level (where sensory input refines higher-level hypotheses about the world) and active inference (where predictions motivate actions to confirm prior beliefs). Mathematically, Bayes' theorem—explicitly conditioned upon a generative model—is articulated by the below expression:

$$p(x|y,m) = \frac{p(y|x,m)\,p(x|m)}{p(y|m)} \quad (5)$$

Where our posterior belief or generative model *m* about the world *p(x|y, m)* consists of the probability of the hidden states *x* conditional upon some observed sensory data *y.* This posterior estimate is calculated as proportional to the likelihood function *p(y|x,m)* times the prior probability *p(x|m)* weighted by the uncertainty (or its inverse, precision) of each. The model evidence *p(y|m)* is a normalisation constant that can be used in Bayesian model comparison (Da Costa et al., 2020; Stephan, Manjaly, et al., 2016).

In this Bayesian framework, beliefs are represented as probability distributions rather than point estimates (Doya et al., 2007; O'Reilly et al., 2012). The method by which organisms might calculate these probabilistic belief estimates is still an open and thrilling topic of debate (Beck et al., 2008; Deneve, 2008; Ma et al., 2006; Ma & Jazayeri, 2014; Pouget et al., 2013). Importantly, harking back to Helmholtz (1866), perception in Bayesian inference is considered the inversion of this generative model (de Lange et al., 2018; Friston, 2003; Hohwy, 2013; Mathys et al., 2014; Press & Yon, 2019; Stephan, Manjaly, et al., 2016; Summerfield & Koechlin, 2008). And the question of how humans invert this generative model is also an unresolved and hotly debated topic (Bassett et al., 2018; Betzel & Bassett, 2017; Frässle et al., 2018; Friston, 2003; Friston et al., 2014). We will not rehash that debate here, choosing instead to focus on one inversion procedure to approximate Bayesian inference as provided by the Hierarchical Gaussian Filter, detailed in **Section 1.4.6**.

There are multiple and distinct sources of noisy sensory information that we receive from both the world and from within our body (Ainley et al., 2016; Bland & Schaefer, 2012; O'Reilly, 2013; Seth & Friston, 2016; Yu & Dayan, 2005). Much of our cerebral operations consist of combining these multiple uncertain information sources, with neural processing of these events being intrinsically uncertain (O'Reilly et al., 2012). The probability of these events can be described in Bayesian systems as using a Probability Density Function (PDF). Given the presumption that the PDF is of a Gaussian distribution, our optimal assessment of the most likely estimate would be the mean. (Where the width of the probability distribution can be

interpreted as the variance [σ; or, its inverse: precision π]). The variance is thus the degree of uncertainty we place on our estimate of the probability of some event. More precise prior beliefs, for instance, dominantly inform posterior belief updates. Returning to Bayesian inference, as touched on above, our posterior belief is a compromise of the likelihood function and the prior probability integrated inversely proportional to our uncertainty (i.e. using a ratio of precisions: the weight of the prior relative to the observation):

$$\frac{\pi_y}{\pi_{x|y}} \tag{6}$$

Where $\pi_{x|y} = \pi_x + \pi_y$. How this modulates posterior beliefs is visually presented below in **Figure 2**. This ratio of precisions describes an equilibrium between sensory input and prior expectations, where assigning more weight or confidence to sensory processing, for instance, is equivalent to lowering the weight on priors and predictions. Importantly, however, disruption to how humans estimate precision is thought to give rise to aberrant posterior beliefs and false inference in various clinical populations (see **Section 1.4.8**, for example, in psychosis, schizophrenia, and autism Friston, 2017; Hauser et al., 2016; Lawson et al., 2014; Parr, Benrimoh, et al., 2018; Parr, Rees, et al., 2018). It is this possibility for altered or biased beliefs and misestimation of uncertainty that motivates the investigation of the effect of anxiety in the work of this thesis (Tappin & Gadsby, 2019). But here, we hypothesise that the healthy general population are also subject to transient changes to Bayesian belief updating, not just neuropsychiatric populations.

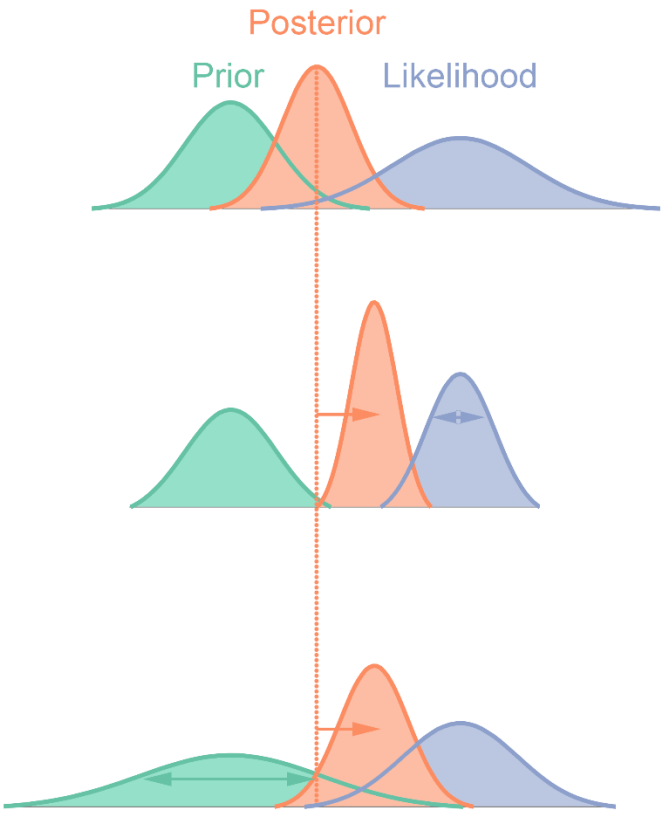


Figure 2. The effect of precision-weighting posterior belief updates. For the middle and lower panels, the mean of the probability density distribution remains the same, while precision (the inverse of the width of the distribution: $\pi = 1/\sigma^2$) is modulated. This shows the effect of precision on posterior belief updates. In the top panel, the prior drives the posterior update as it features higher precision than the likelihood distribution. The mean of the posterior is thus a compromise between the two probability densities and it features a higher precision than its constituent parts. In the middle panel, the prior remains the same, but the precision of the likelihood is increased, drawing the posterior estimate toward the likelihood. In the bottom panel, the precision of both the prior and likelihood are decreased, but the prior features a far lower precision estimate, leading to no relative change in the mean of the posterior estimate, but a concomitant decrease in precision.

The merging of distinct sources of information according to precision results in an intuitive outcome: our posterior belief has higher precision than its constituent inputs (as our certainty increases by adding more unique sources of information, see **Figure 3**, O'Reilly et al., 2012). As an everyday example, consider the hapless pursuit of a mosquito in your bedroom late at night. Low light levels make visual estimates of the gadfly's location of low reliability, while auditory and somatosensory feedback can be used with high reliability to determine proximity. However, to best determine, isolate, and remove the nocturnal nuisance, we need to combine our PDFs over visual, auditory, and somatosensory information using the Bayesian approach, according to our uncertainty. Clearly, in this situation, the ideal strategy is to raise the precision of visual sensory feedback by simply increasing the available light in the room. In experimental work, this has most prominently been shown in visual and haptic feedback experiments and

in the sensorimotor domain. In a study where participants needed to judge the height of a bar under varying conditions of uncertain visual and haptic input, the integration of these distinct sensory inputs followed a Bayesian process, as described by the model of individuals' task behaviour (Ernst & Banks, 2002, see **Figure 3**)

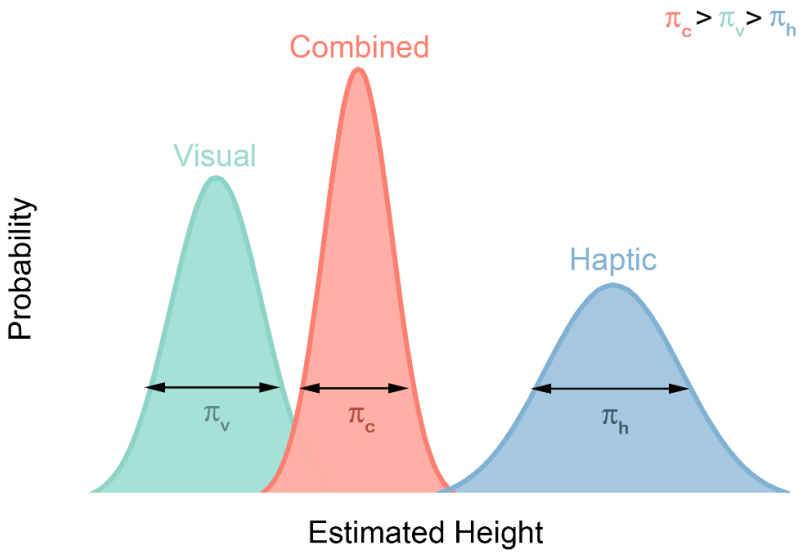


Figure 3. The above depicts how participants in the Ernst and Banks (2002) study combined evidence from multiple sensory origins. The task involved participants judging the vertical measurement of a bar using uncertain visual and haptic sensory input (uncertainty given by precision, $\pi = 1/\sigma^2$). As shown by the probability distributions above, posterior estimates of a bar position are a compromise between each sensory information source, and the precision of the posterior is an assimilated summation of each source of evidence. By increasing the relative precision of either source of information, one can demonstrate how the posterior density estimate is pulled toward the mean of the more precise source. Importantly, the Ernst and Banks (2002) study demonstrated how well individuals' behaviour in this visual and haptic feedback task was described by a Bayesian integration over distinct sensory input. But also, of particular relevance to this thesis, how important precision is for influencing posterior estimates—more on this to come. (Figure adapted with permission from Ernst and Banks [2002], Copyright © 2002, Macmillan Magazines Ltd).

The importance of this framework is that it is a computationally optimal and conceptually efficient tool to perform continuous learning about uncertain states, as we are not throwing any information away (Rushworth et al., 2009). Practically, however, complete integration over the normalising constant (the denominator in **Eq. 6** representing model evidence) is intractable, as in most interesting cases the denominator has to evaluate over the probability for all possible values (Blokpoel et al., 2012; Da Costa et al., 2020; Friston, 2008; Kwisthout & van Rooij, 2013; Mathys et al., 2011, 2014). Some compelling evidence however comes from a non-human primate study where neurons were shown to reason probabilistically (Yang & Shadlen, 2007). In that study, Yang and Shadlen (2007) trained two monkeys to select

among pairs of coloured targets after seeing four shapes. These were displayed in a series. The task was to track the dynamic probabilistic reward by summing the weights associated with the four shapes. The authors showed that the monkeys assimilated the probabilistic shape combinations, also reporting single-cell recordings from the parietal cortex as representing the addition and subtraction of those probabilistic quantities (Yang & Shadlen, 2007). This implies that neurons can track evidence and accumulate evidence over time, calculating the probability of binary outcomes.

In some learning scenarios such as multi-armed bandit tasks with multiple stimulus attributes where the relevant attribute denoting reward changes, it has been shown that human participants use alternative strategies to Bayes-optimal solutions as these are too computationally demanding (Gershman et al., 2010; O'Reilly et al., 2012; Wilson & Niv, 2011). Whereas a similar but simpler binary version elicits full Bayesian learning (Wunderlich, Beierholm, et al., 2011). As will be explored in the subsequent **Section 1.4.6** on the Hierarchical Gaussian Filter model, approximations using PEs and precision-weighted PEs (pwPEs) afford sequential belief updating—thereby reducing the complexity of full Bayesian inference to a simple algorithm that can be implemented by neurons (Whittington & Bogacz, 2017). Thus a hierarchical predictive coding account represents a computationally feasible implementation of Bayesian inference and learning, with multiple biologically plausible alternatives to approximate computation of the posterior distribution (see da Costa et al., 2020; Mathys et al., 2011; Parr et al., 2019).

Until more recently, this mechanistic account of improbability driven ('surprisal'; or, self-information, see Jaynes, 1957) learning of the states of environments has led to reward-learning research being governed by reinforcement learning algorithms (Fletcher et al., 2001; McClure et al., 2003; O'Doherty et al., 2003; Pessiglione et al., 2006; Wunderlich, Symmonds, et al., 2011), while Bayesian approaches have seen more traction in other fields of work. This ranges from Ernst and Banks' (2002) work on visual and haptic feedback to research showing how humans localise a point in space using noise sources from both visual and auditory input (Battaglia et al., 2003), and beyond to sensorimotor learning where movements are guided by combining information about the task and sensory feedback—all using Bayesian operations (Körding & Wolpert, 2004). The overall implication is that accounts of reward-based learning in the brain without probabilistic estimates and uncertainty are too basic. More recent work has formally described dopamine activity using the precision (uncertainty) of beliefs about policies (FitzGerald, Dolan, et al., 2015; Friston et al., 2014), while others have addressed the limitation using hierarchical Bayesian models, which we will unpack in further detail in the later section on learning from volatile environments (**Section 1.4.4**, Aitchison & Lengyel, 2017;

Behrens et al., 2007; Courville et al., 2006; Dayan et al., 1995; de Berker et al., 2016; den Ouden et al., 2012; Doya et al., 2007; Iglesias et al., 2013; Roesch et al., 2012; Yu & Dayan, 2005).

1.4.2 Bayesian learning

The usefulness of Bayes' Theorem should, in light of evolutionary selection, result in behavioural patterns and neural processing that reach some statistical optimum (Diaconescu et al., 2014). That is the model propounded in the "Bayesian brain" hypothesis (Doya et al., 2007; Friston, 2010; Knill & Pouget, 2004; Körding, 2007; Tenenbaum et al., 2011). Accordingly, this indicates that the brain is continually learning about the causes of and changes to its sensory environment. An increasingly accurate generative model of sensory experiences then affords inference on multiple hierarchically related hidden environmental states and the causes of the sensory input that generate those states (Diaconescu et al., 2014; Friston, 2003, 2005).

As was given above, Bayesian learning proceeds as the sequential use of Bayes' rule (posterior $\propto$ likelihood $\cdot$ prior) to continuously learn about the structure of uncertain states in a probabilistic fashion. In what now seems like a tradition for illustrative purposes, the below provides an example of this process from sport. The tradition is to use tennis (de Berker, 2017; Körding, 2007; Vernon, 2020). Yet I play the arguably more thrilling high-paced predecessor: *squash*. The Bayesian equation describing how to localise an opponent's shot, however, remains the same:

$$\mathrm{p}(\mathrm{SB\ at\ Y}|\mathrm{P\ of\ SB\ at\ Y}) \propto \mathrm{p}(\mathrm{P\ of\ SB\ at\ Y}|\mathrm{SB\ at\ Y})\ \mathrm{p}(\mathrm{SB\ at\ Y}) \tag{7}$$

Where the probability (p) of the perception (P) of the squash ball (SB) is given at the location *Y.* In the context of reward learning, the overarching desired outcome of winning in squash is a more elaborate pattern of motivational rewards. However, below we will focus on the narrower but no less difficult goal of predicting the most probable point in space and time that the squash ball will land. To optimally predict the localisation of the squash ball, we take the likelihood *p(P of SB at Y|SB at Y;* or, sensory input, and integrate it with the prior *p(SB at Y)*, giving us a distribution common to all squash shots (say, the serving shot).

Acquiring faithful priors can serve you well in any situation, but especially when facing the return shot in squash. The world record speed is 176 mph, so having a highly reliable distribution of typical serves as your prior can afford you extra time to achieve success

(Williams et al., 2020). A visual illustration of this Bayesian process of combining prior and likelihood information can be found below when trying to return a serve in squash (**Figure 4**).

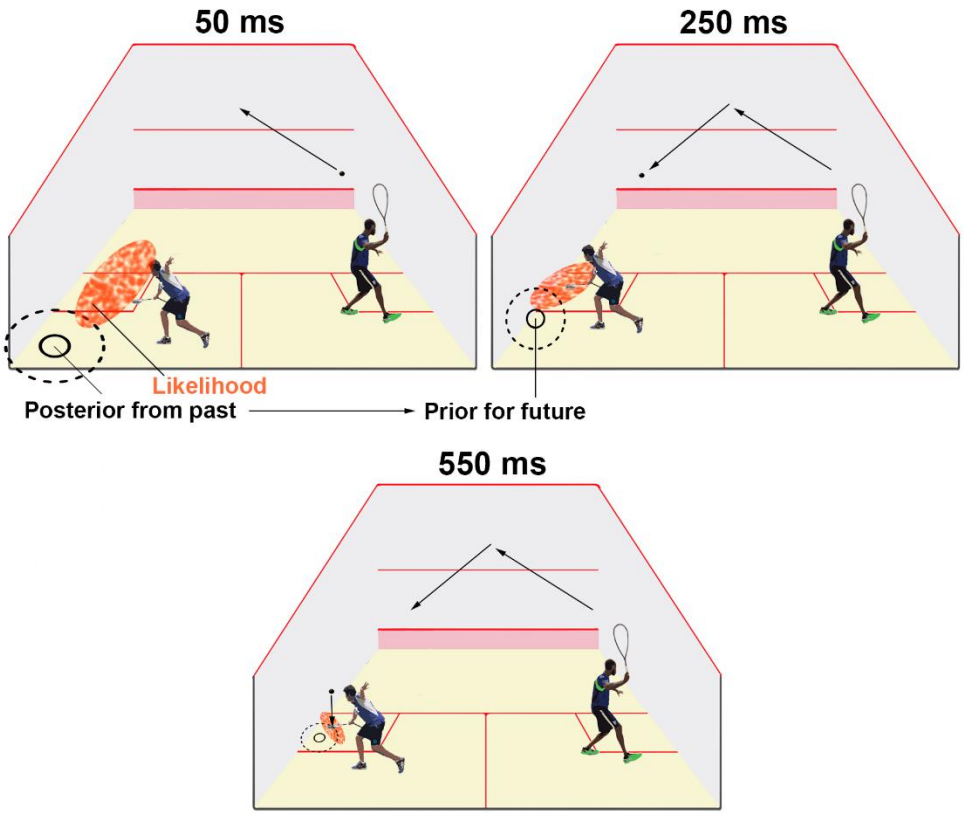


Figure 4. Bayesian integration of priors and likelihoods. A squash player (the left player) attempting to anticipate the most probable site of the ball landing from their opponent's shot (player right). The posterior estimate from the prior (past experience of all serves) is shown at 50 ms (top left panel) after the ball has been struck by the opponent. The likelihood space is given by the indeterminate red ellipse, inferred from the trajectory of the ball shown by a black arrow. By combining these two probability estimates together, weighted by a ratio of precisions, the player can arrive at a new prior for the future (see time at 250 ms in the top right panel). Accompanying the prior is an updated likelihood density provided by the sensory feedback of the ball's trajectory. Again, in a sequential step by step process, these sources of uncertain information are combined using Bayes' rule to arrive at a more precise estimate of the ball's landing position at time 550 ms (bottom panel). (Figure adapted from the original tennis example (Körding, 2007) using images found at Awesome Sports [2018]).

The expression *p(ball at x)* articulates the prior probability of a given landing spot from a serve. This changes based on the inferred movements of the server on approach to hitting the ball; but also, from prior knowledge from a pool of typical serve styles your opponent uses and how often they change this style in response to your actions (these changes to contingencies are to be discussed in further detail in the next **Section 1.4.4** concerning environmental volatility).

If a squash player has a lot of experience and has played many games, they will *a priori* have a good idea that some serves are more likely than others. In the event a new serve is experienced, it can be represented as a probability density function (O'Reilly et al., 2012). And like the above example of combining different sources of sensory information to make a new posterior belief, the prior and the likelihood are also assimilated as a function of their respective estimated reliability, or precision.

One final example from sensorimotor learning research gives compelling support for this Bayesian learning operation in humans (Körding & Wolpert, 2004). In that study, Körding and Wolpert (2004) asked participants to extend their right finger toward a target. Using virtual reality feedback allowed the experimenters to manipulate the degree of shift in visual feedback (laterally) from the participant's finger. The experimenters modulated the degree of displacement on each trial, taken from a Gaussian distribution, and also modulated the fidelity (high, low) of sensory feedback. The results provided good evidence that participants were learning the prior distribution over possible displacements and fine-tuning their sensorimotor performance by integrating prior distributions and their sensory feedback according to their estimated reliability.

1.4.3 Hierarchical predictive coding in the brain

Before moving on to how we purportedly process different forms of uncertainty and how we learn in changing environments, a short section here will tie all the above components concerning human learning—from RL and its associated algorithms to Bayesian inference—into one error-based learning account (Rushworth et al., 2009). This is useful as we need a good functional explanation of how the brain is implementing Bayesian inference in neural processing. Predictive coding (PC) is a process theory (Elias, 1955; Mumford, 1992; Rao & Ballard, 1999; Srinivasan et al., 1982) providing one algorithm in which a fixed-form approximation to Bayesian inference can work in neural processing by representing the difference between sensory input and prediction, as detailed above on TD learning and the dopamine RPE signal (Schultz et al., 1997). While Bayesian inference can be performed using alternative neural processes, like neural firing rates representing log-probabilities in population codes (Aitchison & Lengyel, 2017; Ma et al., 2006; Pouget et al., 2013) or direct variable coding (Olshausen & Field, 1996), PC—characterised by how surprising stimuli elicit higher neural responses than unsurprising ones—accounts for an impressive range of cognitive phenomena under certain assumptions (Huang & Rao, 2011; Spratling, 2017). For a PC account of approximate Bayesian inference, it is commonly assumed that the generative

model is hierarchical and the posterior is Gaussian and resolves hidden states (requiring approximation: Laplace/mean-field, see Mathys & Weber, 2020).

In what is possibly the earliest proposal of error-driven learning in the nervous system, Marr (1969) suggested that cerebellar plasticity (guiding action commission) is shaped by climbing fibre input. This was later substantiated by evidence of complex spikes in cerebellar Purkinje cells sharing similar properties with PE encoded signals, represented in the hypothesised climbing fibres during a visual-sensorimotor task (Kitazawa et al., 1998). Later, in the province of perception, came the work of Rao and Ballard (1999), who built upon Mumford's (1992) processing hierarchy in the brain, propounding a probabilistic hierarchical PC, discriminating between higher levels encoding predictions and lower sensory levels encoding ascending PE signals. The empirical support for this comes from circuit-based modelling of the visual receptive field. Extra-classical receptive field stimuli have been shown to lead to the relative inhibition of activity in V1 (Rao et al., 2016). Also, using a PC scheme by Kalman filtering and a hierarchical hidden Markov model to natural images, Rao and Ballard (1999) showed how cortical neurons demonstrating extra-classical receptive fields are error-detecting neurons.

Later Friston (2005) generalised PC by inferring beliefs and parameters from variational approximation to Bayesian inference. Subsequent implementation of this kind of Bayesian PC came from a rendition with multiple hierarchical levels termed the 'free-energy principle' (a special case wielding a particular type of dynamic probabilistic generative model and class of variational filtering, see Friston, 2005, 2010; Friston & Kiebel, 2009a). The free energy principle ostensibly shares many similar properties with the Rao and Ballard (1999) procedure. Both submit to a laminar specific hierarchical account vacillating between prediction and unexplained 'newsworthy' errors, the latter transmitted using inter-cortical feedforward connections. Yet the variables in the free energy model represent the statistics of the value signals, as opposed to the value of signals themselves, estimating posterior probability densities (Spratling, 2017). The proponents of the free energy principle describe active inference as the minimisation of variational free energy (equivalent to optimising the generative model, mirroring real-world statistics) and expected free energy (obtaining preferred outcomes and circumnavigating surprise, on average, see da Costa et al. [2020)] for more detail). To avoid becoming entrenched in the nuts and bolts here, it is sufficient to say that in using an active inference approach (the normative minimisation of surprise or misprediction [uncertainty] and variational free energy) we also implicitly maximise Bayesian model evidence in executing approximate Bayesian inference (Friston, Parr, et al., 2017; Sengupta et al., 2016; Sengupta & Friston, 2017). And this is precisely where active inference is coterminous with the Bayesian brain (Aitchison & Lengyel, 2017; Friston, 2012; Knill &

Pouget, 2004)—in *self-evidencing* (Hohwy, 2016) and resolving uncertainty (Friston, Parr, et al., 2017).

This free-energy formulation of Bayesian hierarchical PC has contributed to an overwhelming number of new predictions concerning 'surprisal' in the brain, tested using electrophysiological and fMRI data (Cole et al., 2020; Deserno et al., 2020; Diaconescu, Litvak, et al., 2017; Diaconescu, Mathys, et al., 2017; Diaconescu et al., 2019; Garrido et al., 2008; Henco et al., 2020; Iglesias et al., 2021; Kolossa et al., 2015; Liu et al., 2021; Maheu et al., 2019; Mars et al., 2008; Nassar, Bruckner, et al., 2019; Nassar, McGuire, et al., 2019; Powers et al., 2017; Stefanics et al., 2018; Summerfield et al., 2006; Weber et al., 2020). The Bayesian PC framework thus involves a generative model of the environment, optimised through sensory data (Friston, 2005, 2010; Mumford, 1992; Rao & Ballard, 1999; Srinivasan et al., 1982). The model is updated using empirical Bayes in functional hierarchical architecture (Friston, 2005; Knill & Pouget, 2004).

The cortical mechanisms and evidence for PC in the brain has been comprehensively reviewed in several sensational reviews (Bastos et al., 2012; Keller & Mrsic-Flogel, 2018; Shipp, 2016; Summerfield & de Lange, 2014). Below I merely touch upon the basic and necessary components of the requisite neural architecture to stimulate the research that comprises this thesis. PC theory utilises the hierarchical composition of the cortex to implement a hierarchical generative model of the world, transmitting predictions (backward) to suppress the processing of expected (forward) sensory signals (Clark, 2013a; den Ouden et al., 2009; Shipp, 2016). When predictions mismatch with input, a state PE signal is transmitted unimpeded up the cortical hierarchy to refine the internal model (predictions) of the system (Friston, Parr, et al., 2017; Friston & Kiebel, 2009b).

Predictions are thought to be sent down deeper cortical layers (layer 5: larger pyramidal cells) from empirical prior beliefs (layer 3: smaller pyramidal cells), while error signals are transmitted up to meet predictions (layer 4 spiny stellate cells to layer 3: Bastos et al., 2012; Friston, 2008; Haeusler & Maass, 2007; Harris & Shepherd, 2015; Lübke et al., 2000; Peters et al., 2017; Ramaswamy & Markram, 2015; Rao & Ballard, 1999; Shipp, 2016). This global policy of refining a generative model through measuring and reducing PEs is thought to dampen the activity of error residuals arising in supragranular layers in superficial pyramidal cells (Barone et al., 2000; Edwards et al., 2012; Friston & Kiebel, 2009a; Shipp, 2016). It has since become clearer that this interplay of prediction and error across multiple hierarchical levels reaches beyond early sensory processing to higher cognitive processes, including a diverse panoply of cognition, from social learning, decision-making, attentional processes, to emotion, mood,

and interoception (Daunizeau, den Ouden, Pessiglione, Kiebel, Stephan, et al., 2010; Diaconescu et al., 2014; Diaconescu, Mathys, et al., 2017; Hunt et al., 2014, 2017; Hunt & Hayden, 2017; Schwartenbeck, FitzGerald, Mathys, Dolan, Kronbichler, et al., 2015; Seth, 2013; Seth & Friston, 2016; Vossel et al., 2014).

Forward (predominantly excitatory) intrinsic connectivity and backward prediction-based (generally inhibitory) hierarchical message passing, also discussed in the context of PC is thought laminar-specific and realised using 'canonical microcircuits' mediated by distinct oscillatory waves (Arnal & Giraud, 2012; Bastos et al., 2012; Bastos, Litvak et al., 2015; Bastos, Vezoli et al., 2015; Douglas & Martin, 1991; Friston, Parr, et al., 2017; Gilbert, 1983; Haeusler & Maass, 2007; Heinzle et al., 2007; Thomson & Bannister, 2003). The spectral asymmetries of oscillatory patterns between descending predictions and ascending error discrepancy signals are discussed in great detail along with neural and physiological evidence in work by others (see Adams et al., 2013; Bastos et al., 2012; Bastos et al., 2015, 2018, 2020; Friston, 2008; Mumford, 1992; Sedley et al., 2016; Shipp, 2005, 2016; Shipp et al., 2013). Importantly, in **Chapters 3 and 5** we aim to test the impact of anxiety on the neural oscillations of the PC process, discussed in more detail in **Section 1.5**.

It is sufficient for this chapter to take a flying review of the evidence from human electro/magnetoencephalography (MEG/EEG) and monkey electrocorticography. Studies suggest that feedforward PE signals are encoded by faster gamma oscillations (>30 Hz), while backward descending predictions are expressed in lower alpha (8–12 Hz) and beta-band (13–30 Hz) oscillations (Alamia & Van Rullen, 2019; Arnal & Giraud, 2012; Auksztulewicz & Friston, 2016; Bastos et al., 2012; Bauer et al., 2006; Brodski et al., 2015; Cao et al., 2017; Friston, 2005; Mayer et al., 2016; Pinotsis et al., 2016; Van Kerkoerle et al., 2014; van Pelt et al., 2016; Wang, 2010). For PC, alpha/beta-band activity is usually associated with afferent inhibitory effects that gates sensory processing (Arnal & Giraud, 2012; Jensen et al., 2012; Jensen & Mazaheri, 2010; Sedley et al., 2016; van Ede et al., 2011). While gamma-band activity (predominant in superficial layers) mediates the propagation of feedforward PE signals (Arnal & Giraud, 2012; Bastos, Litvak et al., 2015; Bauer et al., 2014; Engel et al., 2001; Michalareas et al., 2016; Sedley et al., 2016; Todorovic et al., 2011; van Pelt et al., 2016; Wang, 2010). This is most evident in visual cortex studies where asymmetry is shown between alpha/beta-band synchronisation in infragranular layers and gamma-band in supragranular layers, with alpha/beta functionally inhibiting the processing of sensory input spiking, suppressing gamma oscillations (Arnal & Giraud, 2012; Bastos et al., 2012; Bastos, Litvak et al., 2015; Buffalo et al., 2011; Gould et al., 2011; Michalareas et al., 2016; Roberts et al., 2013; Xing et al., 2012). Whether these rhythmic attributes in the PC framework represent one neural process by which

affective states like anxiety can influence belief updating and shape learning, however, remains unknown. This will be unpacked and demystified in **Chapters 3 and 5** and in the pages that follow.

A further recent nuance to the PC scheme has been proposed which specifies the lack of unique functional circuits for the computation and transmission of PEs (Bastos et al., 2020). Instead, known inter-cortical connections for distinct processing streams (sensory, attention, working memory and so on) operate the predictive functions outlined above. Among PC and neural oscillation papers, there is typically no consistent distinction made between alpha and beta rhythms as characterising predictions (Auksztulewicz et al., 2017; Bastos, Litvak et al., 2015; Bauer et al., 2014; Buffalo et al., 2011; Friston, 2019; Roberts et al., 2013; Xing et al., 2012). But in the more recent predictive routing, Bastos and colleagues test alpha and beta independently to understand how they transmit top-down signals that suppress the expression of gamma rhythms that aid in the spiking of sensory signals (Bastos et al., 2020). As such, the expression of alpha and beta oscillations encoding predictions that dampen PE responses has now been widely shown across multiple modalities, from visual (Gould et al., 2011; Rohenkohl et al., 2012) and motor (Palmer et al., 2019; Schoffelen et al., 2005), to somatosensory (van Ede et al., 2011) and auditory (Sedley et al., 2016; Todorovic et al., 2015) domains. Yet as of the time of writing this thesis, no studies focusing on the neural oscillations of PC in humans while learning from reward have been reported. We aimed to address this in **Chapter 3** using EEG in both a state anxious and control group, and in **Chapter 5** using MEG in a high trait anxious and low trait anxious group. We aim to show how anxiety impacts reward-based learning and moreover the learning signals used during Bayesian PC belief updating in the brain.

The role of uncertainty estimates in predictive learning within dynamic environments can be observed in the weighting of PE signals according to the estimated reliability of both predictions and incoming sensory events (Bauer et al., 2014; Feldman & Friston, 2010; Kok et al., 2012; O'Reilly, 2013; Payzan-LeNestour & Bossaerts, 2011). This estimate of uncertainty regulates the influence of PE signals on belief updating by modulating the gain of neuronal populations encoding PEs (Feldman & Friston, 2010; Knill & Pouget, 2004; Kok et al., 2012). These are termed precision-weighted PEs (pwPEs). In this way, the relative precision weights serve to enhance or suppress the impact of PEs on updates to posterior beliefs (Feldman & Friston, 2010; Kanai et al., 2015). More plainly stated, the greater the precision, the more influence on belief updates (Peters et al., 2017). Precision is thought to be neurobiologically implemented by neuromodulators such as dopamine and acetylcholine, which we expand upon below in **Section 1.4.8.4**.

In this thesis, we address the question of how pwPEs are represented in both the time (**Chapter 2**) and frequency domain (**Chapters 3 and 5**). Critically, as discussed above, more precise posterior beliefs (predictions) may be associated with a preponderance of alpha and beta waves that suppress forward ascending error signals relayed through gamma waves (Bauer et al., 2014). This equates to slower updates to beliefs. By contrast, when we have little prior knowledge of state statistics or an unexpected change to the environment has recently occurred (or sensory informational uncertainty is high) an increase in gamma waves is anticipated with error signals travelling forward to update beliefs (Arnal & Giraud, 2012; Auksztulewicz et al., 2017; Fries et al., 2002; Womelsdorf et al., 2006). To the authors' knowledge, scant evidence exists on a spectral correlation to modulatory precision. Although, work by Sedley (2016) and later Palmer et al. (2019) suggests precision weights are modulated in alpha oscillations.

To summarise, in a hierarchical PC framework, if we can minimise PEs then we have a sufficient explanation or probabilistic belief about the causes of our sensations. The precision (uncertainty) of these beliefs are vitally important for Bayesian accounts of PC. Precision in this Bayesian PC framework acts as a kind of transistor, a synaptic gain control mechanism modulating the impact of error signals, equivalent to attention (Eldar et al., 2013; Feldman & Friston, 2010). Importantly, precision itself needs to be estimated and optimised to minimise the resultant pwPEs (Parr & Friston, 2018). At the neural transmitter level, several neuromodulators are putatively involved in encoding precision. The cholinergic system through acetylcholine receptors is thought responsible for changes to the precision of the likelihood, as it is known to contribute toward modulating the gain of sensory evoked responses (Disney et al., 2007; Gil et al., 1997; Parr & Friston, 2018). This is further substantiated by evidence of alterations to effective connectivity and behavioural modelling using pharmacological manipulation (Marshall et al., 2016; Moran et al., 2013; Vossel et al., 2014). In a captivating study, acetylcholine was shown to balance the assignment of uncertainty to either noise or environmental changes (Marshall et al., 2016).

Dopamine has also been connected to the encoding of precision, specifically concerning policies, with post-synaptic connections terminating on striatal medium spiny neurons (Freund et al., 1984; Yager et al., 2015). Later fMRI neuroimaging work supplied further evidence of the role dopamine plays in determining precision changes to policies (Schwartenbeck, FitzGerald, Mathys, Dolan, & Friston, 2015), again bolstered by pharmacological manipulation experiments that modelled behaviour (Marshall et al., 2016). Lastly, noradrenaline has also been implicated in the encoding of surprise (low precision input) and transitions (volatility)

through pupil dilation and associated activity in the locus coeruleus (Aston-Jones & Cohen, 2005; Eldar et al., 2013; Lavine et al., 1997; Liao et al., 2016; Parr & Friston, 2018; Vincent et al., 2019).

At the electrophysiological level, precision has been associated with different hierarchically related PEs in a distinct temporal and spatial hierarchy (Diaconescu, Litvak, et al., 2017). The auditory MMN event-related response has consistently been used as evidence for statistical learning of the latent structure of the sensory world (Näätänen et al., 2005; Paavilainen et al., 1999). Predictable trains of stimuli inhibit the MMN while surprising events elicit a large electrophysiological response. This, it is thought, represents model updating across the auditory hierarchy (Garrido et al., 2009; Lieder et al., 2013; Weber et al., 2020; Winkler, 2007). Importantly, several difference waveforms known to ERP research, such as the MMN and error-related negativity (ERN), are interpreted as the difference in pwPE signal between predicted and unpredicted stimuli (Holroyd et al., 2003; Lumaca & Trusbak Haumann, 2019; Sambrook & Goslin, 2015; Weber et al., 2020; Yasuda et al., 2004). Whether it is in visual perception (Brown & Friston, 2012), sensorimotor learning (Palmer et al., 2019), visual or auditory mismatch (Lumaca & Trusbak Haumann, 2019; Weber et al., 2018), or reward-learning tasks (Bellebaum & Daum, 2008; Holroyd et al., 2003; Philiastides et al., 2010; Yasuda et al., 2004). A combined EEG and fMRI study tracking the spatial and temporal sequence of distinct PE signals and their associated precision showed a mirroring of the predicted dynamics and order of hierarchical Bayesian belief computations, with lower-level PEs and precision occurring earlier and across sensory regions, and higher-level PEs and precision represented later and in more frontocentral regions (Diaconescu, Litvak, et al., 2017). Exhilarating advances in EEG studies have also led to researchers distinguishing among lower-level pwPE signals (interpreted as responsible for updates in belief estimates, see Stefanics et al., 2018) and higher-level ones that contribute toward belief computations concerning the rate of change in our sensory environment (see Weber et al., 2020). Crucially, these pwPE signals are thought to be the primary contributor to changes in EEG signals (Friston & Kiebel, 2009a). In the next section, we will learn of the importance of pwPEs and how they have been tracked in tasks where there are multiple sources of uncertainty in dynamically changing statistical environments.

1.4.4 Changing environments and Bayesian adaptation

Until this point, we have been discussing how Bayesian inference in complex environments works, both conceptually and computationally, and within the brain. Yet complex environments are dynamic, unpredictably changing to unknown states (O'Reilly, 2013). These demand

Bayesian computations be calculated continually in real-time, adapting to unforeseen changes to the statistical structure of an environment. The computational burden of which is only intensified when we need to identify and learn which variables are task-relevant (and consequently, participants during experimental tasks may represent irrelevant variables in the Bayesian joint distribution, see Gershman et al., 2010; Gottlieb et al., 2013; O'Reilly, 2013; Wilson & Niv, 2011). Importantly, learning agents in changing environments need to detect both alterations to the probabilistic contingencies which govern predictable outcomes, but also inhibit extemporaneous adaptations to anomalous outcomes brought on by noise (Moens & Zénon, 2019). That is to say, in a reward learning context, agents need to alter flexibly in response to environmental changes to reward contingencies, and that an infrequent absence of anticipated reward does not necessarily indicate a need to switch behavioural policies to an alternative (Daw & O'Doherty, 2014; Dayan et al., 2000; Dayan & Long, 1998; Inglis et al., 2020; Kennerley et al., 2006). From a Bayesian perspective, this means estimating the probability that the present observation was sampled from the same state as prior observations and estimating the likelihood of change points (O'Reilly, 2013; Wilson et al., 2010).

As an everyday example, take the daily commute. Each traveller on encountering new traffic or delays needs to learn the origin of delay to adapt accordingly. Is the unexpected traffic a manifestation of a quick and temporary event such as a road incident? Or is it a symptom of a long term road restoration project which could interfere with commuting for months? Both examples represent changes to the environment. However, a road incident comes from incidental happenstance, leading to a degree of *expected* uncertainty (relevant uncertainty anticipated by our generative model), and therefore does not demand adaptation. While the second represents consequential environmental instability, which may lead to *unexpected* uncertainty used to update beliefs about the world and adapt behavioural responses—more on the different types of uncertainty and their neural underpinnings in due course (Soltani & Izquierdo, 2019). Crucially, learning algorithms need to infer the rate of change to perform optimally. As mentioned above, we can start to approach this by estimating the likelihood that prior and present outcomes are from an identical distribution, and by estimating the prior probability of change (O'Reilly, 2013). The latter restricts how much new data is needed to infer a meaningful change has happened.

That said, inferring the transition function of an environment is also a vital computation that remains a lively area of debate (Courville et al., 2006; Mathys et al., 2011; Moens & Zénon, 2019; O'Reilly, 2013; Soltani & Izquierdo, 2019; Wilson et al., 2010). As will be discussed in later chapters, different a priori beliefs of the probability of change can markedly impact the

estimation of change points and environmental uncertainty. The transition function depends on the true parameters of previous trials. In Gaussian examples of transition probabilities, the change is smooth and diffuse. For environments where the parameters ratchet up or down, the environment can best be described using sudden change points (O'Reilly, 2013). Consequently, the transition function (for example, diffuse or sudden changes) and its parameters (for example change-point probability and the rate of transitions) are both estimated to update beliefs about whether a meaningful alteration to the probabilistic environment has taken place.

The above implies how vital estimating distinct types of (interacting) uncertainty is to executing adaptive actions. Numerous studies have manipulated outcome probabilities or contingencies to test learning in uncertain environments, typically using probabilities approaching 0.5 as more uncertain environments (Bland & Schaefer, 2012; Cohen et al., 2007; Huettel et al., 2005; Krain et al., 2006; Paulus et al., 2004; Polezzi et al., 2008; Volz et al., 2003). However, more recent work has modelled reward-learning behaviour in response to unexpected changes in the contingency mappings (Behrens et al., 2007, 2008; Browning et al., 2015; Krishnamurthy et al., 2017; Nassar et al., 2012; Nassar, McGuire, et al., 2019; Pulcu & Browning, 2017).

Consider one pivotal study by Behrens et al. (2007) that tested this process using a Bayesian formulation of reinforcement learning. For participants tracking the probability of a rewarding outcome associated with either of two images (blue or green) in each trial *i*, the prior probability distribution—updated continuously throughout the experiment—creates a prediction of the expected reward associated with each stimulus. The parameter that Behrens et al. (2007) modelled was thus the reward probability r_i, bound to specific selections. The model was Bayes optimal for learning the reward probability between trials given changes to the rate of change in the reward parameter, governed by volatility, *v*. This proceeded across a stable (low volatility) block where the probability of reward was set at 0.75, to a volatile block (high volatility) where reward probability alternated between 0.8 for blue and 0.8 for green every 35 (±5) trials. This study demonstrated how humans utilised Bayesian learning, evaluating the relevance of each new input, even in a dynamic environment. Highly volatile environments increase the rate and reliability of the integration of new information against prior beliefs. While stable environments benefit from relying on prior beliefs. Consequently, Behren's et al. (2007) showed how rapidly changing environments can make prior beliefs obsolete in favour of new evidence. Happily, these changes to contingencies are readily grappled by a Bayesian learner in using probabilistic belief distributions over potential values with estimated uncertainty.

Although, as was done in Behrens et al. (2007), an emphasis on the utility of describing behavioural responses is useful in respect to learning rates, as is discussed below.

The learning rate describes the rate of adaptation in response to unpredicted outcomes and their influence on updating subsequent beliefs. It is governed by the uncertainty associated with the estimate of reward, which mirrors the degree of predictability in recent outcomes (Behrens et al., 2007; Dayan et al., 2000). If the series of recent outcomes were all predictable, one would expect a decrease in estimated volatility and the reliability of that estimate, and consequently, this will decrease the learning rate. By contrast, if we are in the throes of a sequence of strikingly unpredictable outcomes, then volatility estimates may increase and so too does the learning rate. This also translates upon the reward rate. When little about the reward environment is changing, the estimated reward rate likewise does not change so there is little to learn. By contrast, when all recent experience from the reward environment is changing, the reward rate will fluctuate accordingly with each new observation, meaning we need to increase our learning. Thus, learning rates also fluctuate in volatile environments where the rate of change is fixed but where probabilistic relationships change (see **Chapters 2–5** below but also Iglesias et al. (2013) and de Berker et al. (2016).

The learning rate then determines the relative influence of past to recent information on belief updates (Jepma et al., 2016). Behrens and colleagues (2007) established that this pattern of Bayesian learning, using probabilistic estimates to successfully monitor the statistics of the reward environment and adapting the learning rate to match these changes, was a better predictor of participants' behavioural responses when compared to a delta rule RL model with a fixed learning rate. The insight from Behrens and colleagues (2007) is in revealing how a Bayesian reinforcement learning model can afford dynamic volatility-based adjustments to the learning rate, and moreover the association of the learning rate to volatility estimates and where in the brain this is processed (see **Figure 5**). However, it is important to note that this influenced later work to translate the learning rate to the precision estimate that weights PE signals in pwPEs, as we will see in **Section (1.4.6)** on the Hierarchical Gaussian Filter model (Mathys et al., 2011). But this normative account of learning in adapting the learning rate to environmental change (higher for volatile periods) is consistently observed across tasks using reward (Behrens et al., 2007, 2008; Huang et al., 2017; Nassar et al., 2012) and aversive outcomes (Browning et al., 2015; Pulcu & Browning, 2017).

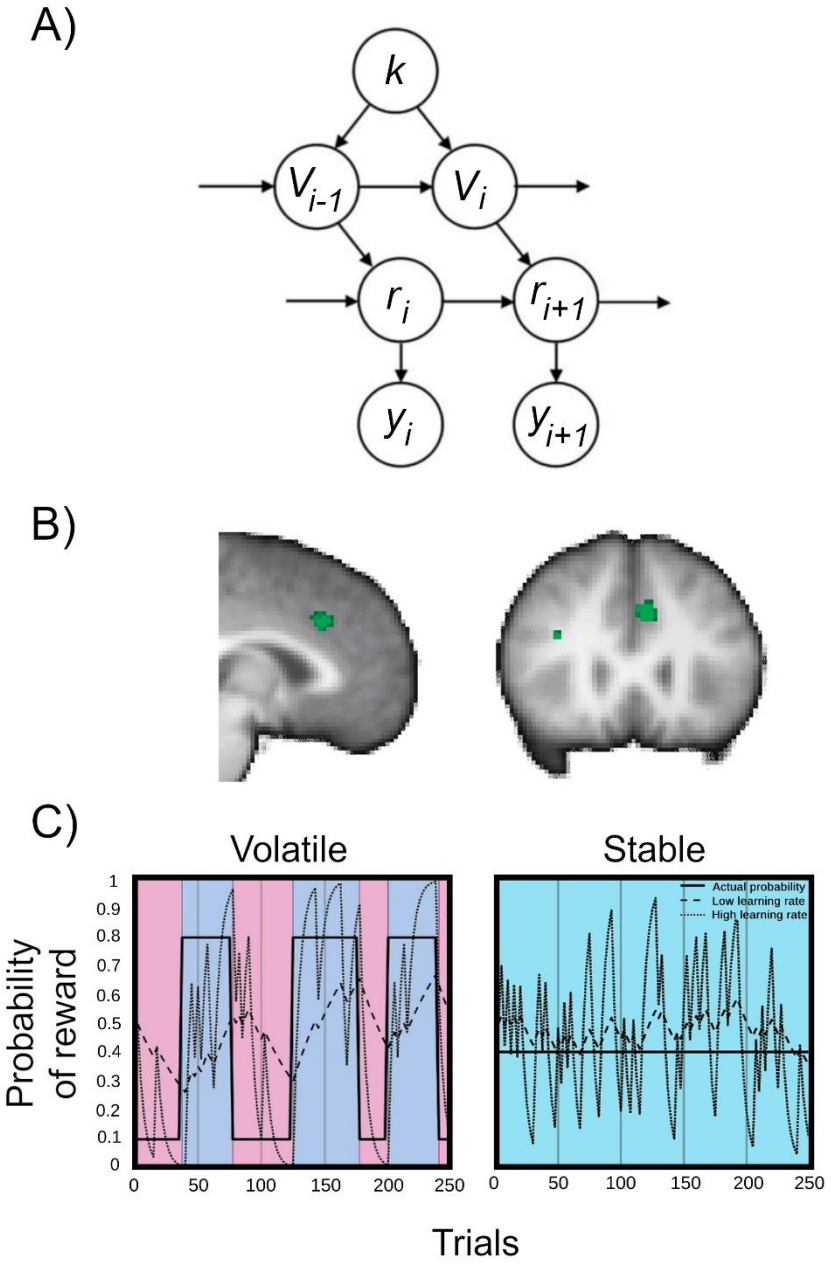


Figure 5. Volatility and the learning rate in a Bayesian model of learning. In the above **A)** is an illustration of the model from Behrens et al. (2007) detailing how reward probability is estimated under changes to contingencies (volatility, V). At the bottom layer, the observed input data y from probability r (that may change across trials i and i+1). The trajectory of change is governed by the arrows (black), and V regulates the rate of change. In the Behrens et al. (2007) study, V also changed (by parameter k). The reward probability distribution for the next trial ($r_{(i+1)}$) is conditional upon the current estimate of reward and the estimate of volatility $V_{(i)}$. This means that as our estimate of $V_{(i)}$ increases, the variance also increases, where $r_{(i)}$ may take other possible values more quickly. Conversely, as $V_{(i)}$ decreases the precision (variance) increases, and thus $r_{(i)}$ is more stable. **B)** These fMRI images from Behrens et al. (2007) show BOLD activity response to reward outcomes in the ACC associated with trial-wise changes to volatility estimates, V. **C)** The impact volatility has on the learning rate. In the left panel is a high volatility block of 250 trials where the reward probability changes from 0.1 to 0.8 (ground truth probability). In the right panel is a stable block (250 trials) where the probability of reward does not change from 0.4 (actual probability). The learning rate when high (----) fits the rapidly changing environment in the volatile environment, but in the stable block the higher learning rate treats infrequent

surprising events as meaningful, thus reacting disproportionately and not reaching a secure estimate. In the stable block, a low learning rate (– –) treats unexpected outcomes as chance events, and thus settles on a better estimate; however, in the volatile environment, the slow learning rate fails to adapt fast enough to form accurate estimates. Panels A and B are adapted with permission from Behrens et al. (2007). Copyright © 2007, Nature Publishing Group. Panel C is adapted from the open access article by Pulcu and Browning (2019). Copyright © 1969, Elsevier, see https://creativecommons.org/licenses/.

The neural underpinnings of reinforcement learning rates (that is to say, fixed learning rates as in RW/TD learning, Barto et al., 1989; Rescorla & Wagner, 1972) has been fit to fMRI data showing the VS and OFC are involved with error and value signals (Hare et al., 2008; O'Doherty et al., 2003). Others have taken up identifying which neural assemblies are linked with dynamic learning rates and belief updating. Some work has indicated that dopaminergic midbrain PEs exhibit a similar scale as the learning rate (Diederen et al., 2016). Several studies have associated volatility and learning rate adaptation to ACC and dorsomedial prefrontal cortex (dmPFC) activity, uncertainty based belief updating to anterior PFC and parietal regions, reward-driven learning to the VS, and the connectivity between the striatum and PFC in the dopaminergic system (Behrens et al., 2007, 2008; Chumbley et al., 2012; Courville et al., 2006; Jocham et al., 2009; Krugel et al., 2009; McGuire et al., 2014; Muller et al., 2019; Nassar et al., 2012; Nassar, McGuire, et al., 2019; O'Reilly et al., 2013; Preuschoff & Bossaerts, 2007; Yu & Dayan, 2005). Also, different authors have pointed to the link between the amygdala and associative learning (Averbeck & Costa, 2017; Phelps et al., 2014). Based on human neuroimaging, non-human primate single-cell recording, and lesion experiments on rats, studies indicate the amygdala is involved in controlling learning rates, or uncertainty and volatility more generally (Averbeck & Costa, 2017; Costa et al., 2016; Holland & Gallagher, 1999; Homan et al., 2019; Li et al., 2011; Roesch et al., 2012). The links between the amygdala and learning will be fleshed out in relation to anxiety in the following **Section 1.5.3**. Single-cell recording studies from primates have demonstrated the involvement of the PFC and dmPFC during probabilistic reward learning (Massi et al., 2018). In humans, however, this is still an unresolved issue (Soltani & Izquierdo, 2019).

Until recently, a more precise understanding of the neuromodulators involved in the encoding of the learning rate also remained largely unknown (Farashahi et al., 2017; Iigaya, 2016). Pharmacological modulation in humans with electrophysiological recording and computational modelling has identified catecholaminergic (noradrenaline and dopamine) involvement in regulating the learning rate (Franklin & Frank, 2015; Jepma et al., 2016). In that study, Jepma and colleagues (2016) used a predictive-inference task and the P3 component from the EEG signal as an index of outcome-evoked phasic catecholamine release in the cortex. They

showed that the P3 ERP (using a single-trial analysis approach) regulated the learning rate proportional to the PE, which they interpreted as a catecholamine driven change in learning as a function of surprising misprediction (Jepma et al., 2016). This means, noradrenaline and dopamine influence the learning rate after unsignalled task changes, but not during periods of low volatility. More details on the broader role of the noradrenaline system in learning from uncertainty will be saved for **Section 1.4.7**: the neural representations of different forms of uncertainty.

To summarise, the learning rate is proposed to change optimally with our estimates of different forms of uncertainty, notably our uncertainty over how much our learning environment is changing over time. And further to this, and in ways that aren't widely appreciated, little is known about how this is accomplished neurally, and the processing of alterations to learning rates and interactions with different forms of uncertainty also remains largely unknown. For now, it is important to take away that the rate in which we learn about certain quantities can vary depending on our uncertainty. Now we will turn to the various classes of uncertainty and evidence of human learners using these estimates.

1.4.5 Classifying uncertainty

Distinct forms of uncertainty can modulate the learning rate, and subsequent work has attempted to understand how the brain processes this unique array of uncertainty in the environment (Yu & Dayan, 2005). For instance, what if the outcome of some event is probabilistic (*expected* uncertainty, Soltani & Izquierdo, 2019; Yu & Dayan, 2005)—as is the case with most interesting events in the world. We may currently know very little about that event and its underlying probabilistic relationships (*estimation* or *informational* uncertainty, de Berker et al., 2016). The contingency relationships that govern those probabilities may even change over time (*unexpected* or *environmental* uncertainty, Bland & Schaefer, 2012; O'Reilly, 2013)—subjectively experienced *volatility* (Soltani & Izquierdo, 2019). To illustrate, an everyday example of volatility is an atypically cold day in an otherwise bright and warm season. Should one readily adapt to this new information about the cold, inferring a change to the climatic environment and so wearing thermal clothing and a jumper the following day? Or did this atypical drop in temperature represent noise in the environment? In this alternative framing, the learner might infer that the atypical climatic event did not indicate environmental change but rather unpredictability in temperature, deciding to wear a tee-shirt. How we distinguish between these different types of uncertainty is still an ongoing problem in the computational neuroscience of statistical inference (Piray & Daw, 2020a, 2020b; Pulcu & Browning, 2019).

Digging a little deeper, *expected* uncertainty also relates to the previously mentioned complexity of establishing the probability that the state of the environment has changed, either by some form of real change or by determining that the state was not as estimated beforehand (not predicted by our generative model, relating to a transition function, O'Reilly, 2013). Thus *expected* uncertainty informs change point detection by using the discrepancy between the new outcome and the expected value, but also the variance of the distribution the expected value is taken from (Dayan & Yu, 2003; Yu & Dayan, 2005).

Mercifully, we are capable of learning and remembering these statistical associations across time by repeatedly observing the outcomes of events. However, *estimation* uncertainty increases as a function of changes to environmental states, as we would need to relearn the contingencies of any new state. Almost instinctively we would expect a learning agent uncertain of the state of the world to be more flexible and disposed to updating their beliefs when encountering new information. Moreover, a-priori beliefs about volatility will influence *estimation* uncertainty; if volatility estimates are high, then predictions using prior beliefs will be less precise and we have more *estimation* uncertainty (O'Reilly, 2013). *Estimation* uncertainty also increases if the observer estimates they have encountered a change point, a surprising outcome indicating a change to the probabilistic structure of the environment, as opposed to a probabilistic fluctuation whose origin is likely *expected* uncertainty (Nassar et al., 2010; Wilson et al., 2010).

Unexpected uncertainty arises from subjectively perceived unexpected changes to learned contingencies (the true value of change is called *volatility*; however, the two can be distinguished by i) *unexpected* uncertainty is a subjective estimate of *volatility* and ii) the relative frequency of change, see Bland and Schaefer [2012]). Infrequent changes to contingencies may indicate change points that abruptly alter the learned environmental statistics, and thus our capacity to accurately predict upcoming outcomes (Nassar et al., 2010; Wilson et al., 2010). High *volatility* reward environments are thus characterised by frequently changing reward contingencies (Bland & Schaefer, 2012; Courville et al., 2006). Low *volatility* reward environments are punctuated with infrequent violations to expectations about reward outcome contingencies that require little adaptation in learning, primarily signalling *expected* uncertainty (irreducible noise). Low volatility environments can also feature meaningful changes (*unexpected* uncertainty; or, change points) that do demand adaptation in order to learn the reward statistics optimally. Consequently, one can define *volatility* as the probability of a change point during some time frame and *expected* uncertainty as subjectively estimated volatility (Pulcu & Browning, 2019).

On top of all that, more recent work highlights the additional complexity of detecting and distinguishing (and processing) *volatility* from outcome noise (*unpredictability*, see Piray & Daw, 2020b). The subtlety is that learning rates are codetermined by teasing apart *volatility* from *unpredictability*—viz. state stochasticity compared to a systematic change (Piray & Daw, 2020b). Piray and Daw (2020b) highlight that this brings to the fore the inconsistent psychological theories of conditioning from Mackintosh (1974) and Pearce and Hall (1980). These theories claim that organisms allocate either increasing or decreasing attentional resources to stimuli that consistently predict outcomes, respectively (Piray & Daw, 2020b). The logic of these theories follows that if a train of cues is reliably predictable, then they are either highly salient (more attention) or already explained away (less attention). It is possible then that *volatility* uncertainty could be misinterpreted as noise and impair task accuracy and learning rates. An alternative means to solve this problem is to assess if a surprising outcome is useful in predicting future states (Nour et al., 2018; O'Reilly et al., 2013; Pulcu & Browning, 2019). Interestingly, humans involved in decision-making tasks under uncertainty tend to make several choices that do not maximize expected reward but rather lower uncertainty about recently unchosen options (Daw et al., 2006; Findling & Wyart, 2021). More recently, Findling et al. (2019) examined computation noise abatement in reinforcement learning that supplies decision variability while humans performed a random-walk alternative to a traditional reversal learning task. The authors found that computation noise induces a significant portion of "non-greedy" choices that would have been allotted to navigating the exploration/exploitation trade-off. Moreover, Findling et al. (2019) distinguish the neural correlates of this process in the ACC and phasic pupil dilation (Findling & Wyart, 2021). Taken collectively, this research suggests that human exploration can stem from environmental uncertainty and estimation uncertainty.

The complexity of estimating these unique forms of uncertainty often leaves our beliefs about the environment either incomplete, imprecise, or outdated; or, in less favourable circumstances, all three. This type of misestimation of uncertainty has more recently been associated with affective disorders and the computational psychiatry literature, to which we will turn in **Chapter 3–5** and in the following **Section 1.4.8** and **Section 1.5** concerning precision pathologies and anxiety, respectively (Pulcu & Browning, 2019). Importantly, to understand how we estimate and process these types of uncertainties, normative behavioural models like the Hierarchical Gaussian Filter (that prescribe how learners need to adjust to uncertainty) explicitly estimate these core types of uncertainty into a hierarchical Bayesian predictive model that is additionally useful in seizing upon neural activity associated with model estimates.

1.4.6 The Hierarchical Gaussian Filter (HGF)

To act adaptively to change, we need to infer upon unobservable and dynamic parts of our environments. One way of modelling these hidden states is to construct a Bayesian filter based on an assumed generative process and apply it to a time series of sufficient statistics (mean, variances) as with the Hierarchical Gaussian Filter (HGF, see Mathys et al., 2011). This sequential mean tracking effectively results in approximate Bayesian inference by the transference of simple mean updated pwPE signals—with precision equivalent to a dynamic learning rate (Mathys et al., 2011, 2014; Mathys & Weber, 2020). The conceptual and mathematical simplicity of the HGF makes it an indispensable tool for the simple and efficient computational modelling of learning behaviour.

The general framework for the modelling of learning in dynamic environments with an HGF is to separate the world (with sensory input [u], and true hidden states [x]) from the agent (with inferred hidden states [λ] leading to action [y] upon the world, (Mathys et al., 2011, 2014), see **Figure 6**). The HGF can then accomplish the modelling of these inferred hidden states by taking the sensory input u over multiple trials k, performing a set of mathematical operations that estimate beliefs on those states using trial-wise values, and inferring upon the hidden states and guiding action by that procedure (Mathys et al., 2011, 2014). The HGF describes the relation between environmental states and an agent's mental states; however, beyond the normative description of behaviour, the HGF stresses individual variability in approximating Bayes optimal learning (Diaconescu et al., 2014). The focus is on capturing individual learning styles occurring over multiple trials. As such, the HGF has been widely applied to understand the differences in learning between groups and populations, such as with clinical groups or between experimental pharmacological populations (Bernardoni et al., 2018; Brazil et al., 2017; Cole et al., 2020; de Berker et al., 2016; Deserno et al., 2020; Diaconescu et al., 2020; Hauser et al., 2014; Henco et al., 2020; Katthagen et al., 2018; Lawson et al., 2020; Lawson et al., 2017; Paliwal et al., 2019; Powers et al., 2017; Reed et al., 2020; Sevgi et al., 2016; Stephan & Mathys, 2014; Weber et al., 2020).

For this thesis, we focus only on the simple HGF for binary reward outcomes in a rewarding environment with three levels, with the top being volatility (Mathys et al., 2011, see **Figure 6**). However, the HGF is not limited to binary outcomes or inputs; it can also be extended to any number of levels beyond the typical three (Mathys et al., 2014). Practically, moving forward in this section, we will examine a learning scenario where the agent is attempting to predict which of two images (a blue image or an orange image) presented simultaneously will lead to a reward outcome (reward or no reward) given a non-stationary correspondence between the

images and reward-outcome probabilities. While the above-mentioned RW model (**Section 1.3**) and the HGF model share similar properties, the HGF recasts learning as a function-approximation problem. These function approximations are precisely where the model earns its name: the HGF proposes that agents infer upon the causes of the sensory inputs using a sequence of hierarchically organised Gaussian functions where the variance (step size) of the level above dictates the learning rate of the Gaussian at the level below (Mathys et al., 2011, 2014).

Each level of the 3-level HGF conforms to a probability distribution and is associated with the aforementioned classifications of uncertainty (*expected*, *unexpected*, and *environmental* uncertainty, see **Figure 6**). The first level encodes the probability of a trial outcome to form a binary prediction gained from a sigmoid transformation of the second level Gaussian above (in our model this is without sensory uncertainty). The first level can be thought of as describing *expected* uncertainty (or, *irreducible* uncertainty)—as the result of a probabilistic outcome is, in most real-world contexts, not wholly knowable (Soltani & Izquierdo, 2019). As touched on above, the learning agent seeks only to learn the sufficient statistics of these distributions (means, variances, see **Figure 6**). The second and third levels are continuous quantities represented as Gaussians. The second level represents belief estimates of the tendency for either the blue or orange image to reward (in logit space), and the variance of this belief distribution can be interpreted as *informational* uncertainty (de Berker et al., 2016). The third level belief estimates represent the environmental uncertainty (or, *volatility*: how much the learner believes the reward outcome contingencies are changing, irrespective of the true rate of change see Soltani & Izquierdo, 2019). Out of that, when a learner estimates a high degree of volatility, they are more uncertain about the outcome and exhibit larger model updates.

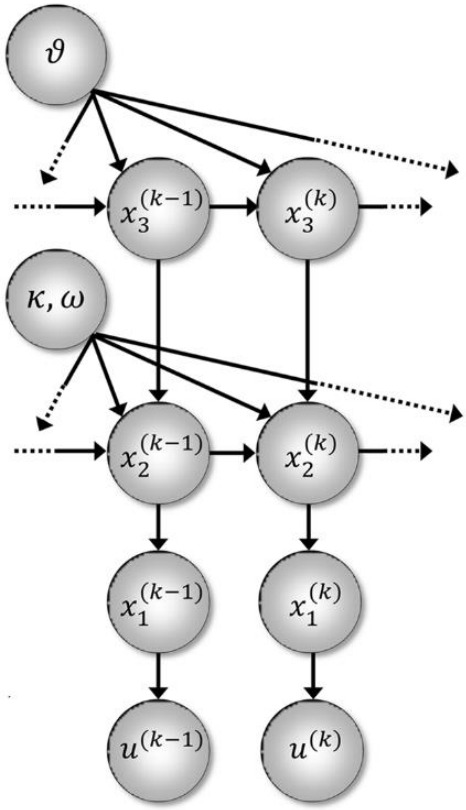


Figure 6. The Generative Model of the Hierarchical Gaussian Filter (HGF). Fitting sensory input or experimental stimuli (*u*) to observed responses using the HGF. The variables and parameters of higher levels dictate the probability of the level below by setting the step size (volatility or variance) of a random walk. At the third level in our 3-level HGF, the step size is a constant parameter ϑ. While x_1 sets the probability of the input *u* on the lowest level. As there is no perceptual uncertainty in our binary reward-learning task, the inputs *u* are equivalent to x_1, the rewarded outcome. The hidden layers x_2 *and* x_3 take the form of Gaussians. x_2 is the stimulus outcome tendency, which depends on a free parameter ω_2 (tonic volatility at level 2), κ, which represents the phasic component of volatility at level 2, and estimates of x_3. At the third level, x_3 represents volatility, whose step size is determined by the parameter ϑ (or, ω_3: a free parameter that can differ across learners). These parameters will be discussed in further detail in the following sections on the HGF. Figure adapted from the open-access article (Mathys et al., 2011). Copyright: © 2011 Mathys, Daunizeau, Friston and Stephan, see https://creativecommons.org/licenses/.

An important distinction is between the generative model of the HGF and its inversion—the inference model. The generative model of the HGF pertains to the outside world (sensory input and true hidden states), and the inference model concerns the agent (inferred hidden states and action). This perceptual model can then be coupled to a response model that links belief estimates with decisions (Mathys et al., 2011). Given the current beliefs, the decision model states the most probable action the participant will take. The response model also has parameters that can be estimated from the sensory inputs and responses of each individual, and the HGF can be used with different response models. Later, we use a fixed decision noise

parameter in **Chapters 2–3** and a dynamic decision noise parameter that depends on trial-wise changes to volatility estimates in **Chapters 4–5**. Together, the parameters of the perceptual model and the response model imply belief trajectories across the course of the task (see **Figure 7**).

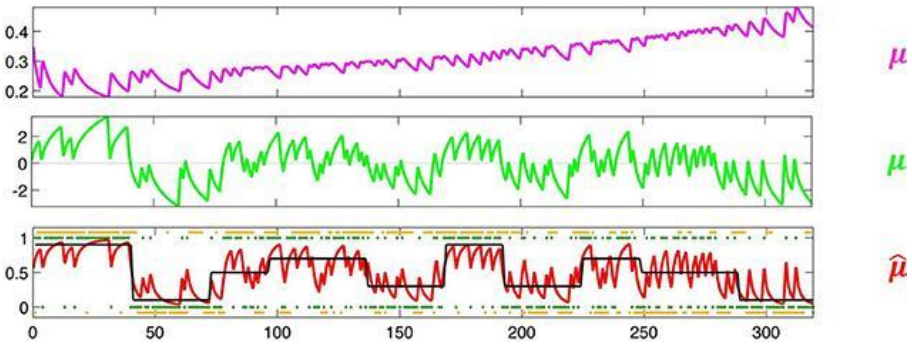


Figure 7. HGF belief trajectories. At the lowest level, the predictions in red (the sigmoid transformation of μ_2) track the black line which shows the 'ground truth' of the probabilistic input for one particular example participant (starting at 0.9 probability of reward for one stimulus, and then switching to 0.1 at around 45 trials, and so on until trial 320). Sensory inputs are shown in green dots, while orange dots show the participant's observed decisions. We can see the participant tracking the rewarding stimulus quite faithfully as the task proceeds. At the level above, in bright green, we see the posterior expectation μ_2 of the stimulus outcome tendency (which image the participant estimates to be more likely to provide reward). And on the top level is the posterior expectation μ_3 of the log-volatility x_3. Figure adapted from the open-access article (Mathys et al., 2014). Copyright © 2014 Mathys, Lomakina, Daunizeau, Iglesias, Brodersen, Friston and Stephan, see https://creativecommons.org/licenses/.

Put simply, the HGF is a set of update equations that tell you how—after providing the model with some inputs—the inferred hidden states are updated. The HGF arrives at the update equations by variationally inverting the HGF generative model (Mathys et al., 2011). Inversion of the generative model proceeds by introducing a mean field approximation and fitting quadratic approximations to the resulting variational energies (Mathys et al., 2011). This leads to simple one-step update equations, supplying estimates for the sufficient statistics (mean, variance; or, its inverse precision) of the approximate Gaussian posteriors of the states x_i (Cole et al., 2020; Mathys et al., 2011, 2014). The updates of these means exhibit the same general form as value updates in the RW model from **Section 1.3**, except that the fixed learning rate of the RW is dynamic in the HGF updates (using precision-weighted predictions errors [pwPEs]). The question of which model of the environment humans actually use (and how they invert it) is still, however, an open and thrilling one (Mathys & Weber, 2020).

The pwPE update depends on how precisely the learning agent thought they could make their prediction, and that is represented in the numerator of the ratio of precisions weighting the posterior update equation—provided in **Chapter 2** Equation 4 and in Mathys et al. (2011). If a learning agent thinks they can make exact predictions and are inaccurate, they may change their beliefs considerably. However, this also depends on how sure the learner is of how much they believe they already know the truth, leading to less changing of beliefs. These two precision quantities have an antagonistic relationship. Together they form the weight on the PE and determine the size of the update. This normative analysis also mirrors the previously mentioned theory of learning under the Pearce-Hall model (Pearce & Hall, 1980) that states unpredicted feedback will raise the learning rate whereas predicted feedback will lower the learning rate.

We refer the reader to **Chapter 2** for more detail on the relevant update equations of the HGF, and to the original HGF methods papers for further detail on the mathematical derivations (Mathys et al., 2011, 2014). Here we will emphasise how uncertainty takes a central role in the HGF (Mathys et al., 2011, 2014). Each level integrates the uncertainty (or variance, its inverse is precision π, interpreted from the width of the probability distributions) of the level above in the hierarchy. The precision expression that assigns a weight to a PE is defined as the inverse variance of the posterior expectation from that level (and comprises the pwPE that dictates the update steps of the HGF):

$$\pi_i^{(k)} = 1/\sigma_i^{(k)} \tag{8}$$

The precision of the prediction of the level below $\hat{\pi}_{i-1}^{(k)}$ is divided by the precision of the current belief $\pi_i^{(k)}$. This makes visible the aforementioned antagonistic relationship between the ratio of precisions that determine the size of the belief update. With this comes the core of the HGF: lower precision beliefs motivate faster learning on that level (unpredicted outcomes signal surprise driving up the speed of subsequent belief updating). Put simply, if we make predictions that turn out to be mistaken, we may need to alter our beliefs to improve future accuracy (the numerator of the precision ratio). However, if we are acutely certain of our beliefs about the state of the world (the denominator of the precision ratio) we may revise our beliefs less. This is a powerful function that leverages the primary advantage of Bayesian models over the above RL counterparts (**Section 1.3**). What appears initially complex in the HGF is in essence reducible to readily deciphered update equations through variational approximation using a small number of parameters.

The work in this thesis concentrates on the importance of uncertainty estimates in belief updating, particularly concerning anxiety and its close relation to uncertainty (see **Section 1.5**). As can be gleaned from the above equations, uncertainty takes a central position affecting the learning rate in the HGF update equations. Thus the HGF deals with all kinds of uncertainty from our environment. If we want to be optimal learners, we need to make an assessment of these distinct forms of uncertainty in our environment and pack it into our learning rate. For that reason, our rationale for using the HGF is that at its heart it makes visible uncertainty in the update equations (Cole et al., 2020; Mathys et al., 2011, 2014). The three types of uncertainty we have are *expected* uncertainty (*irreducible* uncertainty), *estimation* or *informational* uncertainty on level 2 (σ_2, belief uncertainty about outcomes) and level 3 (σ_3, belief uncertainty about volatility, see Mathys, [2014, Eq. 9-10]), and *volatility* that induces *unexpected environmental* uncertainty ([$\exp(\kappa\mu_3^{(k-1)} + \omega_2)$]: uncertainty about the parameters in our environment changing).

This relationship between uncertainty and learning rates has not only motivated the HGF, but also other alternative hierarchical Bayesian inference models expanding upon the foundations of the Kalman filter: learning from the dynamics of a linear dynamical system including estimates of process and observation noise (Behrens et al., 2007; Kalman, 1960; Piray & Daw, 2020b) and generative models that expect sudden changes (Moens & Zénon, 2019; Nassar et al., 2010). The question of whether participants use the optimal model for each task, or some form of general purpose mechanism that can approximate different regimes, is still an active and exciting area of debate. For all practical purposes though, in the research included in this thesis, the HGF and its alternatives are equally efficient at capturing learning dynamics in uncertain environments (Marković & Kiebel, 2016).

1.4.6.1 Application to experimental data

The HGF is especially useful in associating model estimates of learning to neural responses. Although this approach is not unique to the HGF (Stephan et al., 2015), an increasing number of studies are making use of the trial-by-trial estimates of computational quantities extracted from the HGF, and have started to provide evidence for the neural signatures of these parameters from fMRI and electrophysiological data (Auksztulewicz et al., 2017; Diaconescu, Litvak, et al., 2017; Iglesias et al., 2013, 2019; Stefanics et al., 2014). Among other things, these studies have begun to demonstrate how the brain uses hierarchically organised PEs that correspond to the different types of uncertainty in our learning environment. This approach of explaining neural responses with model estimates is also particularly productive when investigating differences in psychiatric groups, among pharmacological manipulations, and

between healthy and experimental populations (Bernardoni et al., 2018; Brazil et al., 2017; Cole et al., 2020; de Berker et al., 2016; Deserno et al., 2020; Diaconescu et al., 2020; Hauser et al., 2014; Henco et al., 2020; Katthagen et al., 2018; Lawson et al., 2020; Lawson et al., 2017; Paliwal et al., 2019; Powers et al., 2017; Reed et al., 2020; Sevgi et al., 2016; Stephan & Mathys, 2014; Weber et al., 2020). We will return to this in **Section 1.4.8** and when examining the impact of anxiety on reward learning in **Chapters 2–5**.

In a critical study by Iglesias et al. (2013), the HGF was used to map model-based quantities to neuroimaging data in healthy participants during a sensory learning task. That study found that pwPEs about visual outcomes modulated dopaminergic midbrain activity (without an association with reward or novelty, see **Figure 8**). In addition, they showed that higher level pwPE on outcome probabilities (visual outcomes conditional on auditory cues) was encoded by responses in the cholinergic basal forebrain (Iglesias et al., 2013, 2019; Stephan & Mathys, 2014)—verifying the hypothesis that the brain is processing a computational hierarchy of PE-based learning signals. Supporting these findings was an experiment conducted later by Diaconescu et al. (2017) using a comparable HGF model that replicated the same pattern of effects but while learning from social cues. As touched on above in **Section 1.3**, the dopaminergic PE is theorised to be a neural signal indicating the discrepancy from an anticipated outcome, both concerning rewards, and as shown above in this section, also sensory information (Gardner et al., 2018; Iglesias et al., 2013; Suarez et al., 2019). Others propound the dopaminergic PE response also updates beliefs concerning the environment through synaptic plasticity (Montague et al., 2004). One example of this is the modulating of NMDA receptors (Cole et al., 2020; Gu, 2002). Ultimately, this demonstrates that the HGF model is capable of capturing quantities that the brain is concerned with and uses, leading to neural activity in concert with the model trajectories for the hypothesised hierarchy of uncertainties (Behrens et al., 2007; Bland & Schaefer, 2012; Mathys et al., 2011; O'Reilly, 2013; Payzan-LeNestour & Bossaerts, 2011; Yu & Dayan, 2005).

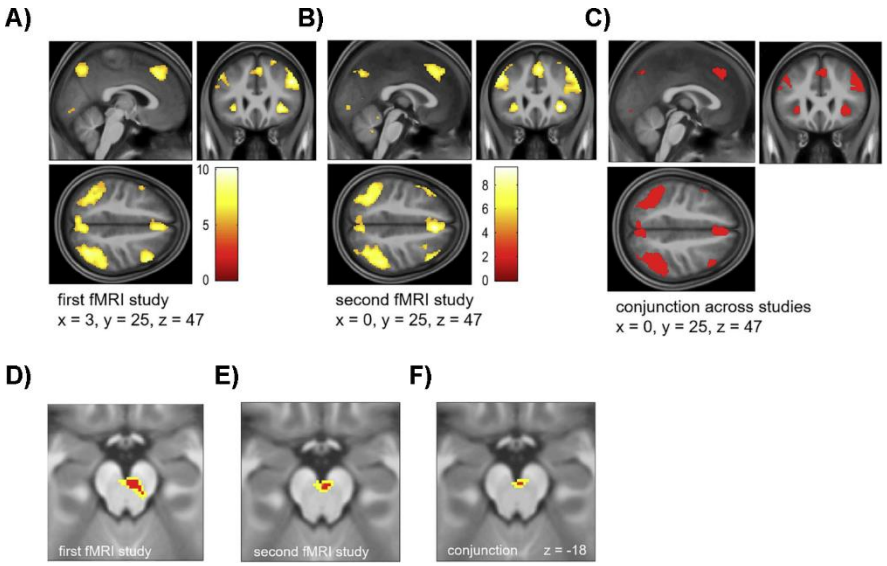


Figure 8. Activation in whole-brain fMRI data by HGF precision-weighted prediction errors on level 2. In the above, the Iglesias et al. (2013) original and replication experiments (yellow) are presented with the conjunction across the studies in red. For lower-level pwPEs about visual stimulus outcomes, significant activation was found in the first study A) and second study B) with the conjunction between the two in C) were encoded in several regions, including the dorsolateral PFC, ACC, and insula. Below that in D) original study E) second study and F) conjunction, Iglesias et al. (2013; 2019) also found activation in dopaminergic VTA and substantia nigra by pwPEs about visual outcomes. From the top level of their 3-level HGF model, they also found higher-level pwPEs about cue-outcome contingencies in cholinergic basal forebrain activity. As an overview, the above shows how much the brain responds from modulation from the visual outcome pwPE at the second level of the model, indicating participants were tracking the learning process in a hierarchical fashion and using hierarchically organised learning signals. Adapted with permission from Iglesias et al., (2013).

The limitations of fMRI neuroimaging data, including poor temporal resolution and a lack of spectral information, have since led to questions concerning the electrophysiological correlates of hierarchically organised PEs and pwPEs. The questions of how these quantities evolve in time and frequency motivated considerable research combining computational modelling and analysis of trial-by-trial EEG and MEG responses (Auksztulewicz et al., 2016, 2017; Diaconescu, Litvak, et al., 2017; Jepma et al., 2016; Kolossa et al., 2015; Liu et al., 2021; Mars et al., 2008; Stefanics et al., 2018; Weber et al., 2020). Event-related potential (ERP) responses such as the P300 are thought sensitive to reward PEs, valence, and surprise, tested using alternative Bayesian inference parameters (Kolossa et al., 2015; Mars et al., 2008; Ostwald et al., 2012). Also, the feedback error related negativity (fERN) is thought to index the degree of prediction violation, examined using RL models (Gehring & Willoughby, 2004; Holroyd & Coles, 2008). As we will see in **Chapter 2** using a single-trial approach, we show that the temporal and spatial dynamics of pwPE learning signals about probabilistic

stimulus outcomes are encoded in central and parietal regions, similarly overlapping with the late P300 in classical ERP analysis.

In Diaconescu et al. (2017), the computational modelling of behavioural responses using the HGF was used to associate EEG activity but in conjunction with fMRI responses. Healthy participants performed a task involving learning from social cues at multiple levels of uncertainty that were equivalent to the uncertainty that the HGF estimates (such as *estimation* and *volatility* uncertainty). Using trial-by-trial EEG responses, Diaconescu et al. (2017) showed that the sequence of hierarchical processing outlined by the HGF was expressed in the series of single-trial ERP responses from EEG activity. Lower-level PEs were represented across occipital, posterior, and occipital-temporal electrodes at an early temporal range 134–258 ms. Interestingly, the precision of these beliefs was also expressed in positive deflections around 352 ms at posterior and central electrodes. Higher-level PEs about volatility, however, were associated with later ~400 ms ERP activity in frontocentral electrodes. These results were later mirrored in an EEG study by Liu et al. (2021) using a visual probabilistic reversal-learning task with vacillating consistent and inconsistent probabilistic contingencies, much like Behrens et al. (2007). In that experiment, the author's showed how healthy participants use hierarchical learning by corresponding HGF low and high-level pwPE quantities to the encoding of the P300 ERP component—independently and adaptively (Liu et al., 2021).

Further elegant studies utilising the HGF have investigated the mismatch negativity MMN/m (Stefanics et al., 2018; Weber et al., 2020). These experiments have provided evidence that EEG recordings can distinguish in time and sensor space between lower-level pwPEs signals, which direct belief updates about outcomes (Stefanics et al., 2018), and higher-level pwPEs directing volatility belief updates (Weber et al., 2020). However, the neural processes of trial-wise ERP activity modulated by hierarchical pwPEs while learning from reward remains understudied. Also, this methodology has not previously been applied to an anxious population. In **Chapter 2**, we address this in an experiment investigating healthy and anxious participants using single-trial EEG analysis, associating HGF model estimates of pwPEs to ERP responses.

The spectral correlates of the HGF computational quantities have, up until this point, been critically overlooked—especially during reward processing. Initial studies using motor responses and visual and auditory tasks indicate, in line with the process theory of predictive coding (PC, Bastos et al., 2012), that PEs manifest at high gamma frequencies (>30 Hz) with concurrent anticorrelated drops in alpha/beta frequencies (8–30 Hz)—driving feedforward error signal transmission (Auksztulewicz et al., 2017; Palmer et al., 2019). Moreover, the PC

framework asserts predictions are transmitted down the cortical hierarchy in alpha/beta oscillations, supported by evidence from studies using the HGF (Auksztulewicz et al., 2017; Palmer et al., 2019). Yet the work so far with the HGF and other alternative models and approaches to modelling estimates of predictions and PEs (Bastos et al., 2012, 2015, 2018, 2020; Bauer et al., 2014; Sedley et al., 2016) has remained in sensory processing, leaving our understanding of how reward-based predictions and error signals are represented in the frequency domain incomplete. The work in this thesis addresses this in **Chapters 3 and 5**, targeting this paucity by examining how states of anxiety and high levels of trait anxiety modulate the oscillatory profiles of activity associated with reward-based learning signals relative to controls.

1.4.7 Representing uncertainty in the brain

A significant amount of evidence, as presented in this **Section 1.4**, suggests that humans learn and make decisions by representing uncertainty; however, precisely how we represent these quantities remains a lively area of research (Bach & Dolan, 2012; Griffiths et al., 2010; Hsu et al., 2005; Knill & Pouget, 2004; Koblinger et al., 2021; Ma & Jazayeri, 2014). Different experiments decompose decision variability into several sources aiming to determine their potentially distinct signatures in the brain (Soltani & Izquierdo, 2019; Vincent et al., 2019). Avenues of investigation have typically taken a pharmacological and neuroimaging approach, but other studies using electrophysiology and pupillometry have started to determine signals linked with quantities including precision (the inverse variance of a belief distribution) and surprise and volatility thought central to Bayesian predictive coding processes (Behrens et al., 2007; Browning et al., 2015; Iglesias et al., 2021; Marshall et al., 2016; Moran et al., 2013; Payzan-LeNestour et al., 2013; Peters et al., 2017; Vincent et al., 2019; Zénon, 2019). In the following section we present a broad neural framework for the different encoding of the different forms of uncertainty in the brain. Chiefly, this evidence serves to validate the hypotheses of Bayesian PC, as uncertainty plays such a core role in the updating of beliefs (Aitchison & Lengyel, 2017; Bastos et al., 2012; Friston, 2005; Shipp, 2016). But in addition, we explore uncertainty and its representations in the brain as uncertainty also significantly contributes to affective states like anxiety, examined in detail later in **Section 1.5** and **Chapters 2**–**5**.

The precise nature of how uncertainty is computed, encoded and distinguished remains unknown (Pouget et al., 2013). There are, however, two general groupings of models that take either a *population encoding* or *summary statistic* approach (Bach & Dolan, 2012). Briefly, the summary statistic approach has yielded several neural correlates of uncertainty, as belief

distributions are assumed to take some form (including Gaussian or Poisson) that are determined by their sufficient statistics. Previous work using summary statistics has most commonly concentrated on the mean and variance of probability distributions. Population encoding, by contrast, highlights the process of populations of neurons and how they may express belief distributions in contrast to point estimates (Ma & Jazayeri, 2014).

1.4.7.1 Summary statistics

For the summary statistics approach, our concentration from here will be limited to reward-based learning. Prior work has established the indispensable role of dopamine (Gershman, 2017; Niv, Duff, et al., 2005), acetylcholine and noradrenaline (Yu & Dayan, 2005), and serotonin (Doya, 2008; Rogers, 2011) in making decisions and learning from uncertainty. As an overview, the evidence so far implicates the involvement of several distributed regions including the PFC, ACC, striatum, basolateral amygdala (BLA), mediodorsal thalamus (MD), and hippocampus (Soltani & Izquierdo, 2019). Due to space restrictions, below we will focus primarily on the PFC, ACC, striatum, and BLA (for further details on the interconnected MD and hippocampus see Soltani & Izquierdo, 2019). In close propinquity and interconnected with the PFC, the anterior cingulate cortex (ACC) has been shown to signal the unsigned RPE, as already touched on above in **Section 1.3** (de Berker et al., 2019; Hayden et al., 2011; Rushworth & Behrens, 2008). However, in a more recent experiment, Monosov (2017) provided evidence that ACC neurons encoded not only expected value but *expected uncertainty* (variability in reward outcomes). Despite the RPE not alone constituting the encoding of uncertainty, the unsigned RPE can contribute robust evidence supporting an approximation of *expected uncertainty* (Soltani & Izquierdo, 2019). Also, research on the process of adapting to volatility changes has strongly associated the ACC and dmPFC with estimated *volatility* in humans (Behrens et al., 2007, 2008; Krugel et al., 2009; McGuire et al., 2014; Muller et al., 2019; Nassar, McGuire, et al., 2019), with reliable replication from single neuron recordings in non-human primates (Massi et al., 2018).

A collection of research utilising electrophysiological, MRI, and single-neuron recordings have broadly implicated the OFC in the PFC (and also parietal regions) in correlating with stimulus value and *expected uncertainty* (risk; or, uncertainty associated with probabilistic outcomes)—similarly to the ACC (Critchley et al., 2001; Hsu et al., 2005; Hunt et al., 2017; Izquierdo, 2017; Jo & Jung, 2016; Kepecs et al., 2008; McCoy & Platt, 2005; O'Neill & Schultz, 2010; Preuschoff et al., 2006; Riceberg & Shapiro, 2017; Rich & Wallis, 2016; Wallis, 2011; Wright et al., 2013). Specifically in the OFC, this capacity to process stimulus value and *expected uncertainty* operates in relation to the changing of environmental states, implying the OFC may be

involved with volatility processing (Massi et al., 2018; Saez et al., 2017; Soltani & Izquierdo, 2019). In pharmacological manipulation and lesion based studies, disengaging the OFC in rats and monkeys vitiates learning from uncertainty (Bradfield et al., 2015; Dalton et al., 2016; Izquierdo, 2017; Mobini et al., 2002; Noonan et al., 2010; Rudebeck, Saunders, et al., 2017; Soltani & Izquierdo, 2019; Stolyarova & Izquierdo, 2017; Winstanley & Floresco, 2016).

Despite strong evidence for the involvement of the PFC in representing parts of uncertainty, it is thought that alternative areas like the BLA and interareal connectivity with cortical and subcortical regions (striatum, MD, and hippocampus) compute and encode *environmental uncertainty* (both *unexpected* and *volatility* uncertainty, for further information we refer the reader to Soltani and Izquierdo [2019] and Bach and Dolan [2012]). The striatum, as mentioned in previous sections, has been linked with adaptive learning in responses to changes in individual outcomes (*expected uncertainty*), potentially through tonically active cholinergic interneurons (Franklin & Frank, 2015), but also as the striatum is the target for inputs concerning *expected* and *unexpected* uncertainty from the hippocampus, BLA, and frontal regions (Averbeck & Costa, 2017; Costa et al., 2016; Onge et al., 2012; Soltani & Izquierdo, 2019; White & Monosov, 2016). Lastly, data on the BLA from studies on rats and non-human primates strongly implicates the BLA in identifying surprise that produces faster adaptive updates (Costa et al., 2016; Morrison et al., 2011; Rudebeck, Ripple, et al., 2017; Saez et al., 2015). This is particularly important considering the overlap between brain regions associated with affective states and the amygdala and prefrontal regions—to be discussed in due course in **Section 1.5** concerning anxiety (Davidson, 2002; Davis, 1992; Robinson et al., 2019; Wassum & Izquierdo, 2015).

Also relevant is evidence that the BLA processes inconstant environments in valenced RPEs in response to the consistent violation of expectation (Roesch et al., 2010; Wassum & Izquierdo, 2015). Explanations for this processing of *unexpected uncertainty* in the BLA have been proposed, like with the aforementioned Pearce-Hall model (Pearce & Hall, 1980), that appeal to the salience of attention capture and so-called ‘associability’ (Roesch et al., 2012). In a paper by Stolyarova and Izquierdo (2017), the BLA was reported to track ground truth changes in volatility. This evidence on the BLA and learning under uncertainty potentially points to the BLA modulating value learning by its communication with the ACC and OFC through dopaminergic pathways (Amaral & Price, 1984; Cassell & Wright, 1986; Lucantonio et al., 2015; Sharpe & Schoenbaum, 2016; Stopper et al., 2014). According to Soltani and Izquierdo (2019, p. 7), these “projections could allow the BLA to compute unexpected uncertainty by comparing changes in stimulus or action values to baselines set by the expected uncertainty”.

1.4.7.2 Population encoding

Summary statistic approaches provide good evidence for how uncertainty is processed in the brain. However, we still need more evidence to understand precisely how *unexpected* uncertainty and *volatility* are processed in the brain. One alternative to the single-cell and the summary statistic approach is the encoding of probability distributions at the population level. We will not delve too deeply into the details here, but one particularly compelling model brought together representations of uncertainty, neuronal variability, and Poisson noise (Ma et al., 2006). The central idea from the Ma et al. (2006) study was that we can calculate probability distributions over neuronal populations as neurons encode probabilistic beliefs about the state of the world in a noisy fashion with inconsistent firing rates in response to identical stimuli (Deneve et al., 2001; Faisal et al., 2008). From our previous treatment of Bayes' theorem in **Section 1.4.1**, it can more readily be understood that these can form estimates of the distribution for the stimulus conditional upon a response. Happily, these Poisson distribution population codes (when linearly added) are equivalent to optimal Bayesian inference (Aitchison & Lengyel, 2017; Friston, 2010; Glaser et al., 2018; Ma et al., 2006; Zemel et al., 1998). Unfortunately, a full discussion about the potential role of population encoding in the processing of uncertainty and Bayesian parameters is far outside the scope of this thesis. There may even be routes to processing uncertainty quantities in ways that aren't widely appreciated yet. For now, it is enough to review some of the evidence for the processing of uncertainty in the brain.

1.4.7.3 Pupillometry

One line of research from early work in animals (Yu & Dayan, 2005) is the representation of volatility in the phasic responses of the noradrenaline/norepinephrine system (Aston-Jones & Cohen, 2005; Lavín et al., 2013) as measured by pupillometry (Koss, 1986; Marshall et al., 2016; Sales et al., 2019; Vincent et al., 2019). This is accordant to earlier theory asserting a general purpose for noradrenaline for amplifying sensory gain (Aston-Jones & Cohen, 2005; Sales et al., 2019) that in turn regulates learning behaviour (equivalent to modulating the learning rate, Nassar et al., 2012). Pupillary dilatation as a proxy for phasic noradrenaline signalling has been reliably demonstrated in non-human primates (Joshi et al., 2016) and healthy humans (Jepma & Nieuwenhuis, 2011; Lavín et al., 2013; Muller et al., 2019; Nassar et al., 2012; Preuschoff et al., 2011) and subsequently linked with estimates of volatility and surprise about environmental conditions in both healthy and clinical populations in a Bayesian

PC framework (Jepma & Nieuwenhuis, 2011; Lawson et al., 2017; Muller et al., 2019; Pulcu & Browning, 2017; Vincent et al., 2019).

Taken together, the evidence here implies a physiological process (pupil dilation) may express surprise from unexpected outcomes and changing environmental states, with this normative response linked with adaptation to environmental volatility (Larsen & Waters, 2018). Of particular interest to this thesis, high levels of trait anxiety have been shown to dampen the pupil response during unstable portions of probabilistic learning from aversive stimuli, which is thought to undergird difficulties adapting to changes in volatility (see Browning et al. [2015] and further detail in **Section 1.5**). More remains to be known about the different ways in which uncertainty is represented in the brain; little is understood, for instance, of the role of neural oscillations in encoding the precision of beliefs, or how these modulate PE signals, and how these are changed by affective states like anxiety. We explore this in part in **Chapters 3 and 5** by examining how pwPEs modulate neural oscillatory activity in the brain using EEG and MEG, respectively. Moreover, we are only just beginning to tease apart how humans distinguish environmental change from unpredictability (the noisiness of outcomes); as such, how these uncertainties are distinguished and processed in the brain is an important but nontrivial challenge. Resolving these issues demands rich datasets and well-suited models to directly distinguish these uncertainties and contrast their processing in the brain (Piray & Daw, 2020a, 2020b).

1.4.7.4 Neuromodulation

Other widely connected neuromodulatory networks are also theorised to regulate uncertainty and perform precision (*attention*) or synaptic gain tuning in the brain (Dayan, 2012b; Doya, 2008; Iglesias et al., 2017, 2021; Sales et al., 2019; Yu & Dayan, 2005). Neuromodulatory systems convey uncertainty at slower timescales, whereas alternatively, ionotropic neurotransmitters have been argued to signal predictions and PEs directly (Friston, 2005; Shine et al., 2021). Theoretically, the neuromodulators act like a transistor, calibrating the activity of particular neural circuits (sensory inputs) corresponding to their uncertainty; turning up the signal to specific sensory PEs in the brain, and as such, the weight of influence on belief updates and the learning rate (Shine et al., 2021). Some notable neuromodulators theorised to play a role in precision include acetylcholine (ACh, see Moran et al., 2013; Yu & Dayan, 2002) and noradrenaline (NA, see Dayan & Yu, 2006; Zhang et al., 2013), and dopamine (DA, see Fiorillo et al., 2003; Friston et al., 2012, 2014; Hart et al., 2015; Iglesias et al., 2013; Schwartenbeck, FitzGerald, Mathys, Dolan, & Friston, 2015). These neuromodulators have been shown to contribute to the precision tuning of predictions and

sensory input and play a part in transmitting uncertainty and RPEs (Iglesias et al., 2021; Sales et al., 2019; Yu & Dayan, 2005). Although the exact implementation is different, these neuromodulators have been associated with precision in PC process theories (Spratling, 2017) and active inference accounts (Friston et al., 2011).

Each neuromodulator is thought to be involved with a different component of the PC process. For instance, ACh is proposed to adjust the gain of feedforward error signals in supragranular pyramidal cells (Dayan & Yu, 2006; Parr & Friston, 2017). Nicotinic ACh has been demonstrated to tune visual sensory precision (Disney et al., 2007; Parr & Friston, 2017), which implies the cholinergic system encodes the likelihood precision (Parr & Friston, 2018). And evidence supporting the involvement of ACh in encoding precision comes from both pharmacological (Marshall et al., 2016; Vossel et al., 2014) and neuroimaging (Moran et al., 2013) studies.

Midbrain DA neurons are hypothesised to reflect the encoding of RPEs (Bayer & Glimcher, 2005) and the precision of beliefs about policies (Schwartenbeck, FitzGerald, Mathys, Dolan, & Friston, 2015). Iglesias et al. (2013) provided some evidence that ACh and DA both contribute to the precision weighting of PEs, although at distinct levels of the cortical hierarchy. Also, the conclusions from Iglesias et al. (2013) need to be taken with some caution, as it is difficult to interpret the fMRI results as exclusively reflecting DA and ACh functioning, for the dopaminergic midbrain and cholinergic basal forebrain where Iglesias et al. (2013) found correlates of lower-level and higher-level PEs are not exclusively comprised of those respective neuromodulators (for example there are also GABAergic and glutamatergic neurons present, see Düzel et al., 2009; Iglesias et al., 2021; Shine et al., 2021). Iglesias et al. (2021) sought to address this by pharmacological manipulation of DA and ACh, showing how these changes might modulate the expression of pwPEs in DA and ACh nuclei; but alas with no clearer results. Higher-level cholinergic pwPEs were not cleanly dissociated from lower-level dopaminergic pwPEs.

Decision variability is thought modulated by noradrenaline (Jahn et al., 2018; Sara & Bouret, 2012), with the precision of transitions (estimated volatility) represented in the locus coeruleus (Marshall et al., 2016; Parr & Friston, 2018). But noradrenaline has also been linked with signal detection (Aston-Jones & Cohen, 2005; Eldar et al., 2013)—which implies a functional compromise among exploratory actions and exploitative actions (Sales et al., 2019; Yu & Dayan, 2005). Other neuromodulators like serotonin (linked to various psychiatric conditions) have been associated with temporal discounting (Miyazaki et al., 2012) suggesting modulation to higher-level beliefs (Shine et al., 2021). For more details on neuromodulators and their role

in predictive processing accounts of brain function, we direct the interested reader to (Parr & Friston, 2018; Shine et al., 2021). Further detailed treatment on current theories and evidence on representing uncertainty in the brain can also be found in Koblinger et al. (2021) and Dehaene et al. (2021). For now, it is sufficient to communicate that there are different processes that may control neural gain and the precision of beliefs in the brain. Interestingly, more recent work has explored the idea that neuromodulators such as DA may be transmitted down different frequency bands in accordance with a multiplexing principle (Gardner et al., 2018; Iglesias et al., 2021; Nakahara, 2014), potentially connecting with phasic/tonic DA firing (Grace, 1991).

1.4.8 Precision pathologies and brain dysfunction

Accurate inference and adaptive learning depend on the subtle harmony of precision in multiple hierarchical levels. Accordingly, a growing number of studies have been linking impaired estimations of uncertainty to neuropsychiatric populations (Corlett & Fletcher, 2014; Friston, 2017; Friston, Redish, et al., 2017; Huys et al., 2021; Parr, Benrimoh, et al., 2018; Parr & Friston, 2018). In the process of belief updating, according to Bayesian PC theory, the brain is required to resolve which predictions to attend to and which error signals to disregard (Friston, 2005). Precision serves this function by attenuating or amplifying PEs, thus mediating what is “newsworthy” (Feldman & Friston, 2010). Alternatively put, attention assigns more weight to reliable information (Parr, Rees, et al., 2018). But precision is a quantity that needs estimating. And accordingly, problems estimating precision can produce inaccurate and atypical posterior beliefs and false inference as shown across several neuropsychiatric conditions (Adams et al., 2012; Fletcher & Frith, 2009; Friston, Redish, et al., 2017; Hauser et al., 2016; Lawson et al., 2014; Pellicano & Burr, 2012; Powers et al., 2017).

From a physiological perspective, relating back to attention, the tuning from precision is conducted by mechanisms that regulate synaptic gain (Feldman & Friston, 2010). The synaptic gain process is theorised in multiple mechanisms, from neuromodulators such as acetylcholine (Moran et al., 2013; Yu & Dayan, 2002), noradrenaline (Dayan & Yu, 2006), and dopamine (Friston et al., 2014; Marshall et al., 2016), to synchronous gain and the ratio of excitation-inhibition (Friston, 2017). These candidate neurobiological mechanisms for precision bridge psychiatric and physiological pathology (Friston, 2017), opening up an investigation that spans Marr’s three levels of analysis: the computational, algorithmic, and implementational (Marr, 1982; Marr & Poggio, 1976). While computational psychiatry is no panacea, this does provide a good framework to understand and refine the modelling of psychiatric disorders (Franklin & Frank, 2015; Hauser et al., 2016).

This line of research into precision pathology has borne many subsequent studies in a diverse array of neuropsychiatric populations (naming but a few: autism, Parkinson's disease, schizophrenia, depression, stress, and anxiety). One recurring motif is a specific precision dysfunction that fails to maintain this equilibrium between sensory input and prior expectations (Friston, Redish, et al., 2017; Williams, 2016). Put simply, precision dysfunction is a departure from the appropriate estimation of uncertainty in a complex and changing environment. Aside from disorders commonly considered psychiatric, this is especially germane in Parkinsonism (Adams et al., 2016).

It is well established that in Parkinson's disease there is atypical dopamine (neuromodulation) signalling in concert with a specific decline in the striatal structure that helps to regulate the aforementioned gain control mechanism—leading to a lack of control over increasing and decreasing sensory attention (Dauer & Przedborski, 2003; Friston, 2017). Deterioration in the substantia nigra pars compacta culminates in a dearth of dopamine in the striatum (Albin et al., 1989). This dopamine deficit produces difficulties generating movements that are thought explainable through a Bayesian PC lens (Parr, Rees, et al., 2018). Decreasing dopamine may bring about a stronger dependence on gradually shifting prior beliefs that fail to evaluate and integrate changes in time for the requisite movement (Jávor-Duray et al., 2017; Parr & Friston, 2018). Interestingly, raising the level of dopamine, as with some Parkinson's medication, may relax the dominance of hierarchically higher-level beliefs over lower levels, reintroducing spontaneous planning and impulsive actions (Parr, Rees, et al., 2018). Of particular pertinence to the reward learning focus of this thesis, elegant work by Frank and colleagues investigated reinforcement learning from both positive and negative outcomes using a probabilistic task, focusing on the encoding of dopaminergic neural firing and breaks in firing (Frank et al., 2004; Maia & Frank, 2011). In a Go and NoGo learning context, Frank et al. (2004) demonstrated that Parkinsonian patients with depleted striatal dopamine learn better relative to controls from negative outcomes, but that medication targeting dopamine reverses this bias (Adams et al., 2016; Maia & Frank, 2011). Importantly, this imbalance may subserve unhealthy compulsively motivated gambling (disproportionately learning from wins but not losses), with the mirror pattern of learning (better learning from losses and not wins) seen in antidopaminergic medicated Tourette's syndrome (who are hyperdopaminergic, see Adams et al., 2016; Palminteri et al., 2009). This is also connected with motivation and task vigour in apathetic populations (Bonnelle et al., 2015) as observed in Parkinson's disease but also in depression and several other neuropsychiatric diseases (Adams et al., 2016; Ang et al., 2017; Le Bouc et al., 2016; Pessiglione et al., 2017).

Autism and autism spectrum disorder (ASD) is one further condition that has been studied in this Bayesian PC framework (Lawson et al., 2014; Pellicano & Burr, 2012). Work so far has been informed by earlier observations that autistic individuals are to some degree insusceptible to sensory illusions (Happé, 1996; Simmons et al., 2009), may be better at tasks involving the detection of low-level features (Happé, 1999; Shah & Frith, 1983, 1993), and exhibit a local processing bias (e.g. Plaisted et al., 1999). However, later empirical work has also found no significant differences between ASD and controls in illusion susceptibility (Milne & Scope, 2008), and further discrepancies exist concerning motion coherence (Del Viva et al., 2006; Milne et al., 2002) and other visual processing tasks in ASD (for a review see, Milne et al., 2005; Simmons et al., 2009). More recent work in understanding perception in autism from a Bayesian perspective has interpreted autism as a condition of lower precision priors relative to high sensory precision, leading to increased estimates of and larger updates about environmental volatility, to the detriment of learning about outcome noise (expected uncertainty, see Lawson et al., 2014, 2017; Pellicano & Burr, 2012).

More specifically, this overreliance on incoming sensory data has been formulated as excessive and disproportionate precision estimates on the likelihood distribution (Lawson et al., 2014) with a potential origin in the low precision prior belief about *environmental volatility* (Lawson et al., 2017). The specific relevance of volatility is that overestimated environmental change thwarts precise estimates of the current state based on previous experience, making priors obsolete. Consequently, incoming sensory PEs exert a greater influence on belief updates.

In a later study utilising the Bayesian HGF model, Lawson et al. (2017) provided compelling evidence that individuals with autism overestimate environmental volatility. Supplementing this computational result, linking to our previous discussion of the relevance of pupil responses and volatility estimates, autistic participants were found to exhibit increased encoding of trial-by-trial changes in phasic noradrenaline as measured by pupillometry—reflecting disproportionate alterations to cortical gain in response to environmental change (Lawson et al., 2017). Additional studies have since expanded upon these findings, supporting the hypothesis that uncertainty and noise play a core role in inference in autism (Bast et al., 2021; de Vries et al., 2021; Manning et al., 2017; Palmer et al., 2015; Pomè et al., 2020; Seth & Friston, 2016; Van de Cruys et al., 2017; van Schalkwyk et al., 2017).

Critically relevant for this thesis, anxiety is also thought to change our beliefs about the nature of the world (Aylward et al., 2019; Grupe & Nitschke, 2013; Paulus & Stein, 2010; Paulus & Yu, 2012) and is proposed connected to misestimation of uncertainty (Huang et al., 2017;

Pulcu & Browning, 2019; Williams, 2016). As an overview, before we investigate in much further detail in the following chapters of this thesis, anxiety has been linked with several biases in cognition, some of which produce conflicting adaptive or maladaptive behavioural responses depending on the context (Grupe & Nitschke, 2013). Notably, there seems to be an important difference between performing sensory-perceptual based tasks (including threat detection) and higher-order executive functioning tasks (like calculating risk and probability, see Grillon et al., 2019; Grupe & Nitschke, 2013; MacLeod & Mathews, 1988; Miu, Miclea, et al., 2008; Robinson et al., 2019). From the sensory-perceptual perspective, anxiety appears typified by a pessimistic bias in evidence gathering as opposed to a bias in prior beliefs (Aylward et al., 2020; Kim et al., 2020). Studies have shown anxiety amplifies the processing of sensory input producing better performance on inhibitory and threat detection tasks (Cisler & Koster, 2010; Grillon, 2008; MacLeod & Mathews, 1988). However, performance in economic decision making and probability learning is impaired by anxiety (Grupe & Nitschke, 2013; Huang et al., 2017; Miu, Heilman, et al., 2008; Remmers & Zander, 2018).

Both sensory-perceptual and executive accounts could be explained by stronger or more resistant prior beliefs, as selectively attending towards threat (greater precision for prior beliefs about threat) may produce the above mentioned adaptive hypervigilance and sensory processing to detect threat (MacLeod & Mathews, 1988; Richards et al., 2014). Likewise, a bias in prior beliefs leads to a negative interpretation of ambiguous stimuli (Hartley & Phelps, 2012; Mathews & MacLeod, 2005) and intolerance to uncertainty (Carleton et al., 2012), which can lead to inflexible prior beliefs as shown in biases toward probability expectation and pessimistic negative outcome evaluation (Borkovec et al., 1999; Butler & Mathews, 1987; Kim et al., 2020; Mitte, 2007; Stöber, 1997).

Current reinforcement and Bayesian models are thus being employed to understand how anxiety influences the learning processes, and have been expanding our comprehension of anxiety as a learning problem (Browning et al., 2015; Grillon et al., 2019; Grupe & Nitschke, 2013; Pulcu & Browning, 2017, 2019; Raymond et al., 2017; Robinson et al., 2013). But precisely how anxiety affects learning in different contexts and interacts with the different forms of uncertainty discussed so far remains difficult to pin down. The work in this thesis attempts to resolve some of these issues. In **Chapters 2–4** we ask whether a temporary state of anxiety induced in the lab can shape learning from probabilistic reward in a volatile task environment using the HGF Bayesian model. And further, in **Chapters 2 and 3**, how this may impact the pwPEs and predictions vital to the proper updating of beliefs in a Bayesian PC process that modulates EEG responses. Whilst in **Chapter 5** we turn to expand this same line of enquiry

by focusing on high levels of trait anxiety and reward learning dynamics under uncertainty using high-resolution MEG recordings.

1.4.9 Summary and relevance

Our environments exhibit copious uncertainty, and this uncertainty represents an unavoidable feature that any learner needs to process to perceive and act efficiently. While not without its critics (Bowers & Davis, 2012), the statistical approach of Bayesian systems gives us a potent and sophisticated means to handle all forms of uncertainty and has established its value in motivating hypotheses concerning human learning and behaviour, especially providing insights into the non-optimal and atypical processes in certain neuropsychiatric groups. There is robust evidence that humans use Bayesian inference while performing sensorimotor and value-based learning and decision-making tasks by assimilating different types of information in line with its uncertainty. This line of enquiry has motivated different forms of hierarchical models of learning, including the Hierarchical Gaussian Filter, where this distinct arrangement of hierarchically organised uncertainty (*expected*, *unexpected*, *estimation*, *volatility*) are updated using an approximation of the Bayesian process in a PC framework. There is substantial data from normative accounts of learning that humans adapt to changes in their environment, and more recent evidence of a similar accomplishment concerning *expected* and *estimation* uncertainty. These investigations imply that we manage and conserve uncertainty estimates and utilise them in refining our learning. Computational modelling has afforded an elegant means to make visible and explicit predictions about a particular participant's or group's behaviour and their resultant neural activity while performing experimental tasks.

It is easy to see how a dysfunction in the proper balancing of precision between prior and sensory evidence may represent a ubiquitous computational mechanism in pathophysiology (failures in neuromodulation or by neurodegeneration) and might explain many features of neuropsychiatric disorders. In this thesis, we present evidence that altered precision estimates are not exclusive to clinical populations, but that we all experience everyday emotional states like anxiety that can alter how we estimate precision and consequently learn from uncertain environments. As we shall see in the pages that remain to us, neurotypical participants with high trait anxiety, or those passing through a transient state of anxiety also experience changes to uncertainty estimates while learning about reward, with changes to M/EEG representations of these processes in the brain (**Chapters 2–5**).

1.5 Emotions and anxiety

1.5.1 Emotional states

Emotional states are tightly coupled to the way we perceive the world, make decisions, and act (Heilman et al., 2010; Lerner et al., 2015). Emotions help to organise and give flavour to our experience (Bach & Dayan, 2017) and are thought evolutionarily adaptive (Darwin, 1956). Emotional and cognitive processes are thus interwoven. They shape ongoing behaviour and neural operations simultaneously (Pessoa, 2017), with subjective experience depending on cognitive systems—including attention and working memory (Baars & Franklin, 2003). A dysfunction in the regulation and control of emotions is a typical underlying factor in many psychiatric disorders (Dolan, 2002). Normative methods, like the previously covered reinforcement learning, have been shown to capture a closer mapping to behaviour when extended to integrate emotions (Sequeira et al., 2011). And more recent work in the Bayesian framework has been used to describe behaviour while incorporating interoceptive feedback and emotions (Garfinkel et al., 2015; Seth, 2013; Seth & Friston, 2016; Smith et al., 2019; Wager et al., 2015). These recast emotional states as inference about internal bodily states. Consequently, emotions can be thought of as statistics holding information about the properties of a state, influencing beliefs and actions about the world—potentially distorting inference (Allen et al., 2008; Eldar et al., 2016; Joffily & Coricelli, 2013; Moutoussis et al., 2018). One example comes from computational psychiatry studies that have examined depressive moods and anhedonic/fatigue states, often treating depression as a result of low meta-cognitive beliefs about the degree of control over bodily states (Barrett et al., 2016; Paulus & Yu, 2012; Rutledge et al., 2017; Schutter, 2016; Smith et al., 2020; Stephan, Manjaly, et al., 2016).

Yet the field of emotion research in psychology is rife with empirical debates, in particular about the definitional differences between potentially distinct emotions. Are emotions cleanly divided into exclusive states, as propounded by James (1922) and Ekman and colleagues (1983)? Alternatively, are emotions continuous and exhibit a precise value on a set of valence and arousal axes (see Russell, 2009)? Psychology has chiefly concerned itself with self-reports of feelings (like with happiness, anger, fear) that lend examination relative to affect (Bach & Dayan, 2017; Kuppens et al., 2013; Russell, 1980), or subjective experience relative to alterations in the body, action or motivation (Bach & Dayan, 2017; Oatley & Johnson-Laird, 2014; Scherer et al., 2001). By contrast, other work has considered first and foremost the cross-species ethological study of emotional facial expressions, speech articulations, and

body mannerisms (Darwin, 1956), making links between humans and non-human animals by categorising ‘anxiety-like’ or ‘fear learning’—explicitly admitting difficulties in interpretation (Bach & Dayan, 2017; Burgdorf & Panksepp, 2006; Calhoon & Tye, 2015; Herry & Johansen, 2014; Stephan, Bach, et al., 2016). As a consequence of this variability in thought about emotions, the theoretical landscape is equivalently varied.

Here we attempt to obviate the complex discussion about definition by taking a cognitive science approach, concentrating on reporting the potential causes and effects of emotions as a result of cognitive processes—like the identification of threat and the gaining of reward (Oatley & Johnson-Laird, 2014). We therefore ask: what are the normative roles of emotion? For example, in a happy or safe emotional state, organisms show a preference for approaching behaviour, as with food or other rewards. By contrast, negative emotional states like fear or anxiety produce escape or behavioural inhibition to threatening stimuli, respectively. On the arousal axes of emotions, this might produce high levels of alertness or freezing responses (De Boer & Koolhaas, 2003; Grupe & Nitschke, 2013; Lang, 1995; Perkins & Corr, 2014).

To achieve this, we need to flesh out the argument for emotion as a feedback process and consider the evidence for the interactions between emotions and decision making and learning. After, we turn to focus on the emotional experience of anxiety; looking more closely at how anxiety shapes the way in which we learn from and interact with complex dynamic environments. We aim to connect environmental uncertainty with anxiety and how this impacts learning behaviour and neural processing. According to recent theory, volatility might provoke feelings of anxiety, as environmental uncertainty concerns the space of uncertainty itself (Dugas et al., 1997, 2005). It is this emotional ‘indeterminability’ that others have associated with intolerance of uncertainty, especially in anxiety (Carleton, 2016; Carleton et al., 2012). Interestingly, intolerance of uncertainty in anxiety has been shown to predict increases in striatal volume (Kim et al., 2017), relating anxiety to striatal circuits for reward learning and decision-making (Balleine et al., 2007). If this is a correct interpretation, we aim to triangulate this claim by investigating in later **Chapters 2–5** how anxiety interacts with learning in uncertain reward environments.

1.5.2 Emotion as a slow decision making system

Emotions provide feedback to cognitive processes like decision making (Paulus & Yu, 2012). They are in that sense predictive. Consider a hyena being chased by a lion, trying to predict whether it will die. Fear can be interpreted as the hyena's prediction of whether it will die (Fanselow, 1994; Lang et al., 1997; Miller, 1948). Similarly, rumination and worry in anxiety disorders can be understood as maladaptive problem solving (Bishop & Gagne, 2018; Kircanski et al., 2015; Szabó & Lovibond, 2006; Treynor et al., 2003; Watkins, 2008). Yet Baumeister and colleagues aver that emotions like fear are too sluggish in the face of life or death scenarios as described above (Baumeister et al., 2007). As such, emotions are thought to slowly modulate future states and responses by experience, potentially accomplishing this through learning biases in memory (Cahill et al., 1994; Strange & Dolan, 2004). There is, however, an important distinction between *affect* as a short term, expressive and specific emotional experience, and *mood*, as a longer-term, latent and diffuse experience (Clark et al., 2018; Lang et al., 1997; Mendl et al., 2010). One excellent study providing evidence for mood-based alterations to decision making conducted an experiment where mood is held constant by the ingestion of a sugar pill that ostensibly causes participants to become unaffected by mood alterations for the following couple of hours (Manucia et al., 1984). That study demonstrated that mood, when believed unalterable, can either promote helping behaviours (when happy) or promote futility in helping (when sad)—also showing that behaviour induces changes to mood. This interaction between emotional feedback and learning and decision making has since enjoyed considerable focus in neuroscience (Aylward et al., 2019, 2020)—with an excellent computational review of mood in (Clark et al., 2018). Articulated in the terminology of the preceding sections of this thesis, emotions as either affect or mood may interact with model-free reinforcement learning processes, or with model-based processes through the biasing of (hyper) priors or modulation of computational quantities like precision. However, the model-based mechanism is particularly prone to biases, like with the emotional bias toward salient stimuli (Gilbert & Wilson, 2009) or biases in attention toward negative facial expressions (Duque & Vázquez, 2015), and more generally toward negative emotions in depression (Browning et al., 2012; Harmer et al., 2009; Harmer & Cowen, 2013).

The following **Chapters 2–5** continue this line of reasoning by testing the effect of anxiety in shaping ongoing reward learning behaviour and brain dynamics. We focus on assessing how anxiety might bias computations of uncertainty during reward learning in changing environments. From a clinical perspective, mood instability is a common complaint across diagnoses (particularly prominent in bipolar disorder and borderline personality disorder, Broome & de Cates, 2015; Henry et al., 2001). And the effects of an anxious mood, as with

states of anxiety, are understudied using a computational approach, especially concerning the engagement with reward feedback. In **Chapter 2** we tested how inducing state anxiety changes how participants learn from rewarding outcomes and unrewarded outcomes, modelling behaviour with the HGF and testing how pwPEs modulate EEG signals. Despite not demonstrating a mechanistic link between these processes, that study demonstrates that anxiety biases uncertainty estimates and alters reward learning performance. Later in **Chapter 3**, we show how these changes to HGF model estimates of predictions about the tendency of reward and pwPEs about stimulus outcomes are correlated to altered spectral content of the EEG signal. While in **Chapter 5** we report how high levels of trait anxiety shape ongoing reward learning performance, HGF model estimates, and pwPE and prediction modulation to oscillatory responses of MEG signals.

1.5.2 Adaptive emotions: anxiety and fear

Emotions serve an adaptive function, with fear and anxiety preparing animals for threats and contributing toward survival (Grillon et al., 2019; Grupe & Nitschke, 2013; Perkins & Corr, 2014; White et al., 2010). In contrast to fear—which is typically in response to an explicit and imminent threat—anxiety is a more diffuse response to equivocal threats with predicted negative outcomes (Grillon, 2008; LeDoux & Pine, 2016; Steimer, 2002). Anxiety prepares organisms for survival; to deliberate, assess risk, and deploy effective strategies from previous experiences (Grupe & Nitschke, 2013; Perkins & Corr, 2014). By this definition, anxiety concerns uncertainty over potential hostile outcomes, producing distinct behavioural, neural, and cognitive states compared with similar adaptive responses to fear or stress (Davis et al., 2010; Grillon, 2008; Grillon et al., 1991; Tovote et al., 2015). However, fear is also convolved with anxiety, producing perturbation and turmoil in response to impending threats, with the special case of posttraumatic stress: where fear exists detached from the experience that first induced the anxious experience (Ehlers & Clark, 2000). It is therefore important to investigate anxiety using several alternative methods to inform a more complete understanding of the neurobiological system (Grupe & Nitschke, 2013; Perkins & Corr, 2014).

Uncertainty is so central to anxiety that assessments of uncertainty play a core role in diagnosing anxiety disorders (Carleton et al., 2012; Quintana et al., 2016). Clinical anxiety disorders are an increasingly high-priority public health issue (Grillon et al., 2019; Stein & Craske, 2017), with anxiety representing the primary driver of mental health disability (Kessler et al., 2005) coming at a tremendous expense to both society and individuals (Chisholm et al., 2016). Accordingly, anxiety has been extensively studied (Blanchard & Blanchard, 1988; Gray & McNaughton, 2000), receiving growing research attention in the last decade, producing

elegant evidence for the neurobiological processes and physiological mechanisms that subserve the creation and maintenance of this state (Bishop, 2007; Calhoon & Tye, 2015; Grillon et al., 2019; Grupe & Nitschke, 2013; Mkrtchian, Aylward, et al., 2017; Quirk & Mueller, 2008; Robinson et al., 2013, 2019).

1.5.2.1 Trait, state, and clinical anxiety

Subclinical "trait" anxiety refers to a reasonably reliable representation of a person's general degree of anxiety (Taylor, 1953; Wiedemann, 2015), typically measured using self-report questionnaires (Carver & White, 1994; Spielberger et al., 1968; Watson et al., 1988). Spielberger's State-Trait Anxiety Inventory (STAI, Spielberger, 1983a) measures trait anxiety by the self-reported frequency of experienced anxiety from everyday events over the lifespan, coupled with an overall pronounced view of the world as threatening (Clark & Beck, 2011; Raymond et al., 2017; Spielberger, 1983a). While the STAI has been shown to closely overlap with components of anxiety and depression, it is a good metric to show a person's risk of developing an anxiety disorder and likelihood of experiencing future anxious states (Grupe & Nitschke, 2013; Nitschke et al., 2001; Raymond et al., 2017; Watson et al., 1995).

A temporary "state" of anxiety, by contrast, is defined as a transient affective state related to temporary apprehension and worry. State anxiety can also be measured using self-reported feelings of anxiety, but concerning the moment of filling out the report (Spielberger, 1983a). State anxiety is also linked with physiological markers, including respiratory and heart rate changes (Endler & Kocovski, 2001; Spielberger, 1979; Wiedemann, 2015). Despite escalating levels of trait anxiety being associated with an increase in the risk of clinical anxiety (Chambers et al., 2004), precisely how this operates remains unknown, but the interrelation between increased threat processing and experiencing anxious states is one hypothesised route (Raymond et al., 2017; Robinson, Overstreet, et al., 2013).

Classifications of disorderly clinical anxiety are commonly specified using the Diagnostic and Statistical Manual of Mental Disorder (DSM-IV and DSM-V, American Psychiatric Association, 2013; Segal, 2010). Clinical anxiety disorders in the DSM-IV include generalized anxiety disorder (GAD), panic disorder (PD), social anxiety disorder (SAD), posttraumatic stress disorder (PTSD), specific phobias, and obsessive-compulsive disorder (OCD). The major change between the DSM-IV and DSM-V worth noting here is the exclusion of post-traumatic stress, acute stress, and obsessive-compulsive disorders from the definition of anxiety (Park & Kim, 2020). The revisions found in the DSM-V (informed by molecular genetics and neuroimaging work) are not without scrutiny (Park & Kim, 2020)—not least because of the lack

of reliable biological markers and processes for anxiety disorders. However, as the focus of the work in this thesis concerns subclinical trait and state anxiety, we will leave the discussion for others to take up (see Park & Kim, 2020). Moreover, while a comprehensive review of the physiological and behavioural sequelae of clinical anxiety is outside the scope of this paper (for a detailed review see Grupe & Nitschke, 2013; Wiedemann, 2015), in what follows below, we present recent computational and neuroscientific efforts to elucidate individual differences in how anxiety changes cognition, focusing on pertinent neuroimaging and electrophysiological work from trait, state, and clinical studies.

1.5.2.2 Anxiety in the lab

Lab-based studies that use periods of shock and safety to induce a state of anxiety have consistently shown, supporting the adaptive account of anxiety, that anxiety sharpens sensory processing and creates a condition of hypervigilance and hyperarousal (Cornwell et al., 2017; Grillon et al., 2019; Robinson et al., 2013). These states are considered adaptive when performing sensory-perceptual based tasks, as anxiety can improve performance. These improvements to task performance under induced anxiety are well documented in the detection of threat (Grillon & Charney, 2011; Robinson et al., 2011) and when applying attention to task-relevant stimuli (Cornwell et al., 2011; Edwards et al., 2010; Kim et al., 2020). Yet it is important to remember that clinical levels of anxiety are maladaptive, disrupting lives with the overapplication of these hypervigilant threat-aware states to everyday settings (settings mostly absent of threat). Interestingly, the attentional bias to negative emotional stimuli like threats has been shown correlated with higher functional connectivity between the dorsal ACC, dorsomedial PFC, and the amygdala (Bishop & Forster, 2013; Etkin et al., 2011; Robinson et al., 2012).

At the neural level, aversive PEs from fearful faces relative to happy faces during lab-induced anxiety have been shown to correlate with increased fMRI activity in the VS (Schmitz & Grillon, 2012). A thought-provoking line of research might be exploring these threat-related biases and their neurobiological correlates in trait anxiety, similarly to those demonstrated by White et al. (2017). Using EEG and MEG, one reliable event-related deflection known to represent unexpected stimuli is the mismatch negativity (MMN, PEs in response to a violation in a series of established stimuli). Or in MEG studies, the magnetoencephalographic MMN (MMNm, see Garrido et al., 2009; Näätänen et al., 2011). A healthy MMN/m response thus represents the adaptive and efficient early identification of a change to the environment (fast orienting to potential threats, Grillon et al., 2019; Wessel & Aron, 2017). The MMN/m is a metric for sensory vigilance (Fan et al., 2018; Gené-Cos et al., 1999). Accordingly, atypical amplification

of the MMN/m is documented in several anxiety disorders, including panic disorder (Chang et al., 2015), PTSD (Ge et al., 2011; Morgan & Grillon, 1999), and phobia (Mager et al., 2001). And also in healthy individuals passing through a transient state of anxiety (Cornwell et al., 2007, 2017). As such, these brain-based results provide empirical support of healthy and patient findings detailing hypervigilant susceptivity to volatility (Grillon et al., 2019) and behavioural inhibition (Marshall et al., 2009).

Other studies have demonstrated alternative effects on performance in different cognitive tasks. For example, anxiety has been shown to improve long-term memory (Singh et al., 1979) but degrade short-term memory (Kalisch et al., 2006; Lavric et al., 2003; Shackman et al., 2006; Vytal et al., 2012). Anxiety is also documented to impede inhibition in emotional-based Stroop tests, where the aim is to identify the colour of a word and disregard semantic content (Richards & Millwood, 1989). Moreover, hypervigilance to threat can impede cognitive control processes such as response inhibition in tasks where the threat is uncoupled from the task (Roxburgh et al., 2020). This ability to control response inhibition was recently shown to be correlated with increases in beta band oscillatory activity in the right inferior frontal gyrus (IFG) of control participants, whereas anxious participants lacked this beta response (Roxburgh et al., 2020). In contrast to these disadvantages, anxiety can improve inhibition-based avoidance during experiments manipulating Pavlovian responses using a go–nogo task (Bishop & Forster, 2013; Grillon et al., 2017; Mkrtchian, Roiser, et al., 2017). Thus, anxious states appear to strike the balance between preparing responses for potential threats that may be wrong or inappropriate and having more effective responses that may be unnecessary—depending on whether Pavlovian responses align with instrumental processes (Grupe & Nitschke, 2013).

Generally, the literature describes a behavioural state of readiness, alertness, and attention to negative environmental input like threat. These behavioural responses are also found when testing anxious individuals who reach the criteria for clinical anxiety (Britton et al., 2013; Duits et al., 2015; Grillon et al., 2019; Grupe & Nitschke, 2013; Howlett et al., 2019; Huys et al., 2016; Maia et al., 2017; Robinson et al., 2013). An alternative to testing anxiety with shocks, however, is anxiety induced by a social stressor, like with a public speaking task (for example, the Trier Social Stress test, Kirschbaum et al., 1993)—sometimes referred to as psychosocial stress (see Grillon et al., 2019). These social stress tests have consistently been shown to produce both anxious and stress responses depending on how they are administered. While participants perform the TSST, stress is induced, confirmed by hormonal changes in cortisol—similar to predictable shocks (Hellhammer et al., 2009; Roelofs et al., 2007; Takahashi et al., 2005). But importantly, when social stressor tasks are used as an undefined anticipatory future threat of social stress, they are known to induce apprehension and worry consistent with

subjective, physiological, and neural changes associated with anxiety (Gorman & Sloan, 2000; Ionescu et al., 2013; Labuschagne et al., 2019). In **Chapters 2–4**, we use this latter psychosocial approach, showing its effectiveness in driving cardiovascular responses (Chalmers et al., 2014; Friedman, 2007) linked with anxiety, impairing reward-based learning, and altering computations of uncertainty thought vital to understanding both anxiety and theories of information-theoretic statistical inference (Friston et al., 2006; Hein et al., 2021; Sporn et al., 2020).

1.5.2.3 Diathesis-stress model

An important model of anxiety that explains how maladaptive avoidance and symptoms of anxiety are intensified by environmental stressors is the diathesis-stress model (Monroe & Simons, 1991). Evidence from subjective self-reports has recently been substantiated by behavioural and computational modelling work supporting the diathesis-stress model of anxiety. In particular, it has been shown that anxious participants display magnified Pavlovian inhibition responses toward aversive outcomes (Mkrtchian, Aylward, et al., 2017) and that higher learning rates about punishment are observed under stress from threat of shock, in spite of a lack of effect from acute stress (Valton et al., 2019). This evidence speaks to the recursive nature of anxiety, where trait anxiety levels repeatedly interact with environmental stressors that serve to exacerbate and increase anxiety (Ghosh et al., 2013)—potentially fitting the profile for anxiety disorders over time (Raymond et al., 2017). Others have noted that engaging with anxious thought processes and producing characteristic behavioural habits can reinforce the associated neural connections (Grupe & Nitschke, 2013). As such, these systems may be chronically activated and turned maladaptive in disorders such as PTSD, even when no threat or experimental anxiety manipulation is present (Grillon et al., 2019).

The relevance of the work included in this thesis is as follows. Just as anxious behavioural responses (e.g. maladaptive avoidance) can be learned and produce higher levels of anxiety (as avoidance tends to extinguish the opportunity to learn about safety), these maladaptive responses can also be unlearned, as demonstrated in therapy (Ishikawa et al., 2007; James et al., 2020; Moutoussis et al., 2018). Yet the way anxious individuals learn from and respond to reward-based signals, potentially in an effort to reduce anxiety, remains critically understudied—especially using a computational approach. The reason for this dearth in research is that anxiety is typically associated with difficulties processing and engaging with aversive stimuli, while depression is associated with difficulties in processing and engaging with reward (Bishop, 2007, 2008, 2009; Grillon et al., 2019; Grupe & Nitschke, 2013). Consequently, the utility of researching reward learning in anxiety, and its potential for

therapeutic use, has been overlooked. To address this, in **Chapters 2–5**, we focus solely on the effects of state anxiety and subclinical trait anxiety on reward-based learning in healthy participants in dynamic and uncertain learning environments. We hope to expand our understanding of how anxiety, uncertainty, and reward signals interact in healthy volunteers, with the ultimate aim of informing future treatments of anxiety and anxiety disorders (Moutoussis et al., 2018).

1.5.2.4 Anxiety and executive functioning

In contrast to the earlier mentioned adaptive benefits of anxiety on sensory-perceptual tasks, when performing executive functioning based tasks, anxiety typically vitiates learning, as with reward-based learning and economic decisions under risk (Hartley & Phelps, 2012). When faced with uncertain decisions, anxiety produces risk averse responses, selecting, for example, more predictable lower payments instead of uncertain higher payments (Charpentier et al., 2017; Maner & Schmidt, 2006) and biasing expectations to pessimistic negative outcomes (Borkovec et al., 1999; Butler & Mathews, 1987; Kim et al., 2020; Mitte, 2007; Stöber, 1997). While performing economic decision-making tasks, anxiety is known to interfere with and lead to suboptimal decisions (de Visser et al., 2010; Miu, Heilman, et al., 2008; Remmers & Zander, 2018). (For excellent reviews of the well-established effect of anxiety on decision-making, see Hartley & Phelps, 2012; Miu, Miclea, et al., 2008; Paulus & Yu, 2012).

To illustrate, Miu et al. (2008) used an Iowa Gambling Task (a multi-arm bandit task with different probabilistic contingencies for each arm, using both reward and loss feedback) to show that high levels of trait anxiety were linked with poorer decision making, selecting the disadvantageous choice more frequently than lower trait anxious participants. In that study, Miu et al. (2008) appeal to several explanations, including difficulties attending to and selecting relevant cues (perhaps looking instead to avoid potential negative outcomes, see Charpentier et al., 2017; Giorgetta et al., 2012), distraction by unrelated stress, and a possible influence of known correlated activity in the amygdala and its potential impact on decision making (making it more probable that anxious people will select the choice with reliable positive emotional feedback, see Gu et al., 2017; Xu et al., 2013).

In a later compelling and influential study, Browning et al. (2015) examined how high levels of trait anxiety influenced feedback learning from aversive outcomes in an uncertain learning environment. As detailed above in the Behrens et al. (2007) study, an ideal Bayesian learner will fit the learning rate (the strength of the reliance on each new outcome in influencing belief updates) to resolve changes in volatility, with higher estimated volatility driving higher learning

rates. The study by Browning et al. (2015) demonstrated that high trait anxiety was negatively correlated with the capacity to calibrate the learning rate to changes in the ground truth probability of receiving a shock. Importantly, in the Browning et al. (2015) study, volatility was between stable blocks, where the probability of one stimulus delivering the shock was 0.75 and the other 0.25, and volatile blocks, where the probability of a stimulus delivering a shock changed every 20 trials between 0.8 and 0.2. Additionally, this study showed that high trait anxiety was linked with diminished correlations between pupil responses and volatility relative to low trait anxious participants (a few seconds post-outcome, also demonstrated above in **Section 1.4.8.3**, see Aston-Jones & Cohen, 2005; Vincent et al., 2019; Zénon, 2019). A paper by Raymond et al. (2017) recently observed that dampened pupil responses to environmental change might reflect a concurrent inhibition of the central amygdala on the Edinger–Westphal nucleus (Lissek, 2012; White & Depue, 1999). An effect consistent with research in fear conditioning, where diminished modulatory activity in the ventromedial PFC that modulates the amygdala has been shown (Indovina et al., 2011; Raymond et al., 2017).

A later experiment by Pulcu & Browning (2017) supported the behavioural modelling results of Browning et al. (2015), presenting evidence for impaired learning rate adaptation, but concerning reward *loss*. One potential reason for this is that anxious learners may estimate a higher level of environmental change in all encountered environments (Bishop & Gagne, 2018). Aberrant adaptation to volatility in learning has also been documented in socially anxious people during an aversive learning task using angry and happy faces (Piray, Ly, et al., 2019). Also in that study, using fMRI, Piray et al. (2019) demonstrated that high trait socially anxious participants' dorsal ACC (dACC) activity did not (as with the low socially anxious group) covary with the learning rate. In other similar work using a multi-arm bandit task with dynamic reward and punishment probabilities, Aylward et al. (2019) provided strong evidence that unmedicated mood and anxiety disorder symptoms exhibited an increased punishment learning rate. This meant that anxious participants more rapidly learned about and updated their behaviour from negative feedback suggesting they assimilate aversive outcomes about threat across a lower number of trials (Aylward et al., 2019). The authors suggest this may indicate the overestimation of the probability of negative outcomes and also may produce avoidance behaviours.

More recently, inspired by these studies, our work in state anxiety and motor learning (Sporn et al., 2020) and experiments run by others on trait anxiety (Huang et al., 2017) have started to explore reward-based probabilistic learning where contingencies are unknown and may change throughout the experiment. In contrast to primary reinforcers like aversive shock, the effects of anxiety on learning from secondary reinforcers like monetary reward remain

unestablished, especially considering the unknown differences between state and trait anxiety. This line of reward-based enquiry has also begun to reveal that anxiety disrupts the learning of probabilistic contingencies and interferes with the estimation of uncertainty. Accordingly, this thesis seeks to more closely examine the largely overlooked probabilistic reward-based feedback learning in subclinical anxiety, especially in dynamic environments where volatility produces uncertainty about outcomes (**Chapters 2–5**). This may aid in our comprehension of suboptimal decision making in anxiety and might help to inform understanding of learning biases in clinical anxiety.

1.5.2.5 Biases in anxiety

One discrepancy worth further examination between the studies of sensory-perceptual processes (Cornwell et al., 2017; Grillon, 2008; Richards et al., 2014) and work in executive functions (Hartley & Phelps, 2012; Miu, Miclea, et al., 2008; Paulus & Yu, 2012) and our work in reward learning (Hein et al., 2021; Sporn et al., 2020) is the enhancement of either sensory input or prior/posterior beliefs. The aetiology of anxiety disorders has been explained using cognitive biases in learning (Grupe & Nitschke, 2013; Kim et al., 2020; Mathews et al., 1997; Mineka & Zinbarg, 2006) that are also thought to perpetuate anxiety (Nelson et al., 2010). These include expected value biases and impaired safety learning (Grupe & Nitschke, 2013). As mentioned above, sensory-perceptual based tasks show anxiety increases the processing of sensory input and leads to better performance on inhibitory and threat detection tasks (Cisler & Koster, 2010; Grillon, 2008). By contrast, higher-order executive functions typically report an impediment to task performance (decision making and probability learning, see Grupe & Nitschke, 2013; Miu, Heilman, et al., 2008; Remmers & Zander, 2018). In addition to these, a negative interpretation of ambiguous stimuli is also considered a core attribute of anxiety (Hartley & Phelps, 2012; Mathews & MacLeod, 2005). We speculate a potential bridge between these two accounts might be found in resistant and biased prior/posterior beliefs.

This is expressed as selectively attending towards threat (Richards et al., 2014), a stronger prior for expecting high levels of environmental change and potential threatening or aversive outcomes, and thus a response of increased vigilance and sensory processing to anticipate, detect and move away from threat (Bar-Haim et al., 2007; Wise et al., 2019). But also, this is expressed in inflexible and recalcitrant beliefs as shown in cognitive biases toward probability expectation and pessimistic negative outcome evaluation (Borkovec et al., 1999; Butler & Mathews, 1987; Kim et al., 2020; Mitte, 2007; Stöber, 1997). These judgment biases are linked with the neural circuitry of online expected value estimation (Grupe & Nitschke, 2013; Levy & Glimcher, 2012), and contextualise our work in the following **Chapter 2 and 5** on biased beliefs

in anxiety during reward learning, as these areas influence decision-making (Hein et al., 2021; Sporn et al., 2020). By contrast, earlier sensory cognition may represent more rapid primed perceptual biases, primarily in subcortical and posterior cortical circuits (Davis et al., 2010; LeDoux & Pine, 2016). The phylogenetic complexity of executive functions like decision making may produce more sophisticated biases such as confirmation bias and conservatism (Hilbert, 2012; Oswald & Grosjean, 2004; Tappin et al., 2017). Yet evidence implies biases like these are not reserved to clinical populations. Healthy individuals put in the same scenarios troubling those with anxiety disorders frequently show similar responses and biases (de Jong et al., 1998; Moutoussis et al., 2018; Pury & Mineka, 1997). Moreover, executive biases (confirmation bias) are connected to PFC processing and to genes linked with prefrontal dopaminergic function thought critical to efficient reward learning (Doll et al., 2011; Kappes et al., 2020)—a region also overlapping with neural changes in anxiety, to be discussed in **Section 1.5.3** below.

From a computational Bayesian PC perspective, these biased beliefs can exert influence over neural processing via two primary routes: i) inhibition from top-down predictions, and ii) postsynaptic gain (precision) regulation (Bauer et al., 2014; Brown & Friston, 2013; Ferguson & Cardin, 2020; Larkum et al., 2004). Relating this to anxiety and uncertainty, increased precision in expectations of adaptive factors like attention to threat may tune the postsynaptic gain of early sensory processing, modulated by both neurotransmitters and attentional processes (Bishop, 2008; Desimone & Duncan, 1995; Feldman & Friston, 2010; Kastner & Ungerleider, 2000). Alternatively, anxious individuals may have more precise priors and predictions that serve to functionally inhibit the processing of lower sensory PEs in order to maintain strong prior beliefs about the world (a resistance to update, Butler & Mathews, 1983, 1987; Mitte, 2007; Stöber, 1997). The devil, it appears, is in the detail of the contextual determinants driving perception and action.

To summarise, greater uncertainty about volatility and higher volatility estimates increase learning from sensory input (Iglesias et al., 2013; Lawson et al., 2017, 2020). Volatility alters decision making under anxiety in aversive feedback settings (Browning et al., 2015). Furthermore, anxious hypervigilance in volatile environments produces heightened stimulus-driven processing, optimising attention toward threat signals (Cisler & Koster, 2010; Cornwell et al., 2017; Robinson, Overstreet, et al., 2013; Shackman et al., 2011; van Marle et al., 2009). Inversely, when prior beliefs are estimated with more confidence, learning slows, potentially preserving safety (avoiding risk). In non-threatening environments, this may lead to descending predictions dampening learning about input. This highlights an important and still unresolved question about how different types of biases in anxiety may be linked with different

learning contexts and behaviours. However, one potential link to help explain these discrepancies could be avoidance behaviours.

1.5.2.6 Avoidance behaviours

Avoidance may be a particularly important component of anxiety that subserves both stronger prior beliefs about reward outcomes and the detection of threatening environmental cues. As anxiety produces an intolerance of uncertainty and potential negative outcomes and emotional and ambiguous stimuli are interpreted in a biased threatening way, a potential defence mechanism employed under anxiety is to avoid prospective catastrophe (Aylward et al., 2019; Bar-Haim et al., 2007; Borkovec et al., 2004; Dugas et al., 2005; Grupe & Nitschke, 2013; Indovina et al., 2011; Mathews & Mackintosh, 1998; Mitte, 2007; Mkrtchian, Aylward, et al., 2017; Wise et al., 2019). One illustration of avoidance behaviours is shrinking from social events for fear of embarrassment (Mkrtchian, Aylward, et al., 2017). Avoiding anticipated uncertainty and biases toward safety may play a critical role in the maintenance of anxiety disorders. A recursive loop may be formed where one increasingly avoids potentially harmful situations as the punishment learning rate increases. Thus a greater number of cues drive fear, and behaviours that would otherwise minimise fear are negatively reinforced (Raymond et al., 2017). Consequently, overestimating negative events accelerates, producing fewer opportunities to extinguish the fear experienced—prioritising avoidance at the cost of exploration and thereby exacerbating anxiety (Krypotos et al., 2015; Mkrtchian, Aylward, et al., 2017). Because of that, avoidance also contributes centrally to the psychological treatment of anxiety (Baum, 1970; Maner & Schmidt, 2006; Pittig et al., 2018). However, an often overlooked implication of safety-seeking biases in anxiety is the treatment seeking itself (Lorian & Grisham, 2011). We thus need to understand more about engaging (not avoiding) in anxiety; more about the motivational components of anxiety and how anxiety and motivation interact with learning from reward, not simply avoiding punishments. This we will explore in **Chapter 3** concerning the potential motivational influence of anxiety.

1.5.3 Neurobiological correlates of anxiety for reward and probability

Having closely examined the behavioural and some potential neural and computational effects of anxiety, we pause here to consider the imaging evidence of the neurobiological correlates of anxiety relating to probability biases. This comes from neuroimaging work using fMRI and electrophysiological work using both EEG and MEG.

Following early neuroimaging studies on emotion processing specifically in the amygdala (Morris et al., 1996; Vuilleumier et al., 2001; Whalen et al., 1998), more recent work has continued this by examining the role of anxiety in modulating amygdala activity and its connections (White et al., 2017). For several core attributes of anxiety, the amygdala has been shown to significantly contribute (Bishop, 2007; Grillon et al., 2019; Raymond et al., 2017). For example, for fear learning (Davis, 2000; Duvarci & Pare, 2014), when detecting threat (Davis, 1992), and for amplifying the processing of emotional and threatening information (Vuilleumier, 2005). Moreover, the amygdala is engaged when detecting novel information (Ousdal et al., 2014; Wessel et al., 2012). Of particular relevance is a study by Cornwell et al. (2007) showing the amygdala is engaged by auditory stimulus deviance (the MMNm) in a MEG based study where the deviance occurred in a similar temporal range to the inferior frontal gyrus. This experiment provides one excellent indication that electrophysiological research in state anxiety can inform clinical investigation. Inducing a state of anxiety in subclinical participants revealed the normative mechanisms (amplified MMNm and shifts between feedforward and feedback error transmission) necessary to provide a link to the potential processes that might also subserve anxiety pathology (Grillon et al., 2019).

The amygdala is thus crucially involved in shaping anxious responses. For this reason, variance in associative learning originating in the amygdala is thought to contribute to the generation of anxiety, pushing particular anxious behaviours, but also in the broader biasing of cognitive systems (Averbeck & Costa, 2017; LeDoux, 2014, 2015; Phelps et al., 2014). The amygdala circuitry strongly indicates the involvement of the PFC in anxiety (Bishop, 2009; Davidson, 2002; Davis, 1992; Rauch et al., 2003; Tye et al., 2011), connecting with higher-order learning and control processing. In recent research, dorsal PFC cortical thickness has been associated with anxiety and reinforcement learning (Abend et al., 2020), supporting earlier work on avoidance behaviours indicating a lower association between PE and decision making structures (ventromedial PFC, VS) while performing a reinforcement learning task (White et al., 2017). Also, anxiety is known to be associated with impoverished recruitment of the PFC (Bishop, 2009; Forster et al., 2015). Several studies have also argued that the amygdala/basolateral amygdala and the ACC are involved in the processing of uncertainty (Stolyarova et al., 2019; Whalen, 2007; Williams et al., 2015). Research using human fMRI, non-human primate single-cell recording, and rat lesion studies have all implicated the amygdala as contributing to volatility uncertainty estimation, and some more specifically argue for amygdala regulation to the learning rate (Averbeck & Costa, 2017; Costa et al., 2016; Holland & Gallagher, 1999; Homan et al., 2019; Li et al., 2011; Roesch et al., 2012).

Others, however, have argued for the wider inclusion of interconnected processing systems, from activity in the hippocampus to modulation by cholinergic sources, and on a wholly different scale: in amygdala molecular composition differences (Lissek, 2012; Raymond et al., 2017). Evidence from human depth electrode recordings has since indicated that ERP activity in the amygdala in response to fearful emotional faces comes in a time window of approximately 200 ms post-stimulus (Krolak-Salmon et al., 2004). And moreover, that other brain areas (occipitotemporal, anterior temporal, and OFC) also demonstrated the effect, however in a later window (300–1300 ms, see Krolak-Salmon et al., 2004). This result supports the involvement in distributed brain areas for the processing of fear in anxiety that overlaps with executive control areas. As mentioned in previous sections of this chapter, the neural systems linked with value estimation, the PFC and OFC, and also PE-based learning, the striatum, also overlap with reported atypical activity in probability biases and exaggerated estimates of threat cost in anxiety (Paulus & Stein, 2006; Paulus & Yu, 2012; Preuschoff et al., 2008; Schultz et al., 1997; Volz et al., 2003). This is typically from interference with aversive PE transmission leading to dysfunctional belief updates when negative outcomes fail to happen (Grupe & Nitschke, 2013; Paulus & Stein, 2006; Paulus & Yu, 2012).

Taken together, this collection of research provides a strong rationale to hypothesise that learning from experience in dynamic and uncertain environments is altered by temporary states of anxiety and the degree of self-reported trait anxiety (Etkin, 2010; Ridderinkhof et al., 2004; Walsh & Anderson, 2012). In the chapters to come, we test this hypothesis by conducting experiments on hierarchical Bayesian learning in anxiety while learning from reward signals (and not negative or punishment signals) in a volatile task environment.

1.5.4 Oscillations and anxiety

Neural oscillation patterns have also been proposed as one mechanistic motif of atypical processing in clinical disorders (Lopes da Silva, 2013; Uhlhaas & Singer, 2012). Exemplative cases include autism (David et al., 2016; Abigail Dickinson et al., 2015, 2018; Ronconi et al., 2020) and schizophrenia (Gonzalez-Burgos & Lewis, 2008; Roach & Mathalon, 2008; Uhlhaas & Singer, 2010). Anxiety involves a number of distinct brain regions, depending on situational factors, but primarily in the prefrontal/amygdala regions (Bishop, 2007; Grupe & Nitschke, 2013; Justin Kim & Whalen, 2009; Shiba et al., 2016; Simpson et al., 2001). The dorsal and lateral medial prefrontal cortex (dmPFC, dlPFC) and the dorsal anterior cingulate cortex (dACC) have also been shown to preserve anxiety through worry and recurrent analysis of threat (Davidson, 2002; Grupe & Nitschke, 2013; Robinson et al., 2019)—linking to rumination characterised by theta (4–8 Hz) and beta band (13–30 Hz) oscillations (Andersen et al., 2009;

Pavlenko et al., 2009). Notably, increased beta oscillations have also been associated with anxiogenic states through the type of fixity of thought observed in obsessive-compulsive disorders, ritualised behaviours, and rumination (Abramowitz et al., 2009; Hamilton et al., 2015; Lang et al., 2015; Williams, 2016)—with beta rhythms thought to maintain an inflexible status quo (Engel & Fries, 2010).

In addition, emotional appraisals during anxiety, such as the processing of fearful faces, are associated with gamma band (30–80 Hz) amygdala and visual cortex activity (Aftanas et al., 2003; Güntekin & Başar, 2014; Schneider et al., 2018), while increases in alpha waves and strengthened anticorrelations with delta (0.5–4 Hz) activity connect to high levels of anxiety and behavioural inhibition (Knyazev et al., 2002, 2004; Knyazev & Slobodskaya, 2003). Interestingly, several researchers have associated the amygdala with associative learning, potentially regulating the learning rate and more broadly uncertainty over environmental change (Averbeck & Costa, 2017; Holland & Gallagher, 1999; Homan et al., 2019; Li et al., 2011; Phelps et al., 2014; Roesch et al., 2012). As anxiety is an emotional response, this may potentially link reported anxiety-induced changes to volatility estimates (Browning et al., 2015) with amygdala modulations.

At any rate, taken together, these studies suggest a role for alpha/beta oscillations in maintaining an anxiogenic state. A recent study by Roxburgh et al. (2020) supported the involvement of beta oscillations by connecting an attenuation of beta activity in threat-induced anxious participants who over respond during a Go/No-Go task. In that MEG based study, deficient control of response inhibition in periods of anxiety was correlated to reduced beta activity in prefrontal areas (bilateral IFG and right dPFC). Although, in a study inducing state anxiety while analysing intracranial recordings in epilepsy patients, Lee et al. (2019) reported increases in beta oscillatory activity in limbic regions during periods of experimentally induced anxiety, also correlating with self-reported anxiety levels. Also, Lee et al. (2019) report safety-related cues dampen the beta oscillation response.

We are not aware of any studies examining the effect of anxiety on the neural oscillations associated with predictive learning from dynamic and uncertain reward learning environments. When applied to learning under uncertainty, beta frequencies might relate to the inflexible belief updating (Browning et al., 2015) and biasing of uncertainty estimates found in anxiety (Huang et al., 2017; Piray, Ly, et al., 2019; Pulcu & Browning, 2019; Sporn et al., 2020). In our recent paper (Sporn et al., 2020), we showed how state anxiety was associated with atypical increases in beta power and burst rate, which inhibit pwPEs during reward-based motor learning. Interestingly, the work by Roxburgh et al. (2020), Lee et al. (2019), and Sporn et al.

(2020) show a potential divide between threat induced hypervigilant states and precision over sensory processing, and executive functioning tasks and precision over prior/posterior beliefs. At present, data is sorely lacking on the role of beta oscillations in anxiety and how these relate to learning and sensory processing, especially in a Bayesian PC framework. In **Chapters 3 and 5**, we address the question of whether state anxiety and high levels of trait anxiety, respectively, are associated with changes in the spectral correlates of PE learning, as theorised by the PC process (Bastos et al., 2012). More specifically, we examine the effect of altered pwPEs about stimulus outcomes and predictions about reward outcome tendencies in anxiety, and how these differently modulate the neural oscillations of continuous EEG and MEG signals.

1.5.5 Summary and relevance

The decisions we make and how we learn from and interact with our environment is coloured by our emotions—a potential system that continuously shapes ongoing behaviour. Healthy emotions serve an adaptive function. Normal anxiety prepares an organism to survive when facing uncertain future threatening situations. Dysfunctional anxiety is normative anxious responses overapplied to unthreatening circumstances, transmuting adaptive to maladaptive. As an example, a pernicious feedback loop between avoiding predicted anxiety-inducing situations, negative emotions, and inference can be instantiated, particularly when activating cognitive biases (attention to threat, confirmation bias on false inference: see Moutoussis et al., 2018). Anxiety disorders disrupt everyday lives and impair overall health and wellbeing. However, even subclinical levels of anxiety can impede cognition and bring about intolerance to uncertainty. In the domain of learning and decision-making, recent work has shown how highly trait anxious individuals exhibit difficulties adapting their rate of learning (in response to unpredicted outcomes) to uncertain (volatile) environments in both aversive and reward-learning contexts. While others have demonstrated how anxiety brings about a state of sensory hypervigilance and hyperarousal, biasing attention toward threat stimuli and inhibition responses, and impairing the extinction of learned fear associations. Each provides evidence for a degree of fixity in beliefs or biased learning when experiencing anxiety, especially when interacting with uncertainty. More recent theoretical accounts also propound a link between affective disorders (like anxiety) and difficulties estimating uncertainty, linking pathologies of precision estimation in neuropsychiatric disorders to subclinical affective states like anxiety.

In **Chapter 2,** we expand the Bayesian PC process theory of anxious learning into the realm of state anxiety, asking whether the threat of a psychosocial stressor can influence ongoing decision-making behaviour by altering model-based estimates of uncertainty. We tested this

by inducing a transient anxious state in a subclinical population and measuring the effect on reward-based learning in a volatile environment. We demonstrate a relationship between temporary anxious states and the misestimation of uncertainty in model-based decision making, additionally reporting impaired overall reward-learning performance. Following up on the results in **Chapter 3,** we associated the model estimates of pwPEs and predictions to changes in EEG time-frequency estimates using convolution modelling. We therefore ask in **Chapter 3** whether a temporarily-induced state of anxiety in healthy human participants alters the neural oscillatory patterns associated with predicting and learning from rewards. We found that pwPEs and predictions were associated with increases in beta oscillations in our state anxious group suggesting beta oscillations as one candidate process for explaining misestimation of uncertainty and maladaptive reward learning in anxiety. In **Chapter 4**, we turn to examine the potential adaptive function of anxiety by coupling a reward-learning task to reducing anxiety. We thereby aimed to test the potential adaptive motivation of an anxious state to reduce uncertainty in a volatile reward-learning environment. Finally, in **Chapter 5**, we aimed to expand the results from **Chapters 2**–**3** by examining the effect of high levels of trait anxiety on model-based reward learning in a volatile environment. We asked whether subclinical anxiety alters model uncertainty estimates and vitiates learning as with state anxiety. Moreover, we sought to address the limitations of poor gamma frequency resolution in the EEG signal in **Chapter 3** by using MEG to resolve gamma frequency responses. We found that high trait anxiety, like state anxiety, impedes overall reward-learning performance. However, in contrast to state anxiety in **Chapter 2**, trait anxiety increases estimates of volatility, uncertainty about stimulus outcomes, and environmental uncertainty—producing more explorative responses. Further, in contrast to state anxiety in **Chapter 3**, pwPEs in trait anxiety decreased the alpha-beta response relative to low trait anxious participants.

Chapter 2: State Anxiety Biases Estimates of Uncertainty and Impairs Reward Learning in Volatile Environments

Hein, T. P., de Fockert, J., & Ruiz, M. H. (2021). State anxiety biases estimates of uncertainty and impairs reward learning in volatile environments. NeuroImage, 224, 117424.[6]

Author contributions M.H.R. and T.P.H. designed the experiment, T.P.H collected the data, M.H.R. and T.P.H. analysed the data, M.H.R., and T.P.H. wrote code for data analysis, M.H.R. performed model parameter simulations, M.H.R. and T.P.H. wrote the manuscript, J.D.F. edited the manuscript. Model parameter estimation/simulations and efficiency of β coefficients in the GLM model were performed by M.H.R. The authors would like to thank Lilian Weber for providing GLM code, for insightful discussions, and for detailed feedback on the manuscript, and Carsten Allefeld and Joram Soch for engaging in fruitful discussions about model comparison and the GLM.

Abstract

Clinical anxiety impairs decision making, and high trait anxiety interferes with learning. Less understood are the effects of temporary anxious states on learning and decision making in healthy populations, and whether these can serve as a model for clinical anxiety. Here we test whether anxious states in healthy individuals elicit a pattern of aberrant behavioural, neural, and physiological responses comparable with those found in anxiety disorders, particularly when processing uncertainty in unstable environments. In our study, both a state anxious and a control group learned probabilistic stimulus-outcome mappings in a volatile task environment while we recorded their electrophysiological (EEG) signals. By using a hierarchical Bayesian model of inference and learning, we assessed the effect of state anxiety on Bayesian belief updating with a focus on uncertainty estimates. State anxiety was associated with an underestimation of environmental and informational uncertainty about the reward tendency, and an increase in uncertainty about volatility. Anxious individuals' beliefs about reward contingencies were more precise (had smaller uncertainty) and thus more immune to updating, ultimately leading to impaired reward-based learning. State anxiety was also associated with greater uncertainty about volatility. We interpret this pattern as evidence that state anxious individuals are less tolerant to informational uncertainty about the contingencies governing their environment and more willing to be uncertain about the level of stability of the world itself. Further, we tracked the neural representation of belief update signals in the trial-by-trial EEG amplitudes. In control participants, lower-level precision-weighted PEs (pwPEs) about reward tendencies were represented in the ERP signals across central and parietal electrodes peaking at 496 ms, overlapping with the late P300 in classical ERP analysis. The state anxiety group did not exhibit a significant representation of low-level pwPEs, and there were no significant differences between the groups. Smaller variance in low-level pwPE about reward tendencies in state anxiety partially accounted for the null results. Expanding previous computational work on trait anxiety, our findings establish that temporary anxious states in healthy individuals impair reward-based learning in volatile environments, primarily through changes in uncertainty estimates, which play a central role in current Bayesian accounts of perceptual inference and learning.

Keywords: Anxiety, uncertainty, hierarchical Bayesian inference, computational modelling, precision-weighted prediction error, single-trial EEG.

1. Introduction

Anxiety is characterised by excessive worry about negative possibilities (Grupe & Nitschke, 2013). It can lead to distinct difficulties when making decisions and learning about the world, as anxious individuals experience negative reactions to uncertainty—known as intolerance of uncertainty (IU, Bishop, 2007; Carleton, 2016). Recent work has established that individuals high in trait anxiety have difficulties adapting their learning rate to changes in probabilistic task environments (Browning et al., 2015; Huang et al., 2017). Less understood is how temporary states of anxiety in healthy subjects interfere with optimal learning and belief updating in the brain. Identifying the computations that subserve learning under state anxiety is important due to the prevalence of highly anxious states in most real-world environments that are filled with uncertainty (Bach et al., 2011; Bishop & Gagne, 2018). In addition, these insights could expand our understanding of the mechanisms by which anxiety biases beliefs about the world, linking to anxiety-related disorders.

Previous computational work has identified three types of uncertainty during decision-making and learning: expected (irreducible) uncertainty, estimation (informational) uncertainty, and unexpected (environmental) uncertainty (Bland & Schaefer, 2012; de Berker et al., 2016; O'Reilly, 2013; Yu & Dayan, 2005). Expected uncertainty emerges from the probabilistic relationships between responses and their outcomes, which is an inherent (irreducible) property of most real-world interactions. Estimation uncertainty arises from the imperfect information about those response-outcome relationships and decreases with learning. Lastly, changes in the environment (volatility) induce unexpected environmental uncertainty. Subjective estimates of volatility should affect learning as the individual should be more willing to update their estimates in a world that is changing (Mathys et al., 2011). To reduce uncertainty, the brain is thought to appraise the inherent statistical structure of the world using probability distributions, continuously updating and inverting a hierarchical model of the causes of the sensory inputs it receives (de Lange et al., 2018; Doya et al., 2007; Friston, 2010; Rao & Ballard, 1999). In this context, each type of uncertainty is expressed by the width (variance, or its inverse, precision) of the probability distribution of the corresponding belief (Feldman & Friston, 2010; Mathys et al., 2011).

Examinations of belief, uncertainty, and precision estimates using Bayesian formulations in perceptual and learning tasks are increasingly used to provide mechanistic explanations for an array of neuropsychiatric conditions. Specifically, difficulties estimating precision have been suggested to explain various clinical expressions, from movement difficulties in Parkinson's

disease to features of schizophrenia and autism (Friston et al., 2016; Lawson et al., 2014, 2017; Parr, Rees, et al., 2018). In the case of anxiety, altered beliefs are also theorised to play a vital role (Paulus & Stein, 2010; Paulus & Yu, 2012). As anxiety relates to worry over uncertainty, volatile task environments have been used to understand how trait anxiety affects learning, providing a mechanistic account of anxiety-related disorders (Browning et al., 2015; Huang et al., 2017). Healthy individuals are known to adapt their learning rate to volatility, with changing environments promoting a higher learning rate as new information needs to be integrated to better predict the future (Behrens et al., 2007). By contrast, high-trait anxious individuals show reduced adaptability of their learning rate to changes in volatility, both in aversive (Browning et al., 2015) and reward settings (Huang et al., 2017). Moreover, they show poorer performance in decision-making tasks (de Visser et al., 2010; Miu, Heilman, et al., 2008). Expanding on those findings, here we evaluated whether temporary anxious states in healthy individuals influence reward learning in a volatile environment through changes in informational and environmental uncertainty. Evidence for a link between anxiety and inaccurate estimation of uncertainty would lend support to recent theoretical accounts suggesting that difficulties learning from incomplete information and misestimations of uncertainty are crucial to understanding affective disorders (Pulcu & Browning, 2019).

Probabilistic inference has been proposed to be achieved through the sequential use of Bayes' rule, by dynamically combining our predictions (prior beliefs) with new evidence (sensory data) and weighting each resultant prediction error (PE) according to its precision (Feldman & Friston, 2010; Friston & Kiebel, 2009a; Kok & de Lange, 2015). This predictive coding scheme relies on a specific message passing between regions of a cortical hierarchy (Bastos et al., 2012; Iglesias et al., 2013; Rao & Ballard, 1999). Predictions are transmitted down the cortical hierarchy (backwards) to meet incoming ascending (forward) sensory PEs thought to arise in supragranular layers in superficial pyramidal cells (Friston & Kiebel, 2009a). Beliefs are then updated by reducing PE signals across each level of the cortical hierarchy, with both priors and PEs weighted according to their estimated precision (Kok & de Lange, 2015). Importantly, developments in Bayesian computational modelling allow us to estimate inter-individual differences in the trial-wise computations and expression of these precision-weighted PEs (Mathys et al., 2011, 2014).

Monkey single-cell recording and human functional magnetic resonance imaging (fMRI) studies have shown that PEs elicited by reward are encoded by phasic responses in midbrain dopamine neurons, and these signals are conveyed to the medial frontal cortex (MFC, Chew et al., 2019; Matsumoto et al., 2007; Morris et al., 2006; Zarr & Brown, 2016). Using electroencephalography (EEG), these reward learning signals can be detectable in the error

related negativity (ERN), an event-related potential (ERP) triggered by overt errors around 100 ms; and the feedback ERN (fERN) that follows negative feedback around 250 ms (Holroyd et al., 2003; Montague et al., 2004; Nieuwenhuis et al., 2004; Yeung et al., 2005). Both components have been shown to originate in the posterior medial frontal cortex (pMFC, including the anterior cingulate cortex, ACC, Holroyd et al., 2003; Montague et al., 2004; Yeung et al., 2005). Relevant to our study, the fERN has been proposed to index the magnitude of prediction violation (surprise), thus reflecting a reward PE signal that can be estimated, for instance, by using reinforcement learning models (Gehring & Willoughby, 2004; Holroyd & Coles, 2008; Holroyd & Krigolson, 2007). Another component of the ERP that may be sensitive to reward PEs, valence and surprise is the P300 (peaking between 250 - 500 ms with a parietal topography, Hajcak et al., 2005, 2007; Polich, 2007; Wu & Zhou, 2009).

Evidence linking PEs, Bayesian surprise, and belief updating to trial-wise fluctuations in ERP responses comes from studies combining computational modelling and analysis of trial-wise EEG responses (Diaconescu, Litvak, et al., 2017; Jepma et al., 2016; Kolossa et al., 2015; Mars et al., 2008; Stefanics et al., 2018; Weber et al., 2020). For instance, recent EEG studies on the MMN were able to spatiotemporally dissociate lower-level precision-weighted PE (pwPE) signals, which drive updates in belief estimates (Stefanics et al., 2018), and higher-level pwPEs, driving volatility updates (Weber et al., 2020). In addition, model-based single-trial analyses of the P300 identified the earlier P3a waveform of anterior distribution as an index of belief updating, whereas Bayesian surprise was represented in the later posterior P3b component (Kolossa et al., 2015). Here we were interested in assessing the neural representation of pwPEs across different levels, including lower-level pwPEs used to update reward tendency estimates, and higher-level pwPEs used to update volatility estimates, as belief updates on these two levels both depend on informational and environmental uncertainty. We therefore aimed to examine the effect of these two hierarchically-related pwPEs on brain activity by analysing trial-wise ERP responses across frontal, central, and parietal brain regions, and within a broad temporal range from 200 to 850 ms, encompassing the fERN and extended P300 components.

To address our questions, we examined cortical dynamics in a state anxious group and a control group using EEG recordings during a reward-based learning task. To link anxiety-induced neural changes to potential alterations in uncertainty estimation, we used a Bayesian model of perception and learning, the Hierarchical Gaussian Filter (HGF, Mathys et al., 2011, 2014). The HGF estimates individual trajectories of trial-wise belief updates governed by hierarchically related PEs based on the behavioural responses of participants. To reveal the effect of hierarchical PEs and precision weights on evoked brain responses, we used the

relevant computational quantities (pwPEs) as regressors in a general linear model (GLM) of trial-wise EEG amplitudes—as done in previous studies (Diaconescu et al., 2014; Diaconescu, Litvak, et al., 2017; Weber et al., 2020).

2. Material and Methods

2.1 Participants

Forty-two healthy individuals (age 18–35, 28 females, mean age 27, standard error of the mean [SEM] 0.9) participated in this reward-based learning study following written informed consent. This experiment was approved by the ethics review committee at Goldsmiths University of London. Our sample size was informed by previous computational work on anxiety (Browning et al., 2015). All participants were healthy volunteers, with no past neurological or psychiatric disorders.

All participants were screened using Spielberger's Trait Anxiety Inventory (STAI, Spielberger, 1983) which has reliably demonstrated internal consistency and convergent and discriminant validity (Barnes et al., 2002; Spielberger, 1983a; Spielberger et al., 1970). Scores on this trait inventory range from low (20) to high anxiety (80). Participants were measured for their trait anxiety level (mean in the total sample was 46, SEM 1.5) and then split into two groups using the median value (43). The sample population range was between 34 and 68 (Low trait = 34–42, High trait = 43–68). This created a high and low trait anxiety sample to then pseudo-randomly draw from to create the experimental (state anxiety, StA) and control groups (Cont). The mean trait anxiety score in the StA group was 47 (SEM = 2.1), while it was 46 (SEM = 2.2) in the Cont group. Importantly, individual trait anxiety levels did not exceed the clinical level (> 70, a cutoff score provided by the authors who developed the Spielberger STAI scale corresponding with the mean and 2SD of the average score for adults, see: Spielberger, 1983a).

Taken together, the ages (mean 27.7, SEM = 1.2) and sex (13 female, 8 male) of the Cont group were commensurate with those from StA (mean 27.5, SEM = 1.3, sex 14 female, 7 male). This demonstrates that no age or sex-related confounds are present for subsequent analysis. This is important in the light of documented age and sex-related effects on heart-rate variability (HRV: see Voss et al., 2015), which we used to assess physiological changes due to state anxiety.

2.2 Experimental Design

We used a between-subjects experimental design with state anxiety being the between-subject factor (StA and Cont groups). Both groups completed our experimental task, which consisted of four blocks: resting state 1 (R1: baseline), reward-learning task block 1 (TB1), reward-learning task block 2 (TB2), and resting state 2 (R2; see **Supplementary Figure 1**). Both resting state blocks were 5 minute-long recordings of EEG and electrocardiography (ECG) with eyes open. After R1, participants conducted a binary choice decision-making task with contingencies that changed over the course of learning as in previous work (Behrens et al., 2007; de Berker et al., 2016; Iglesias et al., 2013). In our task, participants completed two blocks of 200 trials each (TB1, TB2), and their goal was to find out which one of two visual icons (always either blue or orange: see **Figure 1**) would lead to a monetary reward (positive reinforcement, 5 pence). Thus, they had to learn the probability of reward assigned to each stimulus (reciprocal: p, 1−p). Both experimental blocks were divided into 5 segments with different stimulus-outcome contingency mappings that were randomly ordered for each participant and varied in length between 26 and 38 trials. These contingencies ranked from being strongly biased (90/10), moderately biased (70/30), to unbiased (50/50), and repeated in reverse relationships (10/90; 30/70) so that over the two blocks there were 10 contingency blocks in total (de Berker et al., 2016).

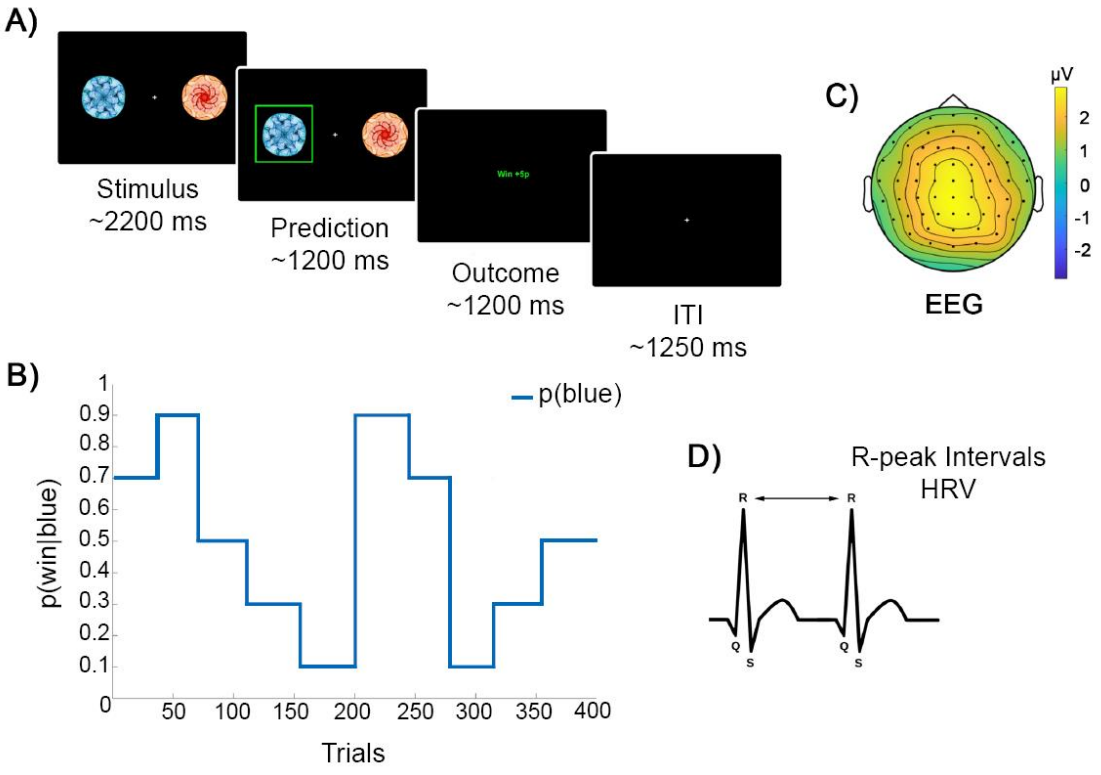


Figure 1. Behavioural task structure and physiological measures. **A)** On individual trials, participants were presented with two visual icons. They were instructed to predict the rewarding stimulus (win = 5p). The stimuli (blue or orange image) were randomly presented to either the left or right of the screen. They remained on the screen until a response was given or the allowed time (2200 ms ± 200 ms) expired—recorded as no-response. When a response of either the left arrow key or right arrow key was pressed, participants immediately saw their chosen image highlighted in bright green, which remained on screen for about 1200 ms (±200 ms) before the outcome was revealed. The outcome, either win or lose, was shown in the middle of the screen for 1200 ms (±200 ms) in green and red respectively. Each trial ended with a fixation cross and an inter-trial interval of 1250 ms (±250 ms). **B)** The probability governing the likelihood of the blue stimulus being rewarded (p(win|blue), with reciprocal probability values for the orange stimulus: p(win|orange) = 1 −p(win|blue)). Probability mappings varied in length (26–38 trials) ranging from heavily biased, p(win|blue) = 0.9, through moderately biased, 0.7, to unbiased, 0.5; and repeated in reverse relationships (0.1, 0.3). Here we display one example of contingency changes for p(win|blue) over the course of the experimental blocks (TB1, TB2, 200 trials each). These blocks were divided into the 5 randomly ordered stimulus-outcome mappings and were randomly generated for each participant. While conducting the experimental task, participants' physiological responses **C)** EEG and **D)** ECG were recorded continuously, with R-peaks from ECG signals being used to calculate heart-rate variability (HRV) and spectral estimates of high frequency (0.15–0.4 Hz) power in HRV.

On individual trials, participants were asked to predict which of the two visual icons was going to reward them with money. Successful predictions were rewarded 5p, while unsuccessful predictions and no-responses were regarded as losses with 0p reward (**Figure 1**). The stimuli were either presented to the left or right of the centre of the screen randomly. They remained on the screen until a response was given or the prediction time (2200 ms ±200 ms) expired. When a response of either the left arrow key or right arrow key was pressed, participants immediately saw their chosen image highlighted in bright green, which remained on screen for 1200 ms (±200 ms) before the outcome was revealed. The outcome, either win or lose, was shown in the middle of the screen for 1200 ms (±200 ms) in green and red respectively. Each trial ended with a fixation cross and an inter-trial interval of 1250 ms (± 250 ms).

The participants were given full computerised instructions for each element of the experiment, including questionnaires. Each questionnaire came with written instructions and was responded to using the numerical keyboard buttons. Just before 10 practice trials of the same probabilistic reward-learning task used in the main experiment, participants were explicitly informed that the reward structure would change throughout the task and that they needed to adjust their predictions in response to inferred changes (de Berker et al., 2016). Importantly, directly after practice trials but before TB1, both the state anxiety and the control groups were informed that this experiment was, in fact, an examination of performance in two subsequent tasks: reward-learning and an additional presentation task (see next section). Their

instructions with regard to the additional task were, however, different as we aimed to induce state anxiety during the blocks of reward-based learning in the state anxiety and not in the control group.

2.3 State Anxiety Manipulation

Participants in the StA group were informed that they had been randomly selected to complete a public speaking task after finishing the reward-learning task (Feldman et al., 2004; Lang et al., 2015). They were told they would be required to present a piece of abstract art and would be allowed to prepare for 3 minutes for a 5 minute public presentation of this artwork to a panel of academic experts. Those in the control (Cont) group were instead informed that they would be given a piece of abstract art, but they were to give a mental description of it for the same time privately to themselves (instead of a panel of experts). After completing the reward-based learning blocks, StA participants were informed of the sudden unavailability of the assessment panel and were ultimately instructed to describe the artwork privately in line with the Cont group.

2.4 EEG and ECG Recording and Pre-Processing

EEG and ECG signals were recorded throughout all task blocks (R1, TB1, TB2, and R2) using the BioSemi ActiveTwo system (64 electrodes, extended international 10–20) with a sampling rate of 512 Hz. The EEG signals were referenced to the average between two electrodes affixed to the left and right earlobes. Four additional external electrodes in a bipolar configuration were used, which included two electrodes positioned to capture vertical and horizontal eye-movements (EOG), one to the zygomatic bone of the right eye, and one to the glabella (between both eyes); and two electrodes to record the ECG. ECG electrodes were placed in a two-lead configuration (Moody & Mark, 1982) calibrated to fit the Einthoven triangle (Wilson et al., 1931). All electrodes used highly conductive bacteriostatic Signa gel (by Parker Laboratories, Inc., 4 Sperry Road. Fairfield, NJ 07004 USA). All events, including presentation of stimuli, participant responses, and trial outcomes, were recorded in the EEG file using event markers.

Analysis of the ECG data was conducted in MATLAB (The MathWorks, Inc., MA, USA) using the FieldTrip toolbox (Oostenveld et al., 2011) and their recommended procedure to detect

the cardiac events. [7] Following this approach, the ECG signal was used to detect the QRS-complex and its main peak, the R wave peak. Next, we extracted the latency of the R-peak, which was used to compute the coefficient of variation (CV = standard deviation/mean) of the difference intervals between consecutive R-peaks (inter-beat interval: IBI). The CV of inter-beat intervals was used as a metric of heart rate variability for statistical testing and is termed HRV hereafter. This measure was recently shown to capture block-wise state anxiety changes using a similar manipulation, as validated by corresponding changes in state anxiety scores (Sporn et al., 2020). See further details below in Section *Measures of Anxiety*.

EEG data were preprocessed in EEGLAB toolbox (Delorme & Makeig, 2004) by first high-pass filtering at 0.5Hz (hamming windowed sinc finite impulse response [FIR] filter, 3381 points) and then notch-filtering between 48–52Hz (847 points) to remove power line noise. Afterwards, artefacts (eye blinks, eye movements, cardiac artefacts) were classified using independent components analysis (ICA, runICA algorithm) and removed (on average 2.3, SEM 0.16, components). Noisy channels were corrected utilising spherical interpolation. All signals were then epoched around outcome onsets (win, lose) from −200 to 1000 ms. Noisy epochs exceeding +/−100μV were identified and removed using a thresholding technique relative to the pre-stimulus baseline. The number of rejected trials for each participant did not exceed 10% of the total number. Additional processing steps related to the use of a General Linear Model in combination with the regressors extracted from the computational model are presented in the below section on EEG analysis and the general linear model.

Cleaned EEG and preprocessed behavioural data files are available in the Open Science Framework Data Repository: https://osf.io/b4qkp/. The results shown in Figures 3, 4, and 5 are based on these data.

2.5 Measures of State Anxiety

One marker of state anxiety used during the experiment was the CV of the inter-beat intervals to assess HRV, as this measure, similarly to other metrics of HRV, has been reported to show reductions during anxious states (Chalmers et al., 2014; Friedman & Thayer, 1998; Gorman & Sloan, 2000; Kawachi et al., 1995). A lower HRV is associated with complexity reduction in physiological responses to stress and anxiety (Friedman, 2007; Gorman & Sloan, 2000), and

[7] http://www.fieldtriptoolbox.org/example/use_independent_component_analysis_ica_to_remove_ecg_artifacts.

is used as a transdiagnostic marker to identify anxiety in psychiatry (Quintana et al., 2016). In our recent work, we validated the use of the CV-based HRV as a proxy for state anxiety by showing that a similar experimental manipulation reduced this HRV index and increased state anxiety scores (Sporn et al., 2020).

Complementing the HRV analysis, we acquired subjective self-reported measures of state anxiety (STAI state scale X1, 20 items: Spielberger, 1983a) four times throughout the experiment using an online version that was embedded within the code for the experiment. However, due to an error in the code, the STAI was presented at the wrong time intervals, rendering it invalid to assess changes in state anxiety following the experimental manipulation. To address this limitation, an additional analysis on the spectral characteristics of the inter-beat-interval data was performed to link our HRV proxy of state anxiety to autonomic modulation and parasympathetic (vagal) withdrawal (Friedman, 2007; Gorman & Sloan, 2000). Reduced high-frequency HRV (0.15–0.40 Hz) and reduced variation between R-R intervals are consistently shown across trait anxiety, worry, and anxiety disorders (Aikins & Craske, 2010; Friedman, 2007; Fuller, 1992; Klein et al., 1995; Miu et al., 2009; Mujica-Parodi et al., 2009; Pittig et al., 2013; Thayer et al., 1996). After obtaining the IBI time series, as described in the previous section, we interpolated it at 1 Hz with a spline function (order 3), with spectral power estimated using Welch's periodogram method (Hanning window: following Rebollo et al., 2018). These power estimates were then normalised to the average power in the baseline (R1) and converted to decibels (dB) for statistical analysis.

2.6 The Hierarchical Gaussian Filter (HGF)

We used the Hierarchical Gaussian Filter (HGF) from Mathys et al. (2011, 2014) to estimate each participant's individual learning characteristics and belief trajectories during our binary reward-learning task. The HGF is a freely distributed open source software available in TAPAS (http://www.translationalneuromodeling.org/tapas), and has been used to model and understand learning across diverse settings (e.g., de Berker et al., 2016; Diaconescu et al., 2014; Diaconescu, Mathys, et al., 2017; Iglesias et al., 2013; Marshall et al., 2016; Stefanics et al., 2018; Weber et al., 2020).

Alternative models to the HGF have been proposed based on a generative model of sudden changes in the environment (change-point models: Moens & Zénon, 2019; Nassar et al., 2010). In our task, changes to the contingencies governing the outcomes were abrupt (see **Figure 1B**), which is in contrast to the generative model of the environment suggested by the HGF, where states evolve as Gaussian random walks and thus change slowly and diffusively

over time. While the HGF has been successful in explaining and predicting human behaviour in such tasks (e.g., de Berker et al., 2016; Iglesias et al., 2013), alternative models (change-point models: Nassar et al., 2010; Hierarchical Adaptive Forgetting Variational Filter: Moens & Zénon, 2019) were formulated to expect sudden changes and could outperform the HGF in environments with diffuse or sudden changes. In practice, however, both approaches (HGF and change-point models) can successfully deal with both kinds of environments (sudden versus diffuse changes), as a recent comparative analysis found (Marković & Kiebel, 2016).

The HGF is a generative model representing an approximately Bayesian observer estimating hidden states in the environment. As such, the HGF is a model of perception where beliefs about states are updated hierarchically. This perceptual model can then be coupled to a response model that associates belief estimates to decisions. More specifically, in the generative model, a sequence of hidden states $x_1^{(k)}, x_2^{(k)},..., x_n^{(k)}$ gives rise to sensory inputs that each participant encounters across *k* trials. Notably, while the perceptual model specifies how inference from observations to beliefs operates hierarchically across those environmental states, the response model probabilistically generates responses (in our case, the choices of the agents) based on those beliefs (see **Figure 2**).

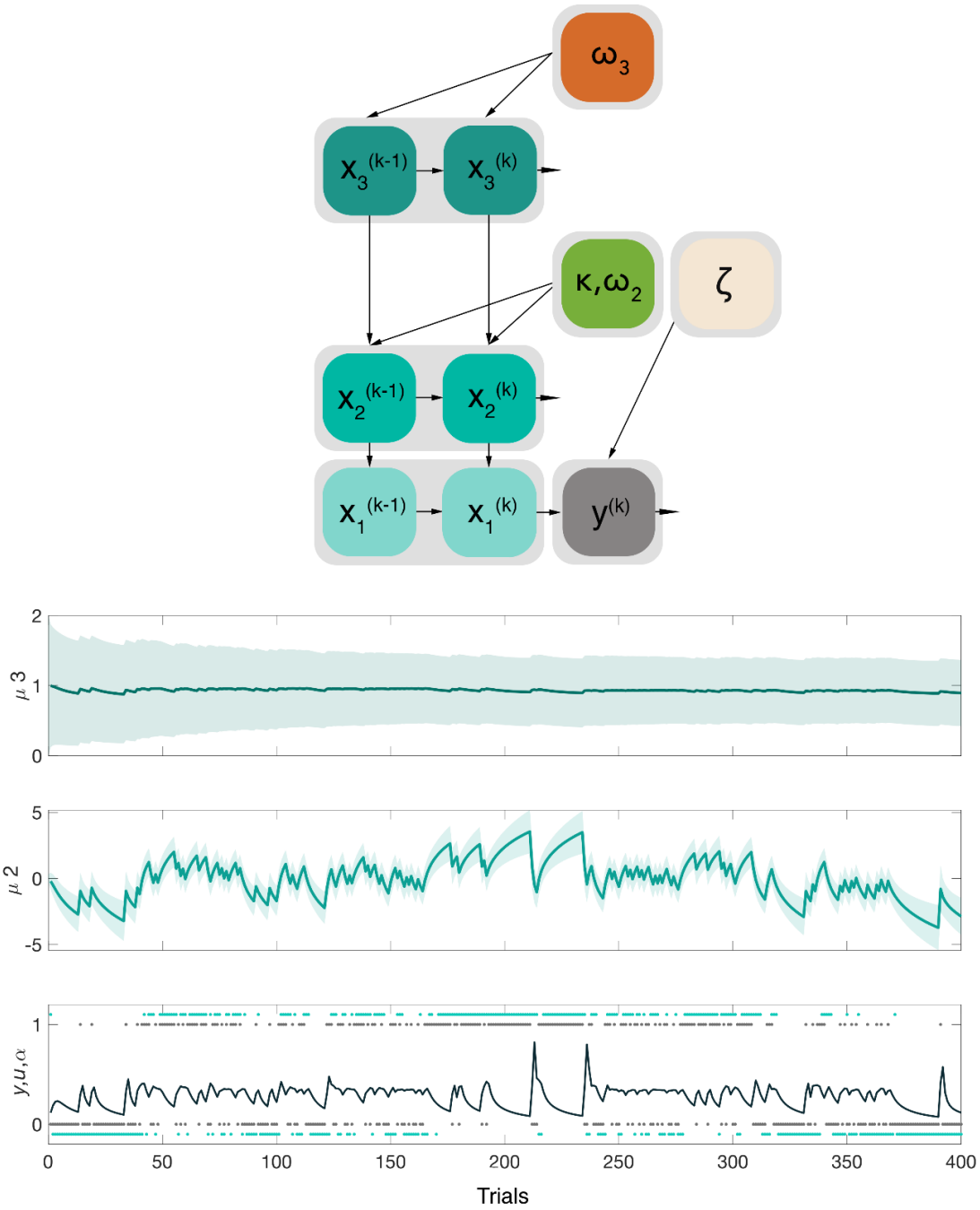


Figure 2. Three levels binary Hierarchical Gaussian Filter for binary outcomes. **Bottom panel.** Representation of the three levels of the HGF for binary outcomes and the associated belief trajectories across the total 400 trials in a representative participant. At the lowest level, the inputs *u* correspond to the rewarded outcome of each trial (1 = blue, 0 = orange; shown as black dots). The participant's responses *y* are shown in light blue dots tracking those trial outcomes. The learning rate (α) about stimulus outcomes at the lowest level is also given in black. The belief on the second level, μ_2 (σ_2), represents the participant's estimate of the stimulus tendency x_2 and the step size or variance of the Gaussian random walk for x_2 depends on parameters κ and ω_2, in addition to the estimates of the level above, x_3. The belief on the third level, μ_3 (σ_3), represents estimates of volatility x_3, whose step size is governed by parameter ω_3. **Top panel.** Schematic representation of the 3-level HGF model with relevant parameters modulating each level. In our study, ω_2, ω_3 and the response parameter zeta ζ were free parameters and were estimated by fitting the HGF to the individual responses and observed

inputs. Generally, parameters ω_2, ω_3 describe an individual's learning motif (see the section below for further details).

We used a 3-level HGF model for binary outcomes, where observed contingencies were used as input (Mathys et al., 2011, 2014). Hence, the trial-wise input $u^k = 1$ if the blue stimulus was rewarding (or orange lose) and $u^k = 0$ if the blue stimulus was not rewarding (or orange stimulus win). Note that all equations of relevant HGF quantities presented below are taken from Mathys et al. (2011, 2014). We refer the interested reader to these papers for the derivation of the perceptual model. At the lowest level of the model, the hidden state x_1 corresponds to the binary categorical variable of the experimental stimuli: whether the blue symbol is rewarded in trial *k* ($x_1^{(k)} = 1$; hence, orange would be non-rewarding) or not rewarded ($x_1^{(k)} = 0$; orange is rewarded). The second and third level states, x_2 and x_3, are continuous variables evolving as Gaussian random walks coupled through their variance (inverse precision). Thus, their value at trial *k* will be normally distributed around their previous value at trial *k-1.* The posterior distribution of beliefs about these true hidden states x_i (i = 2, 3) is fully determined by the sufficient statistics μ_i (mean, corresponding with participants' expectation) and σ_i (variance or uncertainty).

State x_2 describes the true value of the tendency of the stimulus-outcome contingency. It can be mapped to the probability of the binary state x_1 through a Bernoulli distribution $p(x_1 \mid x_2) = \text{Bernoulli}(x_1; s(x_2))$, where s(x) is a sigmoid function $s(x) = 1/(1 + \exp(-x))$. We can then measure the change in expectation at the lowest level and interpret it as an implied learning rate (α). This is defined as the sigmoid transformed difference between μ_2 before seeing the input and after seeing it, relative to the difference between the observed inputs u and its prediction $s(\mu_2)$ (**Figure 2,** lower panel; TAPAS toolbox: tapas_hgf_binary.m). A larger belief update in response to the same observed mismatch between the input u and the prediction amounts to a higher learning rate α. At the top level, x_3 represents the phasic log-volatility within the task environment; that is, the rate of change on the second level.

The coupling between levels 2 and 3 is through a positive (exponential) function of x_3, which represents the variance or step size of the Gaussian random walk that determines how x_2 evolves in time:

$$x_2^{(k)} \sim N\left(x_2^{(k-1)}, \exp(\kappa x_3 + \omega_2)\right) \quad (9)$$

The parameters κ and ω_2 represent the coupling strength and the tonic volatility, respectively. In the associated belief updates, momentarily high volatility estimates (μ_3) increase the speed with which beliefs at level 2 change. Larger values of the tonic (time-invariant) part of the variance (ω_2) generally increase the step size of x_2. This leads to faster belief updates on level 2 irrespective of current levels of (estimated) volatility.

The step size of the volatility state x_3 in our 3-Level HGF is governed by a positive constant, which is the exponential of a constant parameter ω_3:

$$x_3^{(k)} \sim N\left(x_3^{(k-1)}, \exp\left(\omega_3\right)\right) \tag{10}$$

Our analyses of uncertainty estimates focused on informational uncertainty, captured by variance on level 2 (σ_2, belief uncertainty about outcome tendencies) and level 3 (σ_3, belief uncertainty about volatility representing imperfect knowledge about how the reward outcome contingencies are changing across time: Mathys et al., [2014, Eq. 9–10]). Uncertainty about x_2 can be split into two distinct forms of uncertainty (informational uncertainty σ_2, and environmental uncertainty [$\exp(\kappa\mu_3^{(k-1)} + \omega_2)$]), whereas uncertainty about x_3 consists of σ_3. (Note that σ_3 corresponds to "informational" uncertainty about state x_3). Environmental uncertainty (Mathys et al., 2014, Eq. 13–14), determines the step of the random walk for x_2 through a combination of two quantities: phasic volatility ($\mu_3^{(k-1)}$) and tonic volatility (ω_2):

$$\exp\left(\kappa\mu_3^{(k-1)} + \omega_2\right) \tag{11}$$

Formally, the update equations of the posterior estimates for level *i* (*i* = 2 and 3) take the following form:

$$\Delta\mu_i^k = \mu_i^{(k)} - \mu_i^{(k-1)} \propto \frac{\hat{\pi}_{i-1}^{(k)}}{\pi_i^{(k)}}\,\delta_{i-1}^{(k)} \tag{12}$$

Where the posterior mean update term $\Delta\mu_i^k$ is the difference between the posterior expectation in the current trial, $\mu_i^{(k)}$ and the prediction from the previous trial $\mu_i^{(k-1)}$ before seeing the input on the current trial. The update step is proportional to the prediction error $\delta_{i-1}^{(k)}$ term, which denotes the discrepancy between the lower level expectation μ_{i-1}^k and the prediction $\hat{\mu}_{i-1}^{(k)}$.The prediction error is then weighted by a ratio of precisions (the precision of the prediction of the

level below $\hat{\pi}_{i-1}^{(k)}$, before seeing the input; and divided by the precision of the current belief $\pi_i^{(k)}$). Precision is defined as the inverse variance of the posterior expectation:

$$\pi_i^{(k)} = 1/\sigma_i^{(k)} \tag{13}$$

The precision-weights ratio in **Eq. 4** can be interpreted as a learning rate, whereas its product with the prediction error constitutes the precision-weighted prediction error that governs the update steps (pwPE: see also **Eq. 19** and **20** below). Correspondingly, **Eq. 4** above articulates the idea that more uncertain (less precise) belief estimates for the current level should motivate a larger influence of unpredicted outcomes on subsequent belief updating.

As mentioned above, the updates on the first level of our model are equivalent to the input $u^{(k)}$:

$$\mu_1^{(k)} = u^{(k)} \tag{14}$$

While the posterior belief updates on level 2 of our 3-level HGF take the form:

$$\mu_2^{(k)} = \mu_2^{(k-1)} + \sigma_2^{(k)}\,\delta_1^{(k)} \tag{15}$$

With the variance update as follows:

$$\sigma_2^{(k)} = \frac{1}{1/\hat{\sigma}_2^{(k)} + \hat{\sigma}_1^{(k)}} \tag{16}$$

Where the following definitions apply:

$$\delta_1^{(k)} \stackrel{\text{def}}{=} \mu_1^{(k)} - \hat{\mu}_1^{(k)} \tag{17}$$

$$\hat{\sigma}_1^{(k)} \stackrel{\text{def}}{=} \hat{\mu}_1^{(k-1)}\left(1 - \hat{\mu}_1^{(k-1)}\right) \tag{18}$$

$$\hat{\sigma}_2^{(k)} \stackrel{\text{def}}{=} \sigma_2^{(k-1)} + e^{\kappa\mu_3^{(k-1)} + \omega_2} \tag{19}$$

Formulated as precision, the variance step from **Eq. 8** above is:

$$\pi_2^{(k)} = \hat{\pi}_2^{(k)} + \frac{1}{\hat{\pi}_1^{(k)}} \tag{20}$$

A similar form is found for the belief update on level 3:

$$\mu_3^{(k)} = \mu_3^{(k-1)} + \sigma_3^{(k)} \frac{\kappa}{2} w_2^{(k)} \delta_2^{(k)} \tag{21}$$

$$\pi_3^{(k)} = \hat{\pi}_3^{(k)} + \frac{\kappa^2}{2} w_2^{(k)} \left(w_2^{(k)} + r_2^{(k)} \delta_2^{(k)} \right) \tag{22}$$

With

$$\hat{\pi}_3^{(k)} \stackrel{\text{def}}{=} \frac{1}{\sigma_3^{(k-1)} + exp(\omega_3)} \tag{23}$$

$$w_2^{(k)} \stackrel{\text{def}}{=} \frac{e^{\kappa \mu_3^{(k-1)} + \omega_2}}{\sigma_2^{(k-1)} + e^{\kappa \mu_3^{(k-1)} + \omega_2}} \tag{24}$$

$$r_2^{(k)} \stackrel{\text{def}}{=} \frac{e^{\kappa \mu_3^{(k-1)} + \omega_2} - \sigma_2^{(k-1)}}{\sigma_2^{(k-1)} + e^{\kappa \mu_3^{(k-1)} + \omega_2}} \tag{25}$$

$$\delta_2^{(k)} \stackrel{\text{def}}{=} \frac{\sigma_2^{(k)} + \left(\mu_2^{(k)} - \mu_2^{(k-1)} \right)^2}{\sigma_2^{(k-1)} + e^{\kappa \mu_3^{(k-1)} + \omega_2}} - 1 \tag{26}$$

Following the posterior updates from **Eq. 7** and **Eq. 13**, the equations for pwPE on level 2 (ε_2) and level 3 (ε_3) can be written as:

$$\varepsilon_2^{(k)} = \mu_2^{(k)} - \mu_2^{(k-1)} = \sigma_2^{(k)} \delta_1^{(k)} \tag{27}$$

$$\varepsilon_3^{(k)} = \mu_3^{(k)} - \mu_3^{(k-1)} = \sigma_3^{(k)} \frac{\kappa}{2} w_2^{(k)} \delta_2^{(k)} \tag{28}$$

As response model we used the unit-square sigmoid observation model for binary responses (Iglesias et al., 2013; Mathys et al., 2014). This transforms the predicted probability *m(k)* that the stimulus (e.g. blue) is rewarding on trial *k* (outcome = 1)—which is a function of the current beliefs—into the probabilities $p(y^{(k)} = 1)$ and $p(y^{(k)} = 0)$ that the participant will choose that stimulus (blue, 1) or the alternative (orange, 0):

$$p(y|m,\zeta) = \left(\frac{m^\zeta}{m^\zeta + (1-m)^\zeta}\right)^y \cdot \left(\frac{(1-m)^\zeta}{m^\zeta + (1-m)^\zeta}\right)^{1-y} \quad (29)$$

Higher values of the response parameter ζ lead to the participants being more likely to choose the response that corresponds with their current belief about the rewarded stimulus.

Fitting the combination of perceptual and response model to an individual participant's responses allows for a subject-specific characterisation of learning (and response) style by the set of perceptual (and response) parameters. Here, we estimated the parameters ω_2, ω_3, and ζ (see below for free model parameters in an alternative HGF model). The priors on these values were set to be relatively uninformative by choosing a broad variance (16 for ω_2, ω_3 and 1 for ζ as we expected less variation in this parameter). We fixed both the coupling parameter κ and the starting value of the belief on the third level $\mu_3^{(0)}$ to 1 following de Berker et al. (2016), but note that the scale of x_3 is arbitrary in our setting (for details, see Mathys et al., 2014). We chose a neutral starting value for the belief on the second level, i.e., $\mu_2^{(0)} = 0$, assuming participants would not have any initial preference for the outcome to be either rewarding (positive μ_2) or not rewarding (negative μ_2). The initial uncertainties of these beliefs ($\sigma_2^{(0)} = 0.1$ and $\sigma_3^{(0)} = 1$) corresponded to the default settings of the toolbox, and we verified that these values had a negligible impact on the estimated belief trajectories. All prior settings are summarised in **Table 1** (See also model space below for alternative models). Maximum-a-posteriori (MAP) estimates of model parameters were determined using these priors on parameters and the series of inputs, optimised with a quasi-Newton optimisation algorithm and calculated in the HGF toolbox version 3.1.

To assess the reliability of our estimates for the free model parameters in our implementation of the HGF (winning model, see model comparison details below), we simulated behavioural responses of 70 agents for nine different values of ω_2 (total 630 simulations), when observing the input of Cont participant #1. Similar simulations were run to estimate parameters ω_3 and ζ (**Supplementary Figure 2**). This analysis demonstrated high accuracy for estimating ω_2 and ζ, while ω_3 was poorly recovered. Poor estimation of ω_3 has also been reported in a recent study using a different approach (Reed et al., 2020). A complementary analysis using

simulated responses to observed inputs from StA participant #1 provided very similar results (**Supplementary Figure 3**).

Based on these results, we excluded ω_3 from subsequent between-group statistical analyses. Given our stimulus sequence, which exerted a certain level of volatility (changes in the contingencies every 26–38 trials) but did not contain marked changes in volatility, it is thus unsurprising that we could not infer on participants' beliefs about the meta-volatility ω_3. However, even if environmental volatility is constant, participants still need to estimate the adequate level for performing the task, suggesting that learning about volatility is still relevant (consistent with our model comparison results, see below; see also de Berker et al., 2016, which used a very similar task structure with constant true volatility).

In sum, in the current study, the computational quantities of interest for the statistical comparison between the groups were the model parameters ω_2 (tonic volatility estimate) and the decision noise from the response model, ζ. In addition, we assessed the trial-wise trajectories of posterior mean on beliefs about volatility (μ_3), environmental uncertainty, and the variances on levels 2 and 3 (σ_2, σ_3) as a measure of (informational) uncertainty about the hidden states on these levels. Note that due to the poor estimation of ω_3 ('meta-volatility'), which directly modulates precision in level 3 and thus the update steps on the expectation of volatility, μ_3, interpretation of between-group results for μ_3 and σ_3 should be treated with care.

Precision-weighted prediction errors play an important role in current Bayesian theories of perceptual inference and learning (Doya et al., 2007; Feldman & Friston, 2010; Friston et al., 2013; Friston & Kiebel, 2009a; Moran et al., 2013; Rao & Ballard, 1999), and these are the quantities that are considered to predominantly modulate EEG signals (Friston & Kiebel, 2009a). We initially selected the pwPE trajectories from levels 2 and 3 (labelled ε_2, ε_3, Eqs. [19] and [20]) to examine how these are represented in the brain as a function of state anxiety. However, as we identified a very high correlation between the regressors derived from ε_2 and ε_3, our final GLM analysis excluded the pwPE trajectories from level 3 (see GLM analysis section below).

2.7 Model Space

We tested five computational models of learning. The first three were a 3-level HGF (with volatility on the third level: HGF_3), a reduced 2-level HGF excluding volatility (HGF_2) and a modified 3-level HGF where the decision noise parameter that maps beliefs to choices (ζ) depends on trial-by-trial estimates of volatility (μ_3) (in line with Diaconescu et al., 2014; here

termed HGFμ_3). In the modified model HGFμ_3 trial-wise increases in volatility correspond with an individual exhibiting more exploratory or noisier behaviour (smaller decision noise ζ). In this model, in addition to the free model parameters ω_2 and ω_3, we estimated $\mu_3^{(0)}$ and $\sigma_3^{(0)}$. These were all hierarchical Bayesian models implemented using the HGF TAPAS toolbox (Mathys et al., 2011, 2014). The priors on hierarchical Bayesian model parameters are shown in Table 1. The fourth and fifth models were broadly used reinforcement learning models: a Rescorla Wagner (RW) where PEs drive belief updating but with a set learning rate (Rescorla & Wagner, 1972); and a Sutton K1 model (SK1) that permits the learning rate to change with recent prediction errors (Sutton, 1992).

Models were then compared at the group level for fit using random effects Bayesian model selection (BMS, Stephan et al., 2009) with code from the freely available MACS toolbox (Soch & Allefeld, 2018). BMS provided model frequencies and exceedance probabilities, reflecting how optimal each model or family of models performed (Soch et al., 2016). First, the log-model evidence (LME) from all Bayesian models were combined to get the log-family evidence (LFE) and was compared to the LFE of the family of reinforcement learning models (RW and SK1) to assess which provided more evidence. In the winning family, additional BMS determined the final optimal model.

2.8 EEG analysis and the General Linear Model

Prior to single-trial ERP analysis using the general linear model (GLM), a statistical analysis of the main effect of outcome on the ERP was conducted in the total population (N = 42). The aim of this ERP analysis was to assess whether the windows associated with the effect of the outcome (lose versus win) on the EEG signals in our task converge with the windows of the fERN and P300 effects reported in previous studies (see for instance Hajcak et al., 2005; Nieuwenhuis et al., 2004). Accordingly, we assessed the difference between lose and win ERPs in a broad window between 100 and 1000 ms, which includes the latency of those previously reported ERP components. This analysis was carried out using permutation tests with a cluster-based threshold correction to control the family-wise error (FWE) at level 0.05 (dependent samples t-test, 1000 iterations: Maris & Oostenveld, 2007; FieldTrip toolbox, Oostenveld et al., 2011).

To allow for the detection of significant clusters corresponding with positive and negative ERP differences, cluster-based test statistics being in the 2.5th and 97.5th percentiles of the permutation distribution were considered significant (two-sided test). For this statistical

analysis, the ERP data epochs were baseline-corrected by subtracting the mean activation during the baseline period from −200 ms to 0 ms.

For the GLM single-trial analysis, we selected a smaller 200–850 ms interval, primarily based on the observed latency of the fERN and P300 components in our study. This interval also covered the latency of HGF regressors in previous GLM studies (see, e.g. Diaconescu, Litvak, et al., 2017; Weber et al., 2020; although these studies used quite different tasks). Additionally, it should be noted that the modulation by pwPE regressors of the trial-wise ERP responses can peak at different latencies than the model-free ERP effects (Diaconescu, Litvak, et al., 2017; Stefanics et al., 2018; Weber et al., 2020).

In this analysis, EEG data were downsampled to 256 Hz, low-pass filtered at 30Hz and converted to SPM 12 (http://www.fil.ion.ucl.ac.uk/spm/ version 7487) (Penny et al., 2011). In SPM 12 software we converted the EEG data into 3-dimensional volumes (two spatial dimensions: anterior to posterior, left to right across the scalp; one temporal dimension: peri-stimulus time, Litvak et al., 2011). All participants' data consisted of 64 channels and 168 time points using a voxel size of 4.2 mm × 5.4 mm × 4 ms and were spatially smoothed to adjust for between-subject spatial variability in the channel space. The scalp x time 3D images were then tested statistically using statistical parametric mapping and the GLM (see next section, Kiebel & Friston, 2004a, 2004b; Kilner & Friston, 2010). This procedure is firmly established in EEG using SPM (Litvak et al., 2011).

Initially, our GLM was composed of trial-wise estimates of two computational quantities: absolute values of pwPEs in level 2 (ε_2), and pwPEs in level 3 (ε_3). The absolute value of ε_2 was selected because its sign is arbitrary: the quantity x_2 is related to the tendency of one choice (e.g. blue stimulus) to be rewarding ($x_1 = 1$); yet this choice and equivalently the sign of the pwPE at this level was arbitrary (see for instance Stefanics et al., 2018). In addition, we aimed to use as a third regressor the trial-wise win/lose outcome values as we expected this variable to account for much of the signal variance in the EEG epochs.

However, we observed a prominent correlation between the two regressor quantities abs(ε_2) and ε_3. The Pearson correlation coefficient ranged from 0.67 to 0.96 across all 42 participants, mean 0.86, median 0.88; and the correlation was significant in all participants ($P < 0.05$). The effect of collinearity on GLMs in neuroimaging has been assessed and discussed in detail before (see, e.g. Mumford et al., 2015). Collinear regressors can reduce power and lead to unreliable parameter estimates, but researchers should only be concerned with this issue in the case of near-collinearity (very high correlation between regressors, Mumford et al., 2015).

A common practice is to orthogonalise collinear regressors in the model to solve the problem of reduced power and unreliable parameter estimates in the GLM (Mumford et al., 2015). However, other authors argue that despite the potential appeal of orthogonalisation of regressors to remove collinearity from the model, the implications are actually not necessarily beneficial: it does not improve the overall fit of the model, and in most cases, it can lead to a misleading interpretation of the resulting inferences (Vanhove, 2020). Here we followed this second line of argumentation, and instead of orthognalising our pwPE regressors, we updated our GLM to only include trial-wise estimates of the absolute values of pwPEs on level 2, abs(ε_2), and the outcome of each trial (1 = win, 0 = lose). Regressor abs(ε_2) was chosen over ε_3 out of theoretical considerations, but also as abs(ε_2) was associated with much higher efficiency of beta coefficients in the GLM compared to ε_3 (see **Supplementary Materials**, following Mumford et al., 2015).

Having reduced the regressor space, we then assessed the efficiency for the β coefficients associated with each regressor in the final GLM. The efficiency for β_1, modulating the effect of pwPEs about reward outcome on the EEG, was in the same order of magnitude as the efficiency for β_2, associated with the outcome regressor (**Supplementary Materials**). In addition, we observed that, while the efficiency for β_2 was very similar in both Cont and StA groups, the efficiency for β_1 associated with abs(ε_2) was considerably lower in the StA group, relative to control participants. The efficiency values indicated that our final GLM model was a priori well specified for our chosen explanatory variable abs(ε_2), although it had greater efficiency for the regression coefficient on this variable in the control group.

Using these choices for regressors and time interval, we then carried out a whole-volume (spatiotemporal) analysis that searched for representations of our computational quantities in the single-trial EEG responses for each individual participant, before assessing within-group statistical effects at the second level. We corrected for multiple comparisons across the whole time-sensor matrix using Gaussian random field theory (Worsley et al., 1996) with a family-wise error (FWE) correction at the cluster-level ($P < 0.05$). This was performed with a cluster defining threshold (CDT) of $P < 0.001$ (Flandin & Friston, 2019). Importantly, all results reported survived whole-volume correction at the peak-level ($P < 0.05$). We assessed separately within each group, whether the trajectories of our computational quantities were associated with increases or decreases in EEG amplitudes using an F-test. Following within-group analysis, we used a 2-sample t-test to assess between-group StA minus Cont differences in the representations of those same regressors. A standard summary statistics approach was used to perform random effects group analysis within each group (StA, Cont) of 21 participants independently and between groups (Penny & Holmes, 2007).

2.9 Statistics

To assess Group (StA, Cont) and Block (1,2) main effects and interactions in state anxiety measures, behavioural, and computational model variables, we applied non-parametric factorial synchronised permutations tests (Basso et al., 2007) with 5000 permutations. These permutation-based factorial analyses assessed main effects and interaction effects for factors Block (TB1, TB2) and Group (StA, Cont). As shown below (see Results 3.1), we found a significant main effect of factor Block on the HRV index, indicating that the anxiety manipulation led to different physiological changes as a function of the block number. Accordingly, we continued to assess all our dependent variables using the 2 x 2 non-parametric factorial design with factors Block and Group. Factorial analyses were complemented with planned pair-wise permutation tests to assess our specific hypothesis of between-group differences (5000 permutations). This applies to the following dependent variables: (a) model-free behavioural measures (error rate, reaction time: RT); (b) CV as a measure of HRV (CV values in TB1, TB2 blocks were corrected by subtracting the R1 baseline value) and power for spectral analysis expressed in dB; (c) HGF perceptual model parameter ω_2 (tonic volatility modulating the variance of the Gaussian random walk at level 2); (d) Decision noise of the response model, ζ; (e) Relevant HGF quantities: (i) informational uncertainty about the reward tendency x_2 (σ_2); (ii) estimates of belief on volatility (mean, μ_3, and variance, σ_3); and last, (iii) environmental uncertainty—related to volatility in the environment: $\exp(\kappa\mu_3 + \omega_2)$.

Note that the above selected HGF trajectories do not directly reflect the subject-specific sequence of contingency blocks, which was randomly generated for each participant. By contrast, the expectation on the reward tendency, μ_2, was strongly associated with the structure of contingency blocks and therefore necessarily differed across participants by nature of the task design. Accordingly, μ_2 was not selected as a dependent variable.

Pair-wise permutation tests were also used to test within-group differences in RT across blocks. In the case of multiple comparisons (for instance, two between-group permutation tests run separately for each block), we controlled the false discovery rate (FDR) using an adaptive linear step-up procedure set to a level of q = 0.05 (Benjamini et al., 2006). This procedure furnished us with an adapted threshold p-value (P_{FDR}). Prior to these statistical analyses and following BMS, the trial-wise trajectory for each computational quantity of interest (σ_2, σ_3, μ_3, and environmental uncertainty, **Eq. 3**) was extracted from the winning model, followed by an average across trials within task blocks (TB1, TB2). By collapsing the trial information, we

aimed to assess the general block-related or group-related monotonic changes in the HGF quantities using the 2 x 2 factorial analysis with the factors Group and Block described above.

Below, in the Results section, we present the mean and standard error of the mean (SEM) for our dependent variables (either in text or in a figure), alongside non-parametric effect sizes for pair-wise comparisons and corresponding bootstrapped confidence intervals (Grissom & Kim, 2012; Ruscio & Mullen, 2012). In the case of within-group comparisons, the non-parametric effect size was estimated using the probability of superiority for dependent samples (Δ_{dep}), whereas for between-group effects we used the probability of superiority (Δ); both are calculated in line with Grissom and Kim (2012), expressed as the number of values in sample A greater than those in sample B ($\Delta = P[A>B]$). In the case of dependent samples, the comparison between pairs is done for matched pairs. Although in the original formulation by Grissom and Kim (2012), ties were not taken into account; here, in line with Ruscio and Mullen (2012), we corrected (Δ) using the number of ties (difference scores = 0) and estimated bootstrapped confidence intervals (CI) for (Δ).

3. Results

3.1 Heart-rate variability

Using a non-parametric 2 x 2 factorial test with synchronised rearrangements, significant main effects of Block ($P = 0.009$) and Group ($P = 0.04$) were revealed on the normalised HRV index during reward-based learning blocks. No interaction effect was found. Planned between-group comparisons using permutation testing revealed significantly lower HRV in StA during TB1 (mean −0.02, SEM 0.004) when compared to Cont (mean −0.005, SEM 0.005, $P_{FDR} < 0.05$, $\Delta = 0.70$, CI = [0.54, 0.86], see **Figure 3A**). These results indicate that the experimental manipulation induced physiological changes in cardiovascular activity corresponding to an anxious state (Chalmers et al., 2014; Feldman et al., 2004). An additional analysis on the spectral characteristics of the IBI time series corresponded our HRV result to autonomic inflexibility in state anxiety, with significantly reduced high frequency HRV (0.15–0.4 Hz, termed HF-HRV hereafter) in StA (mean −6.3, SEM 0.6) compared to Cont (mean −4.7, SEM 0.5, non-parametric 2 x 2 factorial test with synchronised rearrangements: significant main effect of Group $P = 0.02$ and trend level interaction effect $P = 0.06$, see **Supplementary Figure 4**). There was no effect of Block ($P = 0.8$).

Our analysis demonstrating reduced HRV in state anxiety corresponds to both prior research showing lower HRV in anxiety (Chalmers et al., 2014; Friedman & Thayer, 1998; Gorman & Sloan, 2000) and our previously published work using a similar state anxiety manipulation where we found lower HRV along with higher state anxiety scores using the STAI state scale omitted from the present study (Sporn et al., 2020). Our additional analysis of the frequency content in the IBI time series further links our lower HRV result (as a proxy of state anxiety) to research showing reduced HF-HRV (0.15–0.40 Hz) in anxiety conditions, a physiological expression of inflexible autonomic activity (Aikins & Craske, 2010; Friedman, 2007; Fuller, 1992; Klein et al., 1995; Miu et al., 2009; Mujica-Parodi et al., 2009; Pittig et al., 2013; Thayer et al., 1996).

3.2 Model-free Analysis

The percentage of errors made by each participant across 400 trials was used as a measure to assess whether anxiety impairs reward-learning task performance. Using non-parametric factorial test (synchronised rearrangements), the main effect of factor Group ($P = 0.01$) on error rates was significant, whereas factor Block revealed only a trend ($P = 0.056$). There was

no significant interaction effect (P = 0.66). A planned between-group comparison on the Group factor alone using pair-wise permutation tests revealed a significantly higher total average error rate in the StA group (mean 38.0, SEM 0.97), in comparison to the Cont group (mean 35.6, SEM 0.66, P = 0.001, Δ = 0.70, CI = [0.58, 0.82], see **Figure 3B**).

Turning to the mean reaction times (RT, in milliseconds; averaged across all trials), a significant main effect of task Block on RTs was observed (P = 0.02). No significant main effect of Group or interaction effect was found (P = 0.64, P = 0.26) in line with previous work on anxiety (Bishop, 2009). Post-hoc analyses on the Block effect in each group revealed there was a significant decrease in the mean RT from TB1 (mean 658.3, SEM 32.53) to TB2 (mean 552.8, SEM 20.68) in the StA group (P_{FDR} = <0.05, Δ = 0.73, CI = [0.65, 0.81]). This effect was also significant for Cont, with mean RT dropping from TB1 (mean 657.7, SEM 35.65) to TB2 (mean 591.4, SEM 31.78, P_{FDR} = <0.05, Δ = 0.65, CI = [0.56, 0.72]). As a separate analysis, and given the lack of between-group differences in RTs, we contrasted the total-population (StA + Cont) mean RT of lose and win trials, using the RT value from the current trial. The lose minus win RT difference was highly significant, as expected, reflecting a slower response on trials where participants responded incorrectly (P = 0, Δ = 1, CI = [1, 1]; mean RT for lose trials 639.8 ms [SEM 0.36 ms], and for win trials 600.8 ms, [SEM 0.33 ms], **Figure 3C**).

In a final post-hoc analysis, we assessed RT separately in blocks of unpredictable cues (50-50 contingency phase) and highly predictable cues (90–10 contingency phase). The rationale for this analysis was that previous work using classical attention paradigms revealed that lack of attention leads to larger RTs both on trials with predictive cues (our 90–10 contingency phase) and trials with uninformative cues (as our 50–50 contingency phase; see for instance Prinzmetal et al., 2009). We thus aimed to assess whether state anxiety influenced attention using classical behavioural measurements. We examined whether the strength of the contingency bias shapes RT differently in each group. The results demonstrate no significant difference in RT between the two groups for either unpredictable or predictable contingency phases (P > 0.05: See **Supplementary Materials**).

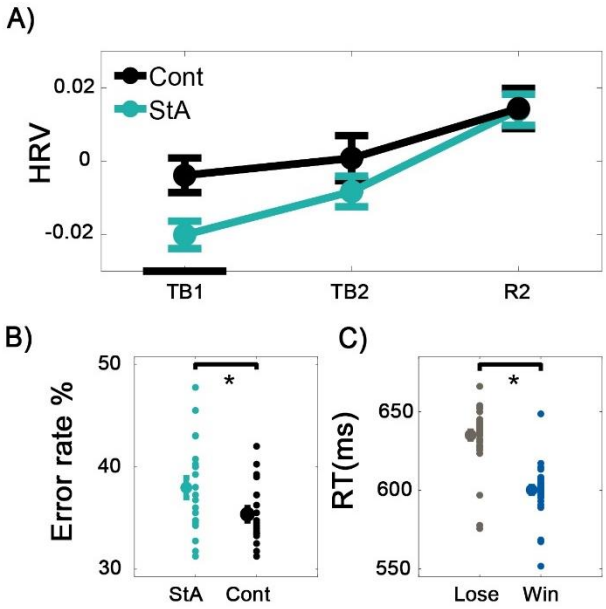


Figure 3. State anxiety modulates heart rate variability and behavioural responses. **A)** Modification in heart-rate variability (HRV) during anxiety manipulation. The average HRV (measured with the coefficient of variation of the inter-beat-interval of the ECG signal) is provided for the state anxiety (StA) and Control (Cont) groups across task block 1 (TB1), task block 2 (TB2) and final resting state (R2). The average of the resting state (R1: baseline) has been subtracted from each subsequent task block to normalise HRV values. Significant between-group differences assessed in learning blocks TB1 and TB2 are identified by black bars on the x-axis (paired permutation test, $P_{FDR} < 0.05$ after control of the FDR at level q = 0.05). **B)** The effect of anxiety on reward-based learning performance: error rates. Here, the average error rates of each group, the state anxiety (StA) and the control group (Cont) are presented using a central point flanked by SEM bars. To the right of each mean and SEM are the individual data points in each group to show group population dispersion. Anxiety significantly increased the error rate in the StA group when compared to Controls ($P = 0.001$). **C)** The main effect of outcome (win, grey; lose, blue) on mean reaction times on that same trial (RT: $P = 0$). On the left, the average RT of each outcome is presented using a central point with SEM bars. To the right of each mean and SEM are the individual data points of each group to show group population dispersion.

3.3 Bayesian Model Selection

After fitting each model (HGF: 3-Levels [HGF_3], 2-Levels [HGF_2], HGFμ_3, the Rescorla Wagner [RW], and Sutton K1 [SK1]) individually in each of the 42 participants and obtaining log-model evidence (LME) values for each, we compared the five models using Bayesian model selection (BMS). Results from BMS revealed that the family of Bayesian models (HGF_3, HGF_2, and HGFμ_3) had much stronger evidence than the reinforcement-learning models (RW, SK1), with an exceedance probability of 0.99, and an expected frequency of 0.74 (leftmost columns: **Figure 4A**). Next, within the Bayesian models, an additional BMS step using the LME for each subject and model demonstrated much stronger evidence for the HGF_3 model relative to the HGF_2 and HGFμ_3 model versions, with an exceedance probability of 0.98 and an expected

frequency of 0.61 (rightmost columns: **Figure 4A**). The HGF_3 model was the winner model also when performing BMS separately in the StA and Control groups.

Although a previous study found the HGFμ_3 model to outperform the 3-level HGF (Diaconescu et al., 2014), here we found the latter to provide more model evidence. One possible explanation is that the HGFμ_3 might be particularly useful in paradigms where (at least some) participants are exposed to a scenario of alternating low and high volatility (Diaconescu et al., 2014). For a fixed value of true volatility, as in our study (constant rate of change of contingency blocks), the standard 3-level HGF with a decision noise parameter that is not modulated by the dynamics of μ_3 performed the best.

To further determine the quality of the fit of the winning HGF_3 model, we simulated responses using the estimated model parameters for each individual (ω_2, ω_3, ζ). Similarly to Aylward et al. (2019), we computed the probability of response switch (choosing orange or blue) across trials in each individual and separately for simulated and empirical responses. We found a high significant non-parametric Spearman rank correlation between both variables across participants (N = 42): $\rho = 0.8679$, $P = 4 \times 10^{-14}$ (**Supplementary Figure 5**). A similar outcome was obtained when assessing correlations within each group, suggesting that the winning model captured the dynamics in the data well.

3.4 Model-based analysis

3.4.1 State anxiety is associated with a lower learning rate about stimulus outcomes

We observed significant differences between the groups in the perceptual model parameter ω_2, with smaller values obtained in StA (mean −3.1, SEM 0.23) when compared to the Cont group (mean −2.0, SEM 0.15, $P = 0.002$, effect size: $\Delta = 0.75$, CI = [0.55, 0.90]). The decision noise parameter, ζ, did not differ between groups ($P = 0.62$), and was moderately low in both groups: Cont (mean 1.98 [0.26]) and StA (2.20 [0.41]).

The values of ω_2 influence, among other HGF trajectories, the learning rate at the lowest level, α (through modulation of μ_2), driving the step of the update about stimulus outcomes (Mathys et al., 2014). More negative ω_2 values—as found in StA—lead to smaller updates, and thus to smaller learning rates (See illustration in **Figure 4B).** To illustrate the impact that this has on the evolution of beliefs about reward contingencies and environmental volatility in our task, we additionally provide simulations of belief trajectories for both μ_2 (σ_2) and the log-volatility μ_3 (σ_3)

(**Supplementary Figures 6-7**). The results demonstrate that decreasing ω_2 reduces the estimation uncertainty about the reward tendency σ_2 with smaller update steps on μ_2 (**Supplementary Figure 6**). The effect of decreasing ω_2 on high-level beliefs is to reduce the update steps for the expectation of log-volatility μ_3 and increase uncertainty σ_3 (**Supplementary Figure 7**). In the following, we explore whether the two experimental groups did indeed differ in the uncertainty of their beliefs as a consequence of the observed change in ω_2.

3.4.2 Informational Uncertainty about the outcome tendency is lower in state anxious individuals.

As indicated in the previous section and illustrated in **Supplementary Figure 6**, the informational (belief) uncertainty about the outcome tendency, σ_2, is reduced for smaller ω_2 values, while it also depends on the volatility estimate μ_3 from the previous trial and other quantities (see Eq. 11 and 13 in Mathys et al., 2014). Here we found a significant main effect of factor Group on σ_2 ($P = 0.008$). There were no significant effects of block and no interaction effect ($P = 0.58$, $P = 0.78$). In addition, planned comparisons showed that anxiety significantly lowered the total average σ_2 for StA in comparison to Cont, as expected from the lower ω_2 values in StA (**Figure 4C**; $P = 0$, $\Delta = 0.75$, CI = [0.55, 0.89]). A lower belief uncertainty about the outcome tendency in StA individuals means that new information had a smaller impact on the update equations for beliefs about x_2 in this group.

3.4.3 Environmental uncertainty is underestimated in state anxiety.

Environmental uncertainty—induced by changes in the environment—depends on the tonic volatility estimate, ω_2, and the trial-wise volatility estimate $\mu_3^{(k-1)}$ (see **Eq. 3** above; the coupling constant κ was fixed to one). More volatile environments lead to greater environmental uncertainty. We found that environmental uncertainty was significantly modulated by factor Group ($P = 0.02$), while there was no significant main effect for factor Block or interaction effect ($P = 0.58$, $P = 0.75$). Further pair-wise analyses demonstrated that the StA group underestimated the environmental uncertainty, relative to control participants, when averaging across both experimental blocks (**Figure 4D**; $P = 0$, $\Delta = 0.74$, CI = [0.54, 0.89]), consistent with their reduction in ω_2.

3.4.4 Uncertainty about volatility is higher in state anxious individuals.

In contrast to the effect on σ_2 reported above, state anxiety increased belief uncertainty on level 3 (σ_3; uncertainty about volatility). We found both a significant main effect of Block (P = 0.0006) and Group on this parameter (P = 0.0002), yet no interaction effect (P = 0.99). Across blocks, the uncertainty about volatility generally decreased. Planned comparisons demonstrated that separately in the first and second task blocks anxiety significantly increased σ_3 in the StA group when compared to the Cont group (**Figure 4E**; $P_{FDR} < 0.05$, effect size for TB1: Δ = 0.73, CI = [0.53, 0.89]; TB2: Δ = 0.74, CI = [0.53, 0.89]). The larger anxiety-induced uncertainty about volatility is consistent with the effect of decreasing ω_2, as illustrated in our simulations (**Supplementary Figure 7**).

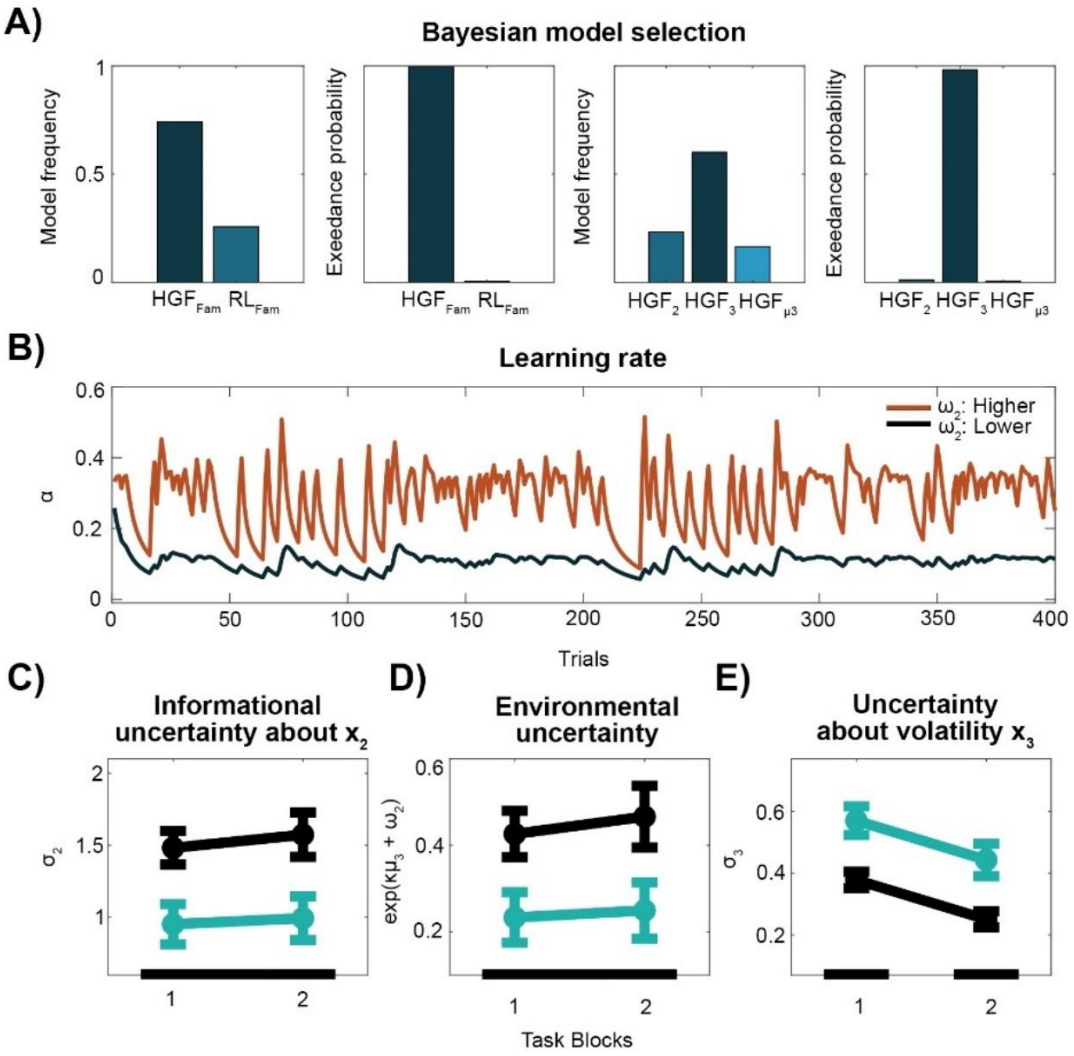


Figure 4. Bayesian Model Selection and Hierarchical Gaussian Filter Results. **A)** Bayesian model selection (BMS). The two leftmost panels represent the model frequency and exceedance probability for the family of models 'HGF Fam' (HGF_2, HGF_3 and $HGF\mu_3$: dark blue) and the family of reinforcement

learning models 'RL FAM' (RW, SK1: blue). The family of HGF models provided the best model evidence. In the two right panels is the comparison between the three HGF models (HGF_2: blue, HGF_3: dark blue and $HGF\mu_3$: light blue). The HGF_3 provided stronger model evidence. **B)** Illustration of the trial-by-trial learning rate about stimulus outcomes (α) in two ideal observers with different values of ω_2. Trajectories were simulated using the same input sequence and parameters (except ω_2): $\mu_2^{(0)} = 0$, $\mu_3^{(0)} = 1$, $\sigma_2^{(0)} = \log(0.1)$, $\sigma_3^{(0)} = \log(1)$, $\kappa = 1$, $\omega_3 = -7$. The two values on ω_2 used in the simulated trajectories are -2 (orange) and -4 (black). This parameter represents the tonic part of the variance in the Gaussian random walk for x_2 and modulates the learning rate about stimulus outcomes at the lowest level. Lower ω_2 values lead to smaller trial-by-trial learning increments. When comparing ω_2 values between groups (StA, Cont), we found more negative values in StA than in the Cont group ($P = 0.002$). **C)** Lower ω_2 in state anxiety leads to decreased informational uncertainty about x_2. There was a significant main effect for factor Group (StA, green; Cont, black; synchronised permutation test: $P = 0.008$) but not for factor Block ($P > 0.05$). Planned between-group comparisons indicated that state anxiety significantly decreased the average uncertainty about beliefs on tendency x_2 ($P = 0$, as given by black bars), after averaging across both blocks; significant effect indicated by black bars at the bottom). **D)** Lower ω_2 in state anxiety leads to decreased environmental uncertainty ($P = 0.02$) (no effect of factor Block [$P > 0.05$]). Thus, StA participants had a lower estimate of environmental uncertainty. **E)** State anxiety increased uncertainty about volatility in the task environment (σ_3). We found a significant main effect for factor Block ($P = 0.0006$) and Group (StA, green, Cont, black; $P = 0.0002$), modulating uncertainty about volatility. Planned between-group comparisons further indicated that state anxiety exhibited significantly higher σ_3, as compared to control participants, separately in each task block (TB1, TB2, $P_{FDR} < 0.05$, as given by black bars).

3.5 Standard Lose versus Win ERP results

Cluster-based random permutation tests demonstrated a significant difference between the effect of the two outcomes (lose, win) on the ERP (N = 42: two significant clusters at level $P < 0.025$). Losing led to a more negative ERP amplitude than winning during a time window between 200 and 350 ms post outcome (negative cluster, $P = 0.003$). This effect at first had a centro-parietal distribution, which later propagated to broader central, frontal, temporal, and parietal electrode regions, occurring approximately in line with the fERN ERP (**Supplementary Figure 8**). In a later time window, between 350 and 860 ms, losing evoked a more positive amplitude when compared to winning (positive cluster, $P = 0.0002$). During this later latency, the difference originated over fronto-central electrodes and later spread to centro-parietal electrodes resembling the P300 component wave (**Supplementary Figure 8**). The latency of the significant clusters confirmed that lose relative to win trials elicited a biphasic ERP modulation consisting of an earlier negative wave resembling the fERN and a later positive and very pronounced deflection corresponding to the P300.

3.6 Single-trial ERP modulations by precision-weighted PEs

The HGF results had confirmed that state anxiety alters informational uncertainty of beliefs about reward contingencies (level 2) and also about volatility (level 3) (**Figure 4**) in an opposing pattern of changes (decrease in σ_2 and increase in σ_3 relative to control participants). We then proceeded to analyse in each group separately the electrophysiological representations of trial-wise pwPEs for level 2—which are a function of the uncertainty estimate as shown in **Eq. 12** (for an illustration of abs(ε_2), see **Figure 5A**). The GLM results of the additional outcome regressor are shown in **Supplementary Figure 9.**

3.6.1 Low-level precision-weighted prediction errors

In the Cont group, abs(ε_2) significantly modulated trial-wise EEG responses from 475 ms to 503 ms post stimulus over central and parietal electrodes, with a maximum effect at 496 ms across a left parietal region ($P_{FWE} < 0.0001$). An additional significant effect of a smaller cluster size was found earlier between 425-464 ms with a peak at 457 ms ($P_{FWE} = 0.001$) across central and frontal electrodes (**Figure 5B**). Details on the cluster effects can be found in **Table 2**. Precision-weighted PEs about the stimulus tendencies abs(ε_2) did not significantly modulate the ERP responses in the StA group. When directly contrasting the groups, there were no significant differences in the representation of abs(ε_2) in EEG activity.

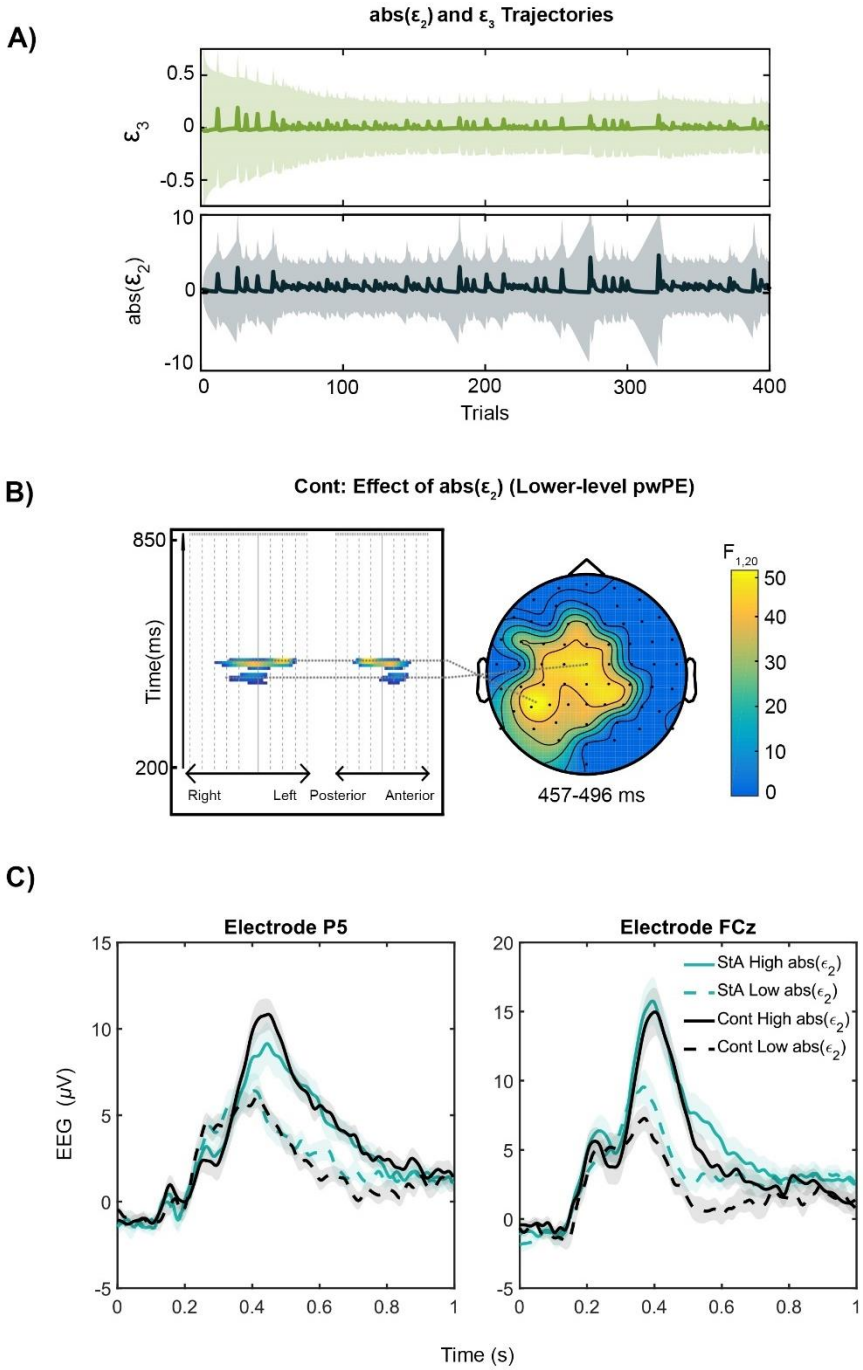


Figure 5. Signatures of precision-weighted prediction errors on trial-wise ERPs. **A)** Trajectories of model-based estimates for both lower-level and higher-level pwPE for one representative control group participant across 400 trials. In green are higher-level pwPEs concerning volatility; in black are the absolute values of the lower-level pwPE concerning beliefs about the rewarding stimulus. This second trajectory was chosen as a regressor in our GLM analysis, whereas ε_3 was excluded due to near-collinearity of ε_3 and abs(ε_2): Individual correlation values were around 0.9 (see main text). **B)** Effect of pwPEs on level 2 (abs(ε_2)) in controls. In the Cont group, pwPEs about reward outcomes correlated with activation changes across a left parietal and central region between 475 ms to 503 ms, as shown in the topographical representation on the top at the time of the maximum peak of the cluster (496 ms post stimulus, $P_{FWE} < 0.0001$, with a cluster-defining threshold of $P < 0.001$). An earlier cluster was also found as shown in the bottom topographical representation, with activation between 425–464 ms ($P_{FWE} = 0.001$) at frontocentral channels. **C)** The bottom panels show the average EEG response to the 40 highest ("High") and 40 lowest ("Low") pwPE values from each participant, and at P5 and FCz

electrodes—representing the significant GLM cluster obtained in Cont participants (shown in B). The averaged EEG responses are displayed separately for StA High, StA Low, Cont High, and Cont Low. Both participant groups show an increased response in EEG activity during "High" relative to "Low" abs(ε_2) trials at both electrode locations and between 475–503 ms. Shaded bars show 1.96∗SEM.

4. Discussion

We combined computational modelling of behaviour and analysis of electrophysiological responses to examine how state anxiety modulates reward-based learning when learning in a volatile environment. Our key finding is that state anxiety was associated with a reduced estimate of tonic volatility, which resulted in an overall lower learning rate, and corresponded to a significant underestimation of environmental and informational uncertainty. At the same time, a reduction of tonic volatility in our paradigm led to a decrease in learning about phasic volatility, a higher-level belief about the current rate of change in the environment. Our modelling results offer a mechanistic explanation for the increase in error rate that we observed in the anxiety group.

Trial-wise estimates of uncertainty—or its inverse, precision—serve to scale the impact of prediction errors (PEs) on the belief updates. We found that precision-weighted PEs (pwPEs) about the stimulus-reward contingency explained trial-wise modulation of observed ERP responses in control participants only. We observed these effects mainly around 425–503 ms across left parietal and central sensors. In state anxiety, there was no significant effect of lower-level pwPEs about reward contingencies on EEG amplitudes. Taken together, the data suggest that temporary anxious states in healthy individuals impair reward-based learning in volatile environments, primarily through changes in uncertainty estimates, potentially mediated by a degraded neuronal representation of lower-level pwPEs about reward contingencies, although the latter remains speculative given the lack of significant differences in pwPE representation between the groups (see below for further discussion).

States of anxiety bias computations of uncertainty during reward-based learning

The threat of a public speaking task used in our experiment reduced both heart rate variability, which is consistent with previous findings on state anxiety (Chalmers et al., 2014; Feldman et al., 2004; Gorman & Sloan, 2000), and high frequency HRV (HF-HRV: 0.15–0.4 Hz), an index of autonomic inflexibility found across anxious conditions (Friedman, 2007; Miu et al., 2009; Mujica-Parodi et al., 2009; Pittig et al., 2013). Beyond the initial induction where the effect of

state anxiety on HRV was strongest, modulation to computational estimates of uncertainty persisted into the second task block. This suggests states of anxiety have a diffuse and transient effect on computational estimates of uncertainty during learning in volatile environments. It is important to note though that we were not able to validate the HRV block-related effects with similar changes in the self-reported STAI inventory, as the data from this inventory were not acquired at the appropriate times. Our results are therefore based on the assumption that we can use the HRV changes observed here as a proxy for the successful induction of state anxiety in our paradigm. This assumption is further supported by a previous study from our lab (Sporn et al., 2020) where we used a similar approach in a motor learning task to successfully induce changes in state anxiety STAI scores (higher in the anxiety group), which was paralleled by a lowering of the same HRV proxy of state anxiety as used here.

Our experimental manipulation had an adverse effect on reward-based learning. Having matched trait anxiety levels across the state anxious and the control group, our results indicate that the changes observed in reward-based learning—lower learning rates due to changes in belief uncertainty—can be linked to temporary anxious states independently of trait levels. These outcomes thus expand prior findings of an association between high levels of trait anxiety and difficulties in decision-making tasks (de Visser et al., 2010; Miu, Heilman, et al., 2008) and learning in volatile task environments (Browning et al., 2015; Huang et al., 2017) to the realm of state anxiety. While we tested only a volatile environment where probabilistic contingencies changed regularly (as in Iglesias et al., 2013; de Berker et al., 2016), a still unresolved question is whether the anxiety-related modulations of uncertainty estimates are exclusive to a volatile environment or would also emerge in stable environments. Given that previous research in trait anxiety showed that learning is affected exclusively during volatile (not stable) experimental phases (Browning et al., 2015; Huang et al., 2017), we predict that during stable blocks state anxiety would not alter belief uncertainty. Moreover, our task did not allow for robust inferences on phasic volatility estimation (as reflected by parameters like the meta-volatility ω_3). Additional follow-up work should extend the current paradigm to also consider an environment with dynamic (as opposed to fixed) volatility, to systematically assess whether state anxiety affects the estimation of phasic volatility on top of the altered tonic volatility estimates observed here.

By using the threat of public speaking instead of a specified aversive outcome, our approach allowed us to investigate behavioural, physiological, and neural responses in anticipation of a future unpredictable threat. Alterations in anticipatory responses to upcoming uncertain threats have been proposed to be a common explanation for anxious states in healthy individuals and anxiety disorders alike (Grupe & Nitschke, 2013). Accordingly, our findings

that anxiety leads to changes in informational and environmental uncertainty could prove relevant for understanding the alterations in decision-making and learning observed in anxiety disorders (Bishop & Gagne, 2018; Browning et al., 2015; de Visser et al., 2010; Huang et al., 2017; Miu, Heilman, et al., 2008).

Our approach is not the first in proposing a role of uncertainty estimates in cognitive biases in anxiety. A recent account of affective disorders suggested that difficulties with uncertainty estimation underlie some of the psychiatric symptoms in these populations (Pulcu & Browning, 2019). This work distinguished between different types of uncertainty, corresponding to irreducible, informational, and environmental uncertainty as described here, and assigned a particular relevance of environmental ("unexpected") uncertainty in explaining anxiety. In fact, evidence from computational studies converges in linking trait anxiety with difficulties in learning in unstable or volatile environments (Browning et al., 2015; Huang et al., 2017). As shown by Browning et al. (2015), an inability to adapt to changes in a task structure can be measured by comparing a single volatile block to a single stable block. Alternatively, suboptimal learning in anxiety can be captured by focusing on volatile environments alone, in which the probability of reward (or punishment) changes regularly across different blocks (Huang et al., 2017).

Here we followed the second approach to investigate reward-based learning in a volatile environment. We investigated the adaptive scaling of learning rates to estimates of environmental uncertainty on a trial-by-trial basis by applying a computational model that explicitly incorporates learning about volatility in a hierarchical Bayesian framework. The winning computational model that best explained our behavioural data was the 3-level HGF, where the third level is a mathematical description of volatility estimates and their variance. Our inferences about phasic volatility estimation, as represented on this third level, are limited by the fact that our paradigm did not include marked changes in the level of volatility over time. Accordingly, we were not able to recover perceptual parameters related to phasic volatility estimation. The fact that the model that included phasic volatility estimation was still a better explanation of the observed responses suggests that trial-wise updating of beliefs about the level of volatility may nevertheless play a role. Participants still need to infer the adequate level of volatility as they perform the task (Iglesias et al., 2013; Weber et al., 2020). Similarly, the three-level HGF outperformed the two-level HGF in a task with comparable structure (and identical priors), further suggesting the validity of the three-level HGF in identifying learning alterations in threatening or stressful environments (de Berker et al., 2016).

We found that the state anxious participants' estimates of tonic volatility, as captured by the parameter ω_2, were significantly lower than controls, which led to significantly reduced learning rates and estimates of informational and environmental uncertainty. Beliefs about the outcome tendency were thus estimated to be more precise during anxiety, such that new and potentially revealing information about the true nature of hidden states had a smaller influence on the belief updates on that level. Critically, an overly precise belief about the outcome tendency might be inappropriate given the fluctuations in the true underlying hidden state. Thus, a drop in informational uncertainty during state anxiety might lead to biased learning, which here was further characterised by a lower learning rate about stimulus outcomes. This finding was confirmed in a separate model-free behavioural analysis: state anxious individuals exhibited a higher error rate during task performance relative to control participants. Our study thus provides novel and compelling evidence for abnormal precision (uncertainty) estimates underlying impoverished learning in healthy individuals going through temporary states of anxiety. Thereupon, the improper weighting of precision could be a general mechanism underlying a range of cognitive biases observed in healthy and psychiatric conditions, such as "hysteria" or autism (Edwards et al., 2012; Lawson et al., 2017).

Theories of aberrant precision estimates are typically formulated using a Bayesian or predictive coding framework (Parr, Rees, et al., 2018). Precision is formalised as an attentional mechanism, calibrating neural gain to regulate the influence of prior beliefs and sensory outcomes on future expectations (Feldman & Friston, 2010; Friston & Kiebel, 2009a; Moran et al., 2013). Our results provide evidence for this computational account of attention through altered uncertainty estimates. However, more "classical" accounts of attention detailing a limited resource capacity do not wholly explain our behavioural data (Lavie, 1995; Lavie et al., 2004). Our results showed that RT was not affected by the anxiety manipulation (in line with Bishop, 2009). This suggests deficient attentional resources or increased distraction are not the primary driving factor behind our reported impaired learning performance under state anxiety.

We also found that state anxiety led to a decrease in the precision of beliefs about environmental volatility, and reduced learning about this quantity. Learning about higher-level quantities depends upon the transmission of learning signals (precision-weighted PEs) from lower to higher levels. As our simulations show, a reduction in tonic volatility estimates does not only reduce learning about the contingencies governing observed stimuli and outcomes (**Supplementary Figure 6**) but also impairs learning about volatility. In particular, it prevents a trial-by-trial modulation of volatility estimates—learning—which would reduce the uncertainty about this quantity (**Supplementary Figure 7**). Therefore, the model indicates that state

anxious individuals remained uncertain about the current rate of change in the environment in our task. However, to examine whether state anxiety induces changes to phasic volatility estimation above and beyond this consequence of aberrant tonic volatility estimates, future studies will have to confront participants with environments in which the rate of change is dynamic across the experiment.

Changes to the contingencies governing the outcomes in our task were abrupt (see **Figure 1B**), which is in contrast to the generative model of the environment suggested by the HGF, where states evolve as Gaussian random walks and thus change slowly and diffusively over time. While the HGF has been successful in explaining and predicting human behaviour in such tasks (e.g., Iglesias et al., 2013; de Berker et al., 2016), alternative models have been proposed based on a generative model which expects sudden changes (Moens & Zénon, 2019; Nassar et al., 2010). In practice, both approaches (HGF and change-point models) can successfully deal with both kinds of environments (sudden versus diffuse changes), as a recent comparative analysis found (Marković & Kiebel, 2016). However, this analysis also indicated that Bayesian inference and model comparison methods can accurately disambiguate between data generated by the HGF versus a (reformulation of a) change-detection model. To understand whether participants use one or the other to infer on the dynamics of the environment, future work would thus profit from directly comparing the recent reformulations of change-point models (Marković & Kiebel, 2016; Moens & Zénon, 2019) to the HGF.

Overall, the computational results confirm our hypothesis that state anxious individuals choose their responses founded on a biased representation of uncertainty over the current belief states—at least when dealing with volatile environments as assessed here. Overly precise beliefs may represent a strategy to regain a sense of control because uncertainty is experienced as aversive (Carleton, 2016), such as observed in obsessive compulsive disorder (Carleton, 2016) and ritualistic behaviour (Lang et al., 2015). In turn, this emergence of biased estimates could increase the symptoms of anxiety over time through inaccurate recursive assessments of threat from uncertainty, thereby fitting a profile of anxious responses similar to those of anxiety-related disorders (Grupe & Nitschke, 2013; Pulcu & Browning, 2019).

Precision-weighted prediction errors modulate trial-by-trial ERP responses

The modulation of trial-by-trial ERP responses by lower-level pwPEs in the control group aligns with previous studies combining EEG analyses with the HGF, which revealed that low-level pwPEs are reflected in trial-wise ERP responses during learning and perception in

unstable environments (Stefanics et al., 2018; Weber et al., 2020). Some studies also found higher-level pwPEs modulating brain responses, and supported that different hierarchically-related pwPEs (or related HGF quantities) are represented across different brain regions specific to the task demands (Diaconescu, Litvak, et al., 2017; Iglesias et al., 2013; Weber et al., 2020).

Here, however, we excluded higher-level pwPEs from the GLM analysis due to near-collinearity of ε_3 and abs(ε_2) regressors. The fact that we did not observe a significant modulation of EEG responses by lower-level pwPEs in the StA group is consistent with our finding of reduced learning rates in this group. However, EEG responses to pwPEs were not significantly different when directly contrasting the groups, which prevents us from drawing strong conclusions about differential pwPE representations during state anxiety. The complementary visualisation of ERP modulations to high and low pwPEs further suggested a similar profile of ERP amplitude changes for both groups at the peak electrodes showing within-group effects to abs(ε_2) in the control group. Thus, the specific neural mechanism explaining the biased uncertainty estimates on reward contingencies—which are related to lower-level pwPEs—observed in state anxious participants remains elusive.

More generally, the evidence for neural representations of pwPEs in the control group is aligned with current predictive coding proposals. These view the brain as a Bayesian observer, estimating beliefs about hidden states in the environment through implementing a hierarchical generative model of the incoming sensory data (de Lange et al., 2018; Doya et al., 2007; Friston, 2010; Rao & Ballard, 1999). In this framework, superficial pyramidal cells encode PEs weighted by precision, and these are also the signals that are thought to dominate the EEG (Friston & Kiebel, 2009a). This motivated us to assess the representation of pwPEs in brain responses, an approach followed by some of the previous fMRI and EEG studies (Diaconescu, Litvak, et al., 2017; Iglesias et al., 2013; Stefanics et al., 2018; Weber et al., 2020).

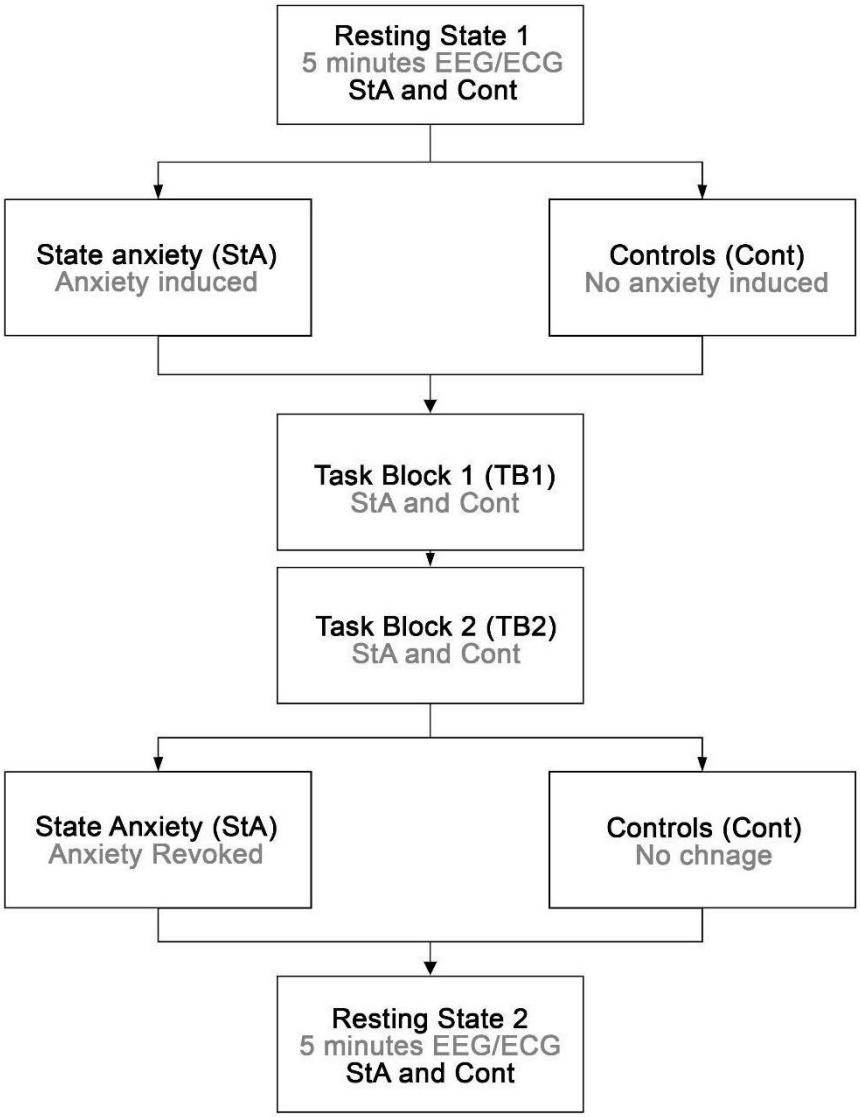
Other model-based studies of trial-wise ERP responses like the P300 assessed alternative Bayesian inference parameters, such as precision or Bayesian surprise (Kolossa et al., 2015; Mars et al., 2008; Ostwald et al., 2012). The centrally-distributed P3a component around 340 ms was identified as an index of belief updating, whereas the later P3b waveform of posterior topography was found to represent Bayesian surprise (Kolossa et al., 2015). Despite these computational approaches to the P300 not being directly comparable to our pwPE results, they share a similar timeline and topography, as the centroparietal cluster in the Cont group overlaps with the location of the P3a and P3b waves as shown in Kolossa et al. (2015). The ERP modulation to low-level pwPEs in our study might thus partially contribute to the

explaining the P300 amplitude changes obtained in the standard lose minus win ERP analysis conducted here, which itself showed the expected topographic gradient of the P300 component from central to posterior regions as shown in classical model-free ERP studies (Hajcak et al., 2005, 2007; Polich, 2007; Wu & Zhou, 2009). Collectively, these results suggest that future studies assessing the effect of subclinical (trait, state) anxiety on the neural representation of computational quantities related to prediction updates could specifically target the topography and latency of the trial-wise P300. A state anxiety manipulation using the widely-used method of the threat of shock (Grillon et al., 2019) could potentially induce more consistent neural responses in StA participants and thus allow for discrimination of the neural bases of pwPE in this group when compared to control participants.

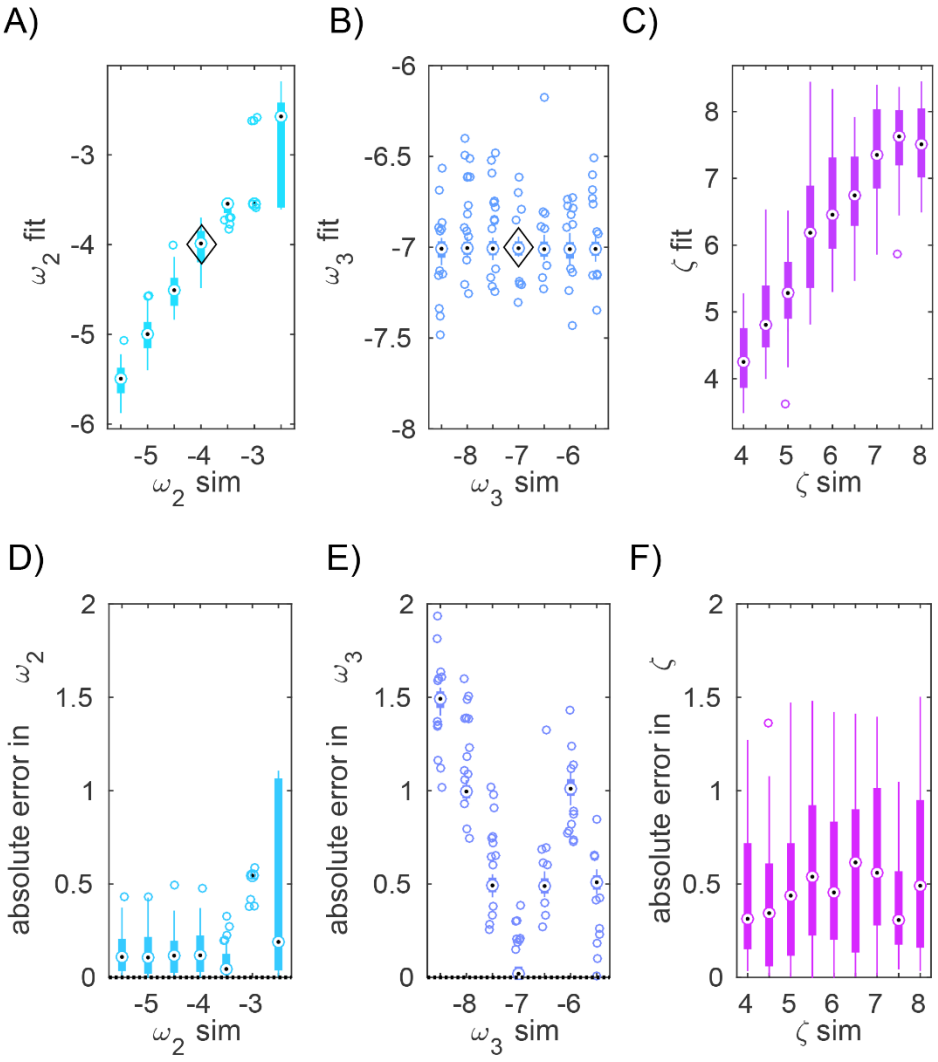
It is important to note, interpretations concerning neuroanatomical regions are limited in our EEG study as it provided exclusively sensor-level results. The anterior cingulate cortex (ACC) has been shown to contribute to encoding lower-level pwPEs in a task with a similar structure (Iglesias et al., 2013). Intriguingly, state anxiety has been shown to deactivate the ventrolateral prefrontal cortex (PFC) and rostral ACC during cognitive control tasks that crucially depend on these areas (Bishop et al., 2004). Attention bias for threat in anxiety is also associated with alterations in ACC/PFC, specifically in the connectivity between dorsal ACC / dorsomedial PFC and the amygdala (Grillon et al., 2019). Thus, one hypothesis that could be tested in future combined fMRI-EEG studies is whether state anxiety disengages these ACC and PFC regions during reward-based learning, undermining their proper contribution to tracking pwPE about stimulus outcome tendencies, but also volatility.

Of particular interest, decreased dorsolateral prefrontal cortex activity also characterises elevated trait anxiety levels, with detrimental consequences for performance and attentional control (Bishop, 2009). And portions of the cingulate cortex and prefrontal cortex are part of the central network underlying anxiety disturbances (Grupe & Nitschke, 2013). Thus, an additional interesting question for future studies would be to assess the role that these brain regions play in the modulation of hierarchically-related pwPEs that may lead to the computational biases described in trait anxiety (Browning et al., 2015; Huang et al., 2017).

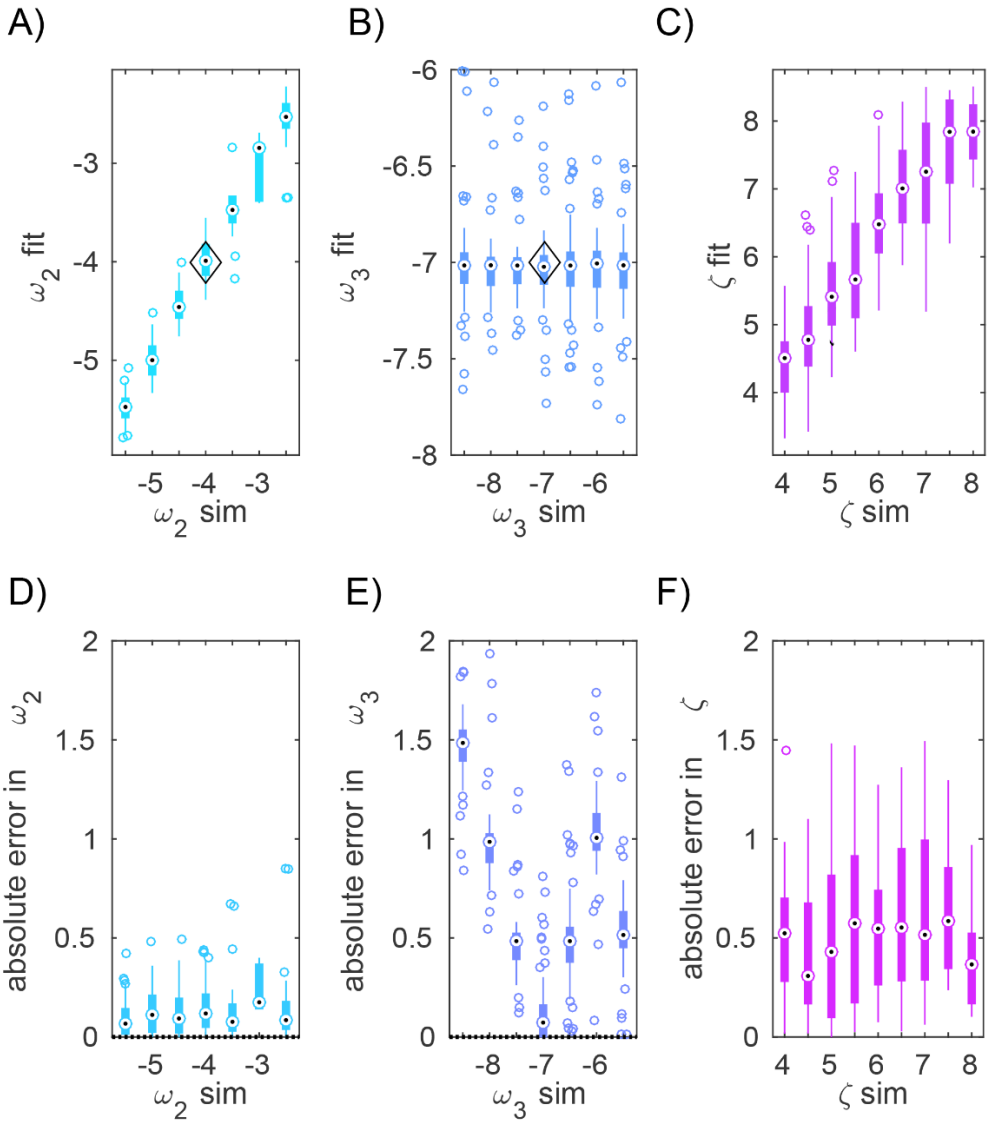
Supplementary Figures



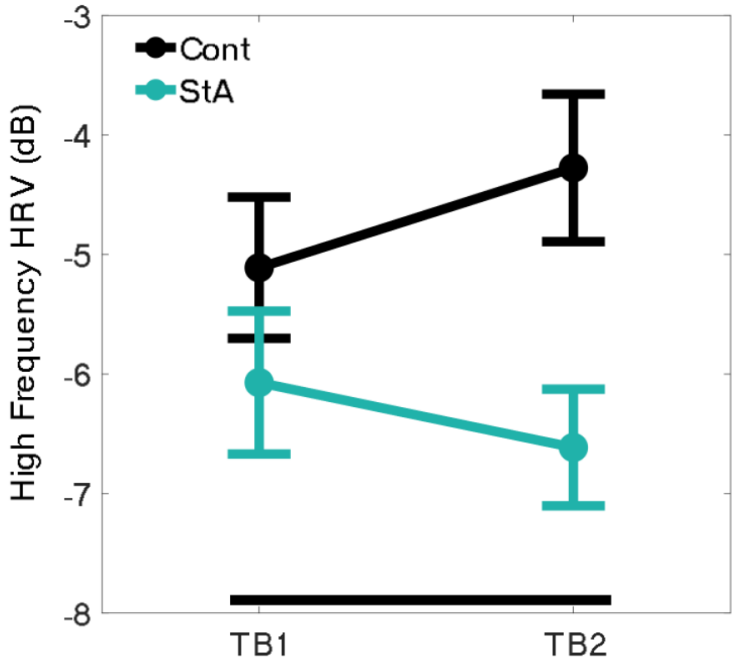
Supplementary Figure 1. Scheme of the experimental task structure. The experiment was split into four blocks: the first resting-state block (R1: baseline), task block 1 (TB1), task block 2 (TB2), and the final resting state block (R2). Both groups (StA and Cont) started with 5 minutes of resting state (R1) consisting of EEG and ECG recording. Prior to TB1, the StA group were informed of the experimental manipulation of public speaking about an unknown artwork in front of 3 experts, to be carried out after the computer learning task (TB1, TB2) had finished. This aimed to induce anxiety during TB1 and TB2 for the StA group, as indicated by hatched lines. The Cont group were told to expect to describe the same artwork to themselves for an identical period of time. Both groups undertook the reward-learning task across two blocks (TB1, TB2). After completing TB2, the StA group were informed they would not need to speak publicly about the artwork, and that they would, in an identical fashion to the Cont group, only need to present the artwork to themselves. When finished with this self-presentation, the final resting state block of ECG and EEG was recorded (R2).



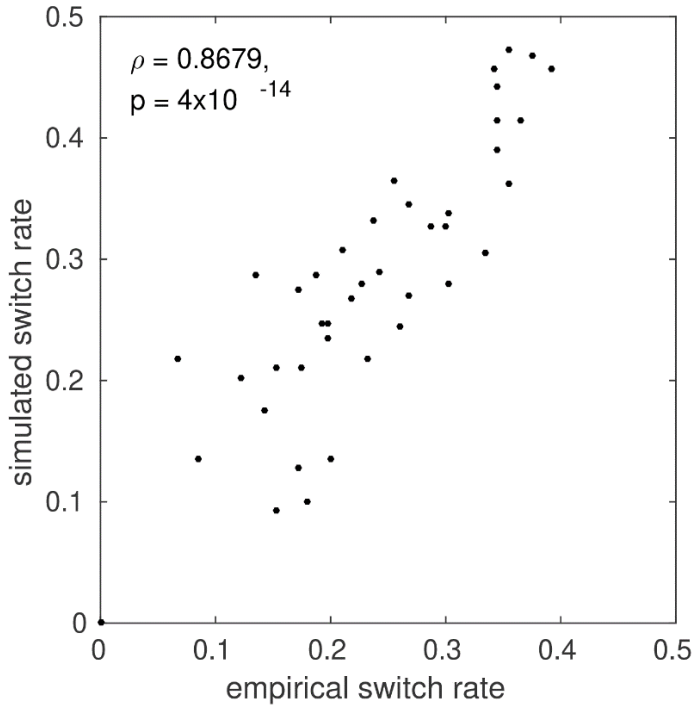
Supplementary Figure 2. HGF parameter estimation. We simulated responses in an ideal observer, receiving the input from participant #1 in the control group. The responses were simulated using as model parameter values those listed in Table 1 (fixed parameters $\mu_2^{(0)} = 0$, $\sigma_2^{(0)} = 0.1$, $\mu_3^{(0)} = 1$, $\sigma_3^{(0)} = 1$, $\kappa = 1$) but we also set the values of ω_2, ω_3 and the response model parameter ζ. Different simulated responses were created by using different values of ω_2 (−5.5 to −2.5 in steps of 0.5; 70 repetitions in each case), ω_3 (−8.5 to −6.5, in steps of 0.5, 70 repetitions), ζ (4 to 8 in steps of 0.5; 70 repetitions). Next, we fitted the simulated responses and the input with the HGF winning model used for the empirical data and estimated the parameter values that best account for the simulated behavioural data. The prior values used for ω_2 and ω_3 were −4 and −7, respectively (variance 16 in both cases). Prior values are denoted by the diamond shape in the top panels. Panels A-C are boxplots (median, 25 and 75 percentiles) illustrating the results of parameter estimation for ω_2 (A), ω_3 (B) and ζ (C). The x-axis represents the set parameters introduced in the simulated responses (labelled "sim"), while y-axis data reveal the corresponding estimated value of that same parameter (labelled "fit). Parameter recovery was excellent in the case of ω_2 and ζ, as there was a high significant correlation between simulated and estimated (fit) values: Pearson R = 0.9497, $P << 1 \times 10^{-6}$ for ω_2, R = 0.8167, $P << 1 \times 10^{-6}$. Parameter ω_3 could not be well recovered: R = 0.02, $P = 0.5383$.



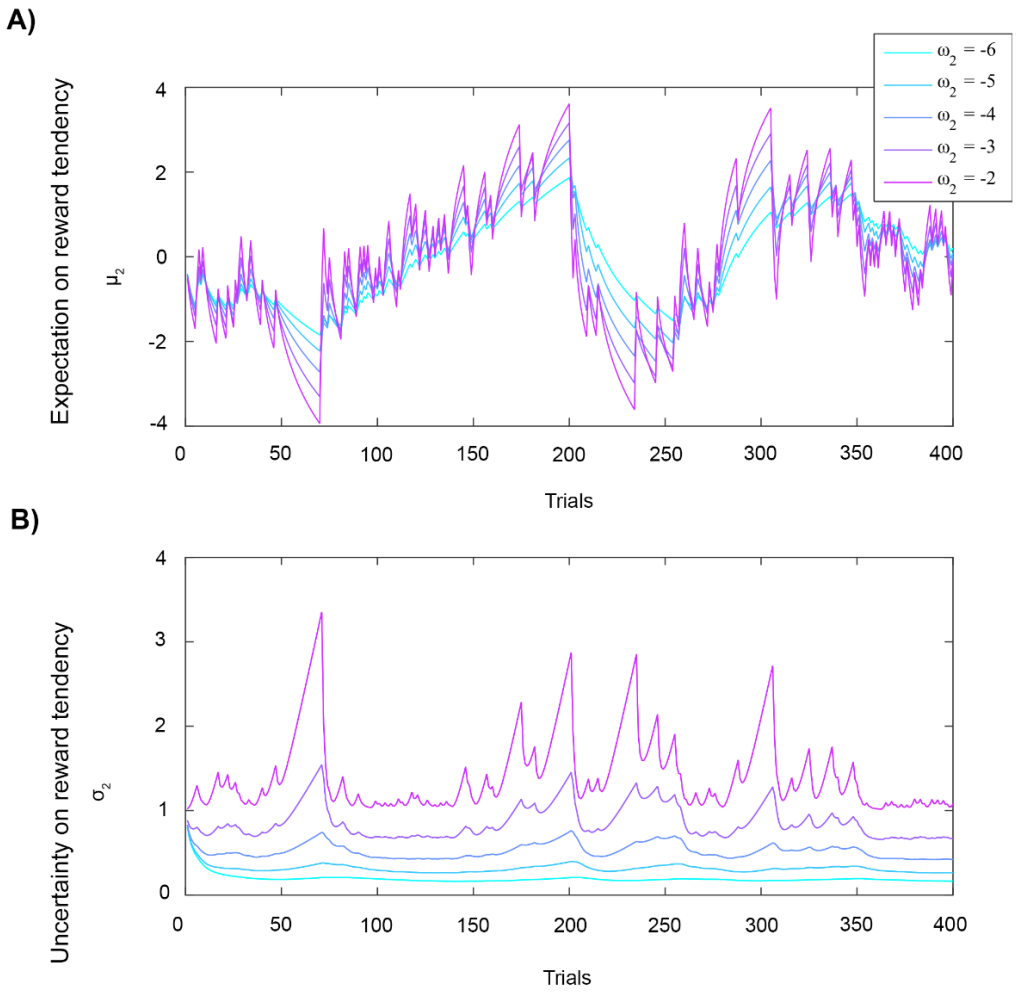
Supplementary Figure 3. HGF parameter estimation. Same as **Supplementary Figure 2** but using simulated responses in an ideal observer receiving the input from participant #1 in the StA group. Parameter recovery was excellent for ω_2 and ζ, as simulated and estimated (fit) values were highly significantly correlated: Pearson R = 0.9834, $P = 0$ for ω_2, R = 0.8225, $P << 1 \times 10^{-6}$. ω_3 could not be well recovered: R = −0.0898, $P = 0.047$.



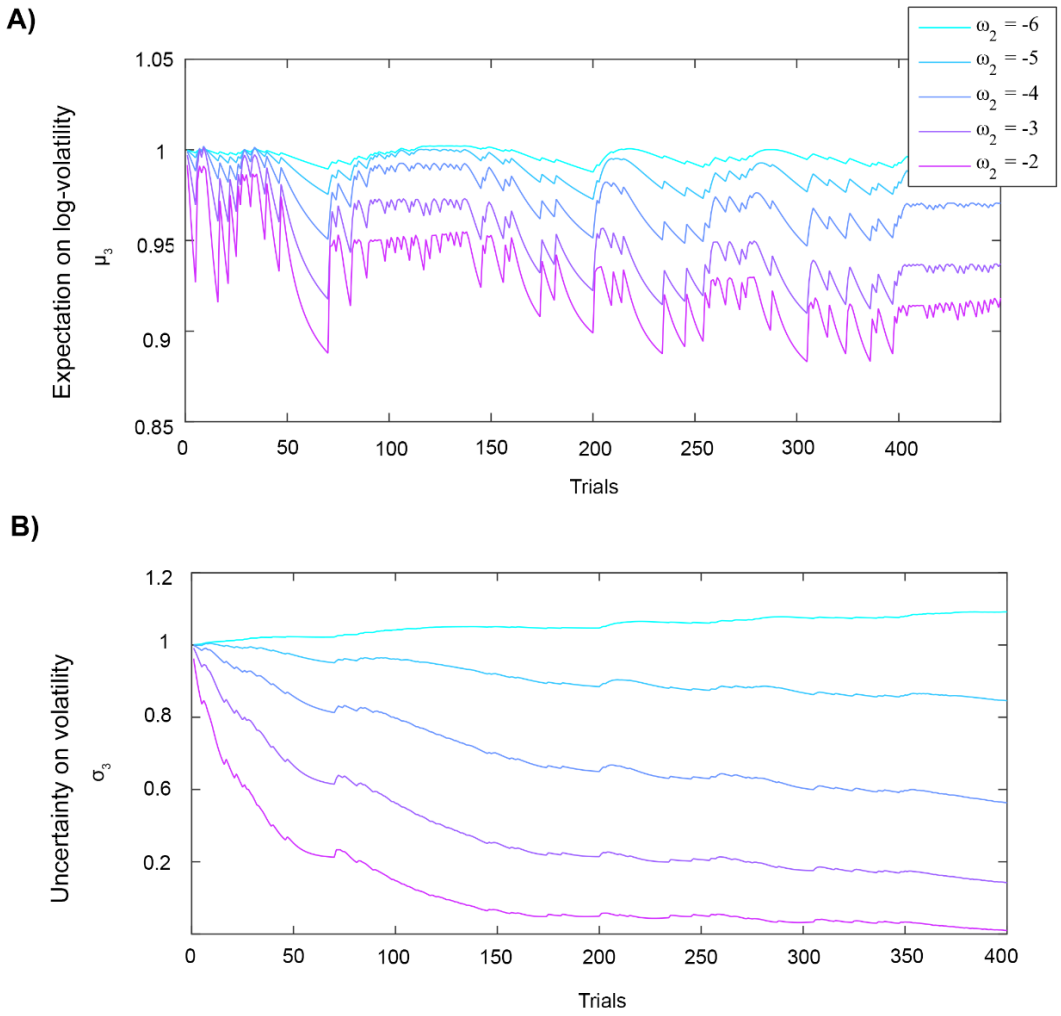
Supplementary Figure 4. Spectral analysis of IBI time series data. To complement our HRV proxy of state anxiety, an analysis of the frequency content of IBI data was performed to link our finding of reduced HRV to evidence of autonomic inflexibility (parasympathetic vagal modulation) in anxiety. The results showed reduced high frequency (0.15–0.4 Hz) content in the StA (mean −6.3, SEM 0.6) compared to Cont group (mean −4.7, SEM 0.5, *P* = 0.02). With a trend level interaction (*P* = 0.06) but no Block effect (*P* = 0.08).



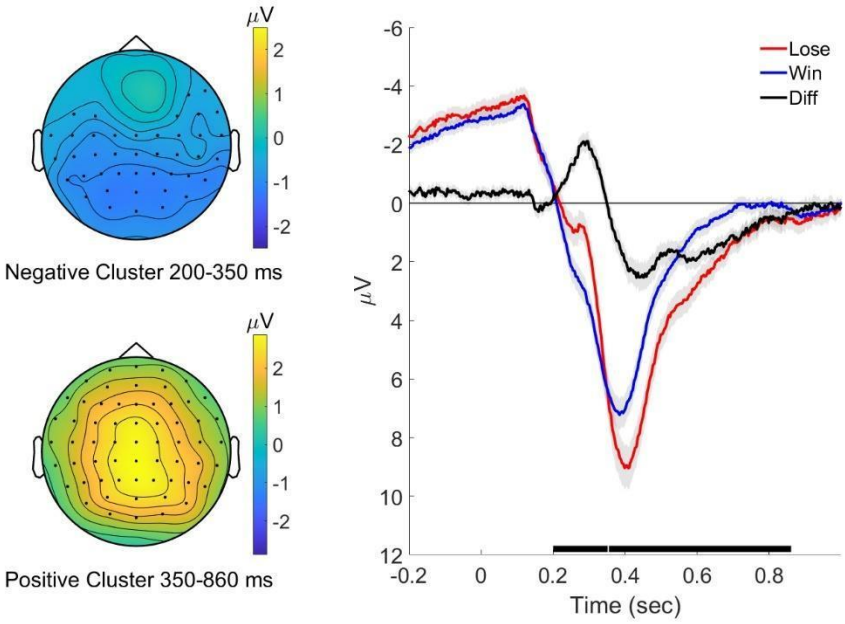
Supplementary Figure 5. Model check. Simulated responses were generated using the estimated model parameters for each individual (ω_2, ω_3, ζ). We compared the probability of trial-to-trial response switch (Orange, Blue) between simulated and empirical data across participants computing the non-parametric Spearman rank correlation. There was a very high and significant rank correlation between both variables (N = 42): $\rho = 0.8679$, $P = 4 \times 10^{-14}$. Inspection of this association within each participant group revealed similar values (N = 21 in each case): $\rho_{Cont} = 0.8263$, $P_{Cont} = 4 \times 10^{-6}$; $\rho_{StA} = 0.8355$, $P_{StA} = 2.5 \times 10^{-6}$.



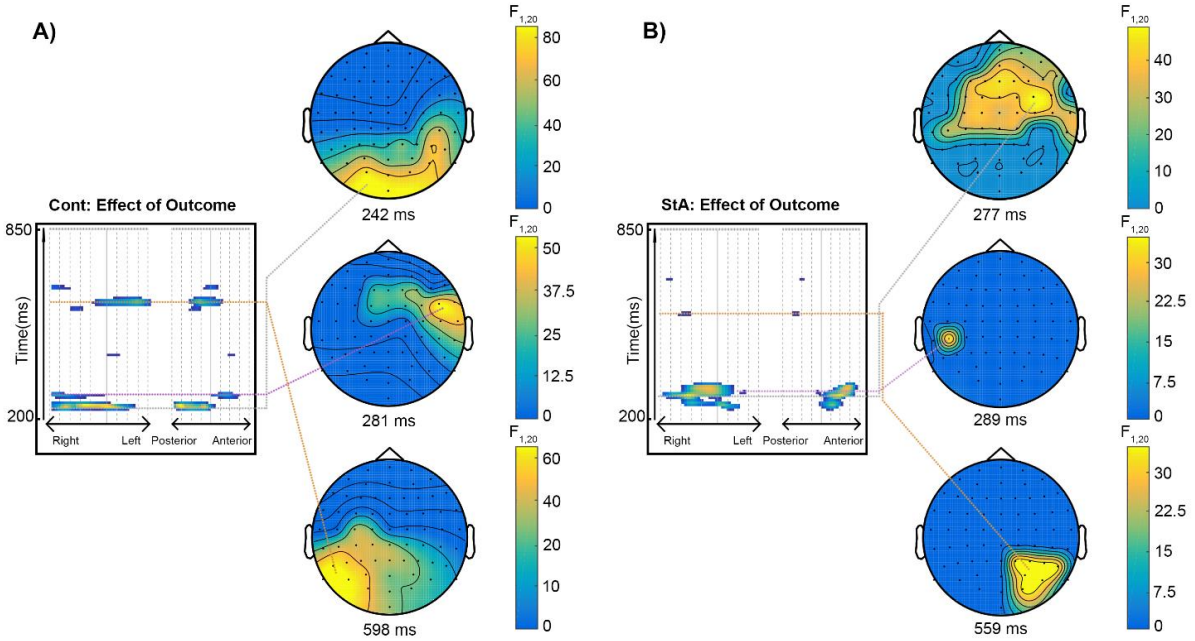
Supplementary Figure 6. Simulated belief trajectories for μ_2 and σ_2. Trajectories were simulated using the input data from participant 1 using the priors $\mu_2^{(0)} = 0$, $\sigma_2^{(0)} = 0.1$, $\mu_3^{(0)} = 1$, $\sigma_3^{(0)} = 1$, $\kappa = 1$, $\omega_3 = -7$, but modulating ω_2. This parameter represents the tonic part of the variance in the Gaussian random walk for x_2 and modulates the learning rate about stimulus outcomes at the lowest level. Here we show that **A)** decreasing ω_2 is linked to smaller update steps on the reward tendency μ_2, while **B)** decreasing ω_2 also reduces the estimation uncertainty about the reward tendency, σ_2.



Supplementary Figure 7. Simulated belief trajectories for μ_3 and σ_3. Trajectories were simulated using the input data from participant 1 using the priors $\mu_2^{(0)} = 0$, $\sigma_2^{(0)} = 0.1$, $\mu_3^{(0)} = 1$, $\sigma_3^{(0)} = 1$, $\kappa = 1$, $\omega_3 = -7$, but modulating ω_2. Here we show that decreasing ω_2 leads to **A)** smaller update steps for the expectation of log-volatility μ_3 and **B)** higher uncertainty σ_3 about volatility. As we presented in our results, StA individuals had significantly lower ω_2. This can explain why StA exhibited higher uncertainty about volatility σ_3 when compared to Cont. The StA group did, moreover, exhibit generally a higher (less reduced) expectation of volatility μ_3, despite this effect being non-significant.



Supplementary Figure 8. Results of the ERP response comparison for the main effect of outcome: wins and losses. **Left panels**. Cluster-based random permutation analysis of ERP responses was carried out in the total population (N= 42) to assess the effect of the outcome (win, lose). Maps given for each cluster show the scalp topography of the significant cluster ERP differences between outcomes (win, lose). Black dots on the topographical maps indicate electrodes pertaining to a significant cluster ($P < 0.025$, two-tailed test). **Right panel**. Grand-mean ERP waveforms of the two outcomes (lose, red; win, blue) and the difference (lose minus win, black) are presented from all electrodes between −0.2 and 1 seconds, with SEM given as grey shaded areas. Significant clusters are denoted by black bars on the x-axis.



Supplementary Figure 9. Signatures of the representation of trial outcomes on trial-wise ERPs. **A)** Effect of trial outcomes in controls (Cont) correlated with EEG response changes across parietal-occipital and central-parietal electrodes between 229 ms to 257 ms, as shown in the topographical representation at the time of the maximum peak of the cluster (242 ms post stimulus, $P_{FWE} < 0.0001$, with a cluster-defining threshold of $P < 0.001$). A smaller cluster was also found as shown in the middle topographical representation at 281 ms, with activation between 267–287 ms ($P_{FWE} = 0.002$) at right frontal and temporal electrodes. A later cluster was also shown between 583 ms to 618 ms with a peak at 598 ms over left parietal and central electrodes (PFWE < 0.0001).**B)** Effect of trial outcomes on EEG activity during state anxiety. In the state anxiety group (StA), trial outcomes were associated primarily with trial-wise ERP changes in right frontocentral electrodes. This effect, ranging from 228 ms and 316 ms, is shown in a topographic scalp map at the time of the maximum peak of the cluster (277 ms post stimulus, $P_{FWE} = < 0.0001$, with a cluster-defining threshold of $P < 0.001$). A second later cluster peaking at 371ms had a left parietal distribution. An additional smaller significant cluster was found between 286–297 ms in right parietal electrodes peaking at 289 ms (PFWE = 0.04), and later between 558–569 ms peaking at 559 ms (PFWE = 0.036).

Supplementary Materials

Results: Reaction time by contingency phase

A study by Prinzmetal et al. (2009) demonstrated that both voluntary and involuntary attention impact reaction times. Voluntary (exogenous) and involuntary (automatic or endogenous) attention were assessed in a spatial cueing task developed by Posner and colleagues (Posner, 1980). Voluntary attention was evaluated in trials with informative or predictive cues (80% valid trials) whereas involuntary attention was assessed using uninformative cues (25% of valid trials, for four possible target locations). The authors found that both voluntary and involuntary attention shorten RT (lack of attention, therefore, increases RT), and voluntary attention additionally affects accuracy. Follow up studies have further validated this effect of attention on RT.

In our study, RT was not affected by the anxiety manipulation, which argues against “classical” attention being the main driving factor of the altered learning effects. Similarly, a previous study explicitly assessed the relation between attention and anxiety and showed that state anxiety was not linked to slower reaction times (Bishop, 2009).

We additionally assessed RT separately in blocks of unpredictable cues (50–50 contingency phase) and highly predictable cues (90–10 contingency phase) and assessed how they were affected by anxiety. We averaged the RT across all 50–50 contingency trials and separately 90–10 trials and performed a pair-wise permutation test comparing the mean RT between StA and Cont groups separately for each contingency category. The results demonstrate that in the 90–10 contingency phase there is no significant difference between StA (M = 667, SEM = 34) and Cont (M = 641, SEM = 49, $P = 0.76$). Likewise, there was no difference in the 50–50 contingency phase between the StA group (M = 683, SEM = 40) and Cont group (M = 684, SEM = 58, P = 0.98).

Efficiency of β coefficients in the GLM model

Collinearity of regressors was an issue in our initial GLM design, as abs(ε_2) and ε_3 were highly correlated. A common practice is to orthogonalise collinear regressors in the model to solve the problem of reduced power and unreliable parameter estimates in the GLM (Mumford et al., 2018). However, other authors argue that despite the potential appeal of orthogonalisation of regressors to remove collinearity from the model, the implications are actually not necessarily beneficial: it does not improve the overall fit of the model, and in most cases, it can lead to a misleading interpretation of the resulting inferences (Vanhove, 2020). We followed this second line of argument and chose to keep only one of the pwPE regressor trajectories instead. To inform our decision, we calculated the **efficiency for β coefficients** as proposed by Mumford et al. (2018) as a useful index to assess a priori our chosen explanatory variables (abs(ε_2) and ε_3):

$$eff(\beta_1) = (N-1)Var(x_1)\big(1 - cor^2(x_1, x_2)\big),$$

The equation above is taken from Mumford et al. (2018) and is valid for a GLM with two regressors, x_1 and x_2, associated with two β coefficients β_1 and β_2. N is the number of observations (400 trials in our case).

We estimated the efficiency for β_1 (regression coefficient for abs(ε_2)) and β_2 (regression coefficient for ε_3) in each control and experimental group separately, and we obtained the following (mean and SEM across subjects):

Control group:
Efficiency for β_1 (ε_2): 47 (16)
Efficiency for β_2 (ε_3): 0.33 (0.03)

StA group:
Efficiency for β_1 (ε_2): 15.7 (6.9)
Efficiency for β_2 (ε_3): 0.57 (0.08)

Although the exact values of the efficiency indices are not informative, what this analysis reveals is that the efficiency for the β regression coefficient associated with abs(ε_2) is much higher than for ε_3, due to larger variance in abs(ε_2). Accordingly, we kept abs(ε_2) in our GLM analysis and excluded regressor ε_3. In addition, as indicated in the main manuscript, we used

regressor outcome (win/lose trials), as it was expected to explain a large proportion of the variance in the EEG data.

A similar analysis of the efficiency of regression coefficients for the final GLM model using abs(ε_2) (β_1) and outcome (β_2), demonstrated that efficiency for beta coefficients was in the same order of magnitude for both regressors:

Control group:
Efficiency for β_1 (abs(ε_2)): 94 (10)
Efficiency for β_2 (outcome): 62 (2)

StA group:
Efficiency for β_1 (abs(ε_2)): 34 (12)
Efficiency for β_2 (outcome): 63 (5)

Notably, while the efficiency for the β regression coefficient associated with the outcome regressor was very similar in both groups, the efficiency for β_1 associated with abs(ε_2) was considerably lower in the StA group, relative to control participants.

Modulations of single-trial ERPs by trial outcome.

Trial outcomes were represented by a large cluster of activity in both the Cont and StA group and with a similar latency. In the Cont group, trial outcomes significantly modulated EEG activity from 229 ms to 257 ms post stimulus over parietal-occipital and central-parietal electrodes, with a maximum effect at 242 ms in central parietal-occipital sites ($P_{FWE} < 0.0001$). A second cluster of similar size occurred between 583 ms to 618 ms with a peak at 598 ms over left parietal and central electrodes ($P_{FWE} < 0.0001$). Four further significant effects of a smaller cluster size were found earlier between 267–287 ms ($P_{FWE} = 0.002$) and 642-654 ms ($P_{FWE} = 0.018$) in right frontal and temporal electrodes, and between 567–583 ms ($P_{FWE} = 0.016$) in right parietal electrodes, and in frontocentral electrodes between 408–424 ms ($P_{FWE} = 0.028$, see **Supplementary Figure 9;** Details on the cluster effects can be found in **Table 2**).

In the StA group, trial outcomes significantly modulated trial-wise EEG responses from 228 ms to 316 ms across frontocentral regions, with a maximum effect at 277 ms in right frontal central electrodes ($P_{FWE} < 0.0001$, **Supplementary Figure 9**). An additional smaller significant cluster was found between 286–297 ms in right parietal electrodes ($P_{FWE} = 0.04$), and later between 558–569 ms ($P_{FWE} = 0.036$) and 669–681 ms ($P_{FWE} = 0.042$) in right parietal electrodes (cluster effects can be found in **Table 2**). There were no between-group differences in the representation of trial outcomes in EEG activity.

Model	Prior	Mean	Variance
3-level HGF	κ	1	0
	ω_2	-4	16
	ω_3	-7	16
	$\mu_2^{(0)}$	0	0
	$\sigma_2^{(0)}$	0.1	0
	$\mu_3^{(0)}$	1	0
	$\sigma_3^{(0)}$	1	0
	ζ	48	1
2-level HGF	κ	0	0
	ω_2	-4	16
	ω_3	-7	0
	$\mu_2^{(0)}$	0	0
	$\sigma_2^{(0)}$	0.1	0
	$\mu_3^{(0)}$	1	0
	$\sigma_3^{(0)}$	1	0
	ζ	48	1
HGFμ_3	κ	1	0
	ω_2	-4	16
	ω_3	-7	16
	$\mu_2^{(0)}$	0	0
	$\sigma_2^{(0)}$	0.1	0
	$\mu_3^{(0)}$	1	1
	$\sigma_3^{(0)}$	1	1

Table 1. Means and variances of the priors on perceptual parameters and starting values of the beliefs of the HGF models. Values are shown for 3-level HGF, 2-level HGF and HGFμ_3 models. Free parameters are estimated in their unbounded space. Accordingly, parameters that are restricted to a confined interval are log-transformed, to allow for estimation in an unbounded space. In our study, ζ was estimated in the log space (3-level HGF and 2-level HGF models). Model HGFμ_3 had as free parameters ω_2, ω_3, $\mu_3^{(0)}$, and $\sigma_3^{(0)}$ with μ_3 governing the decision noise through a negative exponential (Diaconescu et al., 2014). Here, $\sigma_3^{(0)}$ was estimated in the log space.

	Activation size (voxels)	Cluster *p* value (FWE corrected)	Peak *p* value (FWE corrected)	Peak *F* statistic	Peak equivalent *Z* statistic	Peak latency (ms)
Control Group: Lower-level pwPE: abs(ε_2)						
	1194	0.000	0.004	52.70	4.89	496
	342	0.001	0.021	38.36	4.43	457
State Anxiety Group: Outcome						
	2293	0.000	0.003	49.19	4.70	277
	10	0.040	0.041	30.26	4.09	289
	16	0.036	0.041	30.23	4.08	559
	6	0.042	0.043	29.89	4.07	680
Control Group: Outcome						
	1305	0.000	0.000	80.75	5.50	242
	1110	0.000	0.002	58.35	5.04	598
	298	0.002	0.003	51.97	4.87	281
	55	0.018	0.019	37.50	4.40	648
	63	0.016	0.025	35.60	4.32	574
	24	0.028	0,038	32.83	4.20	414

Table 2. Test statistics for lower-level precision-weighted prediction errors and trial outcomes. Each significant activation is ordered according to size (leftmost column). We provide both the cluster and peak *p* values with the family-wise error correction applied. Also given are the relevant statistics (*F* and peak equivalent *Z*) for each activation cluster and within each activation.

Chapter 3: State anxiety alters the neural oscillatory correlates of predictions and prediction errors during reward-based learning

Hein, T. P., & Ruiz, M. H. (2021). State anxiety alters the neural oscillatory correlates of predictions and prediction errors during reward learning. bioRxiv.[8]

[8] *This chapter is based on a manuscript that is currently under review and has been published on a preprint server.: doi/10.1101/2021.03.08.434415.*

Author contributions M.H.R. and T.P.H. designed the experiment, T.P.H collected the data, M.H.R. and T.P.H. analysed the data, M.H.R., and T.P.H. wrote code for data analysis, M.H.R. and T.P.H. wrote the manuscript.

Abstract

Anxiety influences how the brain estimates and responds to uncertainty. The consequences of these processes on behaviour have been described in theoretical and empirical studies, yet the associated neural correlates remain unclear. Rhythm-based accounts of Bayesian predictive coding propose that predictions in generative models of perception are represented in alpha (8–12 Hz) and beta oscillations (13–30 Hz). Updates to predictions are driven by prediction errors weighted by precision (inverse variance), and are encoded in gamma oscillations (>30 Hz) and associated with suppression of beta activity. We tested whether state anxiety alters the neural oscillatory activity associated with predictions and precision-weighted prediction errors (pwPE) during learning. Healthy human participants performed a probabilistic reward-based learning task in a volatile environment. In our previous work, we described learning behaviour in this task using a hierarchical Bayesian model, revealing more precise (biased) beliefs about the tendency of the reward contingency in state anxiety, consistent with reduced learning in this group. The model provided trajectories of predictions and pwPEs for the current study, allowing us to assess their parametric effects on the time-frequency representations of EEG data. Using convolution modelling for oscillatory responses, we found that, relative to a control group, state anxiety increased beta activity in frontal and sensorimotor regions during processing of pwPE, and in fronto-parietal regions during encoding of predictions. No effects of state anxiety on gamma modulation were found. Our findings expand prior evidence on the oscillatory representations of predictions and pwPEs into the reward-based learning domain. The results suggest that state anxiety modulates beta-band oscillatory correlates of pwPE and predictions in generative models, providing insights into the neural processes associated with biased belief updating and poorer learning.

Keywords Anxiety, predictive coding, oscillations, EEG, convolution, uncertainty

1. Introduction

Affective states closely interact with decision making (Lerner et al., 2015). For example, altered computations—such as learning rates and estimates of belief uncertainty—during decision making are considered central to explaining clinical conditions including anxiety, depression and stress from a Bayesian predictive coding (Bayesian PC) perspective (Browning et al., 2015; de Berker et al., 2016; Paulus & Yu, 2012; Pulcu & Browning, 2019; Williams, 2016). The Bayesian PC framework proposes that the brain continuously updates a hierarchical generative model using predictions optimised through their discrepancy with sensory data—prediction errors (PE)—and weighted by precision (inverse variance; Friston, 2010; Rao & Ballard, 1999; Srinivasan et al., 1982). This hierarchical message passing was hypothesised—in the context of sensory processing—to be mediated by neural oscillations at specific frequencies, in distinct cortical layers and regions (Bastos et al., 2012). Empirical evidence supports this, identifying predictions in alpha and beta frequencies and PEs in gamma frequencies (Arnal & Giraud, 2012; Auksztulewicz et al., 2017; Bastos et al., 2020; Sedley et al., 2016). Yet how affective states modulate the oscillatory activity associated with predictions and PE signals has been largely overlooked.

Uncertainty makes refining predictions particularly challenging. Estimates of uncertainty (or its inverse, precision) regulate how influential PEs are on updating our generative model of the environment (Friston, 2008; Yu & Dayan, 2005), scaling precision-weighted PEs (pwPEs). Uncertain and changing environments may render prior beliefs obsolete, down-weighting predictions in favour of increasing learning about sensory input. Recent studies have highlighted that precision estimates are important in explaining atypical learning and perception in neuropsychiatric conditions (Fletcher & Frith, 2009; Friston et al., 2013; Lawson et al., 2014; Montague et al., 2012). Anxiety, in particular, has been shown to lead to insufficient adaptation in the face of environmental change (Browning et al., 2015; Huang et al., 2017), disruption in learning and maladaptive biases—in both aversive and reward-based learning contexts (Hein et al., 2021; Huang et al., 2017; Kim et al., 2020; Lamba et al., 2020; Piray, Ly, et al., 2019; Pulcu & Browning, 2019). Whether the learning alterations in anxiety are mediated by oscillatory changes representing predictions and pwPEs remains unknown.

Within the Bayesian PC framework, growing evidence supports that, during perception, feedforward PE signals are encoded by gamma oscillations (>30 Hz), while backward connections convey predictions expressed in alpha (8–12 Hz) and beta (13–30 Hz) oscillations (Arnal & Giraud, 2012; Bastos et al., 2015; van Pelt et al., 2016; Wang, 2010). Precision weights are also modulated by alpha and beta oscillations (Palmer et al., 2019; Sedley et al.,

2016). Because precision weights scale PEs during Bayesian inference and learning (Feldman & Friston, 2010; Friston, 2010), the modulation of pwPE signals could be expressed both in gamma and alpha/beta activity, as recent work suggests (Auksztulewicz et al., 2017). Specifically, gamma increases during pwPE encoding are accompanied by attenuated alpha and beta activity (Auksztulewicz et al., 2017). Indeed, gamma and alpha/beta oscillatory rhythms are anticorrelated across the cortex, as shown in investigations of prediction violations (Bastos et al., 2018, 2020) and during working memory (Lundqvist et al., 2016, 2020; Miller et al., 2018). Complementing these findings, we recently showed that abnormal increases in beta power and burst rate can account for the dampening of pwPEs during reward-based motor learning in anxiety (Sporn et al., 2020). Moreover, gamma oscillations in the dorsomedial prefrontal cortex (dmPFC) are associated with unsigned prediction errors during exploration-exploitation behaviour (Domenech et al., 2020), suggesting cortical gamma activity as a relevant correlate of reward-based learning. Accordingly, we speculated that decreased anxiety-related learning during decision making could be associated with abnormally enhanced beta, in addition to reduced gamma, during processing pwPEs.

Predictions have been consistently associated with the modulation of alpha-beta rhythms across multiple modalities, such as visual (Gould et al., 2011), motor (Schoffelen et al., 2005), somatosensory (van Ede et al., 2011), and auditory (Todorovic et al., 2015)—yet frequency-domain evidence for predictions about reward contingencies in volatile environments is currently lacking. This is important to understand learning biases that manifest in anxiety conditions during environmental instability (Browning et al., 2015; Pulcu & Browning, 2019). Crucially, predictions in deep layers are thought to functionally inhibit the processing of sensory input and PEs in superficial layers (Bastos et al., 2015; Bauer et al., 2014; Mayer et al., 2016; Van Kerkoerle et al., 2014). This suggests that aberrant oscillatory states modulating predictions would be an additional route through which encoding of pwPEs is altered, contributing to impaired learning.

Here, we used convolution modelling of oscillatory responses (Litvak et al., 2013) in previously acquired EEG data to estimate the neural oscillatory representations of predictions and pwPEs during reward-based learning in healthy controls and a state anxious group. Our previous computational modelling study (Hein et al., 2021) revealed that state anxiety biases uncertainty estimates, increasing the precision of posterior beliefs about the stimulus-reward contingency. We now ask whether this bias is associated with altered spectral characteristics of hierarchical message passing, which could represent a candidate marker of biased belief updating and poorer reward-based learning in anxiety. We hypothesised that, in state anxiety, increased precision in the predictions about a certain stimulus-reward contingency should be

associated with increased alpha and beta activity. This, in turn, would inhibit the processing of expected inputs in line with PC accounts, resulting in a hypothesised lower gamma activity and concomitantly higher alpha-beta activity for attenuating encoding of pwPEs.

2. Materials & Methods

2.1 Participant sample

The data used in the preparation of this work were obtained from our previous study Hein et al. (2021), which was approved by the ethical review committee at Goldsmiths, University of London. Participants were pseudo-randomly allocated into an experimental state anxiety (StA) and control (Cont) group, following a screening phase in which we measured trait anxiety levels in each participant using Spielberger's Trait Anxiety Inventory (STAI; Spielberger [1983]). Trait anxiety levels were matched in StA and Cont groups: average score and standard error of the mean, SEM: 47 [2.1] in StA, 46 [2.2] in Cont). Importantly, the individual trait anxiety scores were lower than the previously reported clinical level for the general adult population (> 70, Spielberger et al., 1983). Further, the age of the control group (mean 27.7, SEM = 1.2) and their sex (13 female, 8 male) were consistent with those from the state anxiety group (mean 27.5, SEM = 1.3, sex 14 female, 7 male). This is important to consider as there are known age and sex-related confounds to measures of state anxiety (see Voss et al., 2015).

2.2 Experimental design

Both groups (StA, Cont) performed a probabilistic binary reward-based learning task where the probability of reward between two images changes across time (Behrens et al., 2007; de Berker et al., 2016; Iglesias et al., 2013). The experiment was divided into four blocks: an initial resting state block (R1: baseline), two reward-based learning task blocks (TB1, TB2), and a final resting state block (R2). Each resting state block was 5 minutes. Participants were instructed to relax and keep their eyes open and fixated on a cross in the middle of the presentation screen while we recorded EEG responses from the scalp and EKG responses from the heart.

The experimental task consisted of 200 trials in each task block (TB1, TB2). The aim was for participants to maximise reward across all trials by predicting which of the two images (blue, orange) would reward them (win, positive reinforcement, 5 pence reward) or not (lose, 0 pence reward). The probability governing reward for each stimulus (reciprocal: p, 1−p) changed

across the experiment, every 26 to 38 trials. There were 10 contingency mappings for both task blocks: 2 x strongly biased (90/10; i.e. probability of reward for blue p = 0.9), 2 x moderately biased (70/30), and 2 x unbiased (50/50: as in de Berker et al., 2016). The biased mappings repeated in reverse relationships (2 x 10/90; 2 x 30/70) to ensure that over the two blocks (TB1, TB2) there were 10 stimulus-outcome contingency phases in total.

In each trial the stimuli were presented randomly to the left or right of the centre of the screen where they remained until either a response was given (left, right) or the trial expired (maximum waiting time, 2200 ms $\pm$ 200 ms). Next, the chosen image was highlighted in bright green for 1200 ms ($\pm$ 200 ms) before the outcome (win, green; lose or no response, red) was shown in the middle of the screen (1200 ms $\pm$ 200 ms). At the end of each trial, the outcome was replaced by a fixation cross at an inter-trial interval of 1250 ms (± 250 ms).

Specific task instructions to participants were to select which image they predicted would reward them on each trial and adjust their predictions according to inferred changes in the probability of reward (as in de Berker et al., 2016). All participants filled out computerised questionnaires (state anxiety STAI state scale X1, 20 items: Spielberger, [1983a]) and conducted practice trials as detailed in Hein et al. (2021). Critically, the state anxiety manipulation was delivered just before the first reward-based learning block (TB1) to the StA group (see the following section).

2.3 Manipulation and assessment of state anxiety

Our StA group was instructed to complete a public speaking task in line with previous work (Feldman et al., 2004; Lang et al., 2015). This meant, as detailed in Hein et al. (2021), that StA participants were told just before TB1 that they would need to present a piece of abstract art for 5 minutes to a panel of academic experts after completing the reward-based learning task, with 3 minutes preparation time. By contrast, the Cont group were informed that they would need to give a mental description of the piece of abstract artwork for the same time privately (rather than to a panel of experts, see Hein et al., 2021). Importantly, the state anxiety manipulation was then revoked in the StA group directly after completing the second reward-based learning block (TB2) and before the second resting state block (R2). They were informed that the panel of experts was suddenly unavailable. Both groups, therefore, presented the artwork to themselves after completing the reward-based learning task.

To assess state anxiety, in our previous work we used the coefficient of variation (CV = standard deviation/mean) of the inter-beat intervals (IBI) as a metric of heart rate variability

(HRV), as this index has been shown to drop during anxious states (Chalmers et al., 2014; Feldman et al., 2004; Gorman and Sloan, 2000; Kawachi et al., 1995; Quintana et al., 2016). Additional to this, the spectral characteristics of the IBI data were analysed to obtain an HRV proxy of state anxiety associated with autonomic modulation and parasympathetic (vagal) withdrawal (Friedman, 2007; Gorman and Sloan, 2000). HRV and high-frequency HRV (HF-HRV, 0.15–0.40 Hz) measures were derived from the R-peaks extracted from the EKG signal recorded throughout the experimental sessions (see details in Hein et al., 2021, and section **EEG acquisition and analysis** below). The HRV and HRV-HF measures during performance blocks were normalised with the average baseline levels during R1, after we established that StA and Cont groups did not differ in these indexes in the initial resting state phase (P = 0.76, 0.66 for HRV and HRV-HF, respectively). This outcome suggested that control and anxious participants were not significantly dissociated in these physiological measures at the beginning of the experiment. Hereafter we refer to R1-normalised measures when summarising the results from Hein et al. (2021) on the HRV/HRV-HF measures during task blocks.

In line with prior research, our previous study showed reduced HF-HRV and reduced HRV in state anxious participants relative to controls (**Figure 1C**). Reduction in these measures has been reliably shown across trait anxiety, worry, and anxiety disorders (Aikins & Craske, 2010; Friedman, 2007; Fuller, 1992; Klein et al., 1995; Miu et al., 2009; Mujica-Parodi et al., 2009; Pittig et al., 2013; Thayer et al., 1996), and thus, significant changes to these metrics suggested physiological responses consistent with state anxiety. Subjective self-reported measures of state anxiety (STAI state scale X1, 20 items: Spielberger, 1983) were taken at four points during the original Hein et al. (2021) study, but the data could not be used due to an error in STAI data collection. We showed in a separate study, however, that HRV can effectively track changes in state anxiety, as validated by concurrent changes in STAI scores (state scale; Sporn et al., 2020).

2.4 Behavioural analysis and modelling

The behavioural data in our paradigm were analysed in Hein et al. (2021) using the Hierarchical Gaussian Filter (HGF, Mathys et al., 2011, 2014). This model describes hierarchically structured learning across various levels, corresponding to hidden states of the environment $x_1^{(k)}$, $x_2^{(k)}$,..., $x_n^{(k)}$ and defined as coupled Gaussian random walks. Belief updating on each level is driven by PEs modulated by precision ratios, weighting the influence of precision or uncertainty in the current level and the level below. The HGF was implemented with the open-source software in TAPAS http://www.translationalneuromodeling.org/tapas, version 3.1.0).

To model learning about the tendency towards reward for blue/orange stimuli and the rate of change in that tendency (volatility), we used three alternative HGF models, and two reinforcement learning models (Hein et al., 2021). The input to the models was the series of 400 outcomes and the participant's responses. Outcomes in trial *k* were either $u^{(k)} = 1$ if the blue image was rewarded or $u^{(k)} = 0$ if the orange image was rewarded. Trial responses were defined as $y^{(k)} = 1$ if participants chose the blue image, while $y^{(k)} = 0$ corresponded to the choice of the orange image. We tested a 3-level HGF (HGF_3, with volatility estimated on the third level), a 2-level reduced HGF (HGF_2, that fixes volatility to a constant level), and a HGF where decisions are informed by trial-wise estimates of volatility (see Diaconescu et al., 2014). We additionally tested two widely used reinforcement models, a Rescorla Wagner (RW, Rescorla & Wagner, 1972) and Sutton K1 model (SK1, Sutton, 1992). Following random effects Bayesian model comparison, the model that best explained the behavioural data among participants was the 3-level HGF for binary outcomes (see **Figure 1A**). In this winning model, the first level represents the binary outcome in a trial (either blue or orange wins) and beliefs on this level feature expected or irreducible uncertainty due to the probabilistic nature of the rewarded outcome (Soltani & Izquierdo, 2019). The second level $x_2^{(k)}$ represents the true tendency for either image (blue, orange) to be rewarding on trial *k*. And the third level represents the log-volatility or rate of change of reward tendencies (Bland & Schaefer, 2012; Yu & Dayan, 2005). In the HGF update equations, the second and third level states, $x_2^{(k)}$ and $x_3^{(k)}$, are modelled as continuous variables evolving as Gaussian random walks coupled through their variance (inverse precision). Hereafter we drop the trial index *k* in most expressions for simplicity.

Variational inversion of the model provides the trial-wise trajectories of the sufficient statistics of the posterior distribution of beliefs about x_i (i = 2,3): μ_i (mean, denoting participant's expectation) and σ_i (variance, termed informational or estimation uncertainty for level 2; uncertainty about volatility for level 3). The coupling function between levels 2 and 3 is as follows:

$$f_2(x_3) \stackrel{\text{def}}{=} exp(\kappa x_3 + \omega_2) \qquad \text{(Eq.1)}$$

In equation (1), ω_2 represents the invariant (tonic) portion of the log volatility of x_2 and captures the size of each individual's stimulus-outcome belief update independent of x_3. The κ parameter establishes the strength of the coupling between x_2 and x_3, and thus the degree to which estimated environmental volatility impacts the learning rate about the stimulus-outcome probabilities—κ in Hein et al. (2021) was fixed to one.

Another relevant parameter in the HGF equations is ω_3, which seizes upon 'metavolatility': how estimates of environmental volatility evolve—with larger values articulating a belief that the changeability of the task is itself changing. Note that in our experimental task, however, the rate of change (true volatility) was constant, as the stimulus-outcome contingencies changed every 26–38 trials (similarly to de Berker et al. [2016] and Iglesias et al. [2013]). Environmental uncertainty is defined as $exp(\kappa\mu_3^{(k-1)}+\omega_2)$, which depends on the phasic log-volatility estimates on the previous trial ($\mu_3^{(k-1)}$) and the tonic volatility (ω_2). Thus, the higher $\mu_3^{(k-1)}$ or ω_2 are, the greater the environmental uncertainty (see Mathys et al., 2014, page 15, Eq. 11).

In our implementation of the winning model, the 3-level HGF, we estimated the perceptual model parameters ω_2, ω_3, while we fixed κ and the initial values of the mean and variance of the belief trajectories ($\mu_2^{(0)}$, $\mu_3^{(0)}$, $\sigma_2^{(0)}$, $\sigma_3^{(0)}$). This choice was based on the previous work that we used as reference for our study (de Berker et al., 2016). The prior values on the model parameters can be found in **Supplementary Table 1** and Hein et al. (2021). Hein et al (2021) also includes the results of simulations carried out to assess how well the HGF_3 estimated each free model parameter. In brief, ω_2 could be estimated well, whereas ω_3 was not recovered, in line with recent findings (Reed et al., 2020).

Paired with this perceptual model of hierarchically-related beliefs is a response model that obtains the most likely response for each trial using the belief estimates. The winning HGF model from Hein et al. (2021) used the unit-square sigmoid observation model for binary responses (Iglesias et al., 2013; Mathys et al., 2011, 2014) and the response model parameter ζ, which represents decision noise, was additionally estimated for each participant (see **Supplementary Table 1**). Simulations carried out in Hein et al. (2021) revealed that the decision noise parameter ζ was also estimated well. We refer the reader to the original HGF methods papers for more detail on the mathematical derivations (Mathys et al., 2011, 2014), and to Hein et al. (2021) for equations included in the original results.

In the current study, we used two types of subject-specific trajectories of HGF variables as parametric regressors for convolution GLM analysis: (a) unsigned predictions about the tendency towards a certain stimulus-reward contingency ($|\hat{\mu}_2|$); (b) precision-weighted prediction errors on level 2 ($|\varepsilon_2|$) updating the beliefs on the tendency towards a reward contingency. The arguments supporting our choice of unsigned (absolute) values for these computational quantities are given in section **Spectral Analysis** below. The update steps for

the posterior mean on level 2 on trial k, $\mu_2^{(k)}$, depend on the prediction error on the level below, $\delta_1^{(k)}$, weighted by a precision term according to the following expression:

$$\mu_2^{(k)} = \mu_2^{(k-1)} + \frac{1}{\pi_2^{(k)}}\delta_1^{(k)}. \quad \text{(Eq.2)}$$

The prediction about the tendency towards a stimulus-reward contingency before observing the outcome, $\hat{\mu}_2^{(k)}$, is, in our winning model, the expectation in the previous trial, $\mu_2^{(k-1)}$. The pwPE term on level 2, on the other hand, is the product of the estimation uncertainty (inverse precision π_2) and the PE about the stimulus outcome: $\sigma_2^{(k)}\,\delta_1^{(k)}$. Thus, the influence of PEs on updating μ_2 decreases with greater precision on that level, π_2 , or smaller estimation uncertainty, σ_2. The intuition from this expression is that the less certain we are about level 2, the more we should update that level using new information (prediction errors) from the level below. See Mathys et al. (2011, 2014) for detailed mathematical expressions for the HGF. Details on the free parameters of the HGF model that were estimated, including prior values, can be found in Hein et al. (2021), and are briefly mentioned in **Figure 1A**.

Trial-by-trial trajectories of the unsigned predictions about the stimulus-reward tendency $|\hat{\mu}_2|$ and pwPEs about the stimulus outcome $|\varepsilon_2|$ for an exemplar participant are provided in **Figure 1B**.

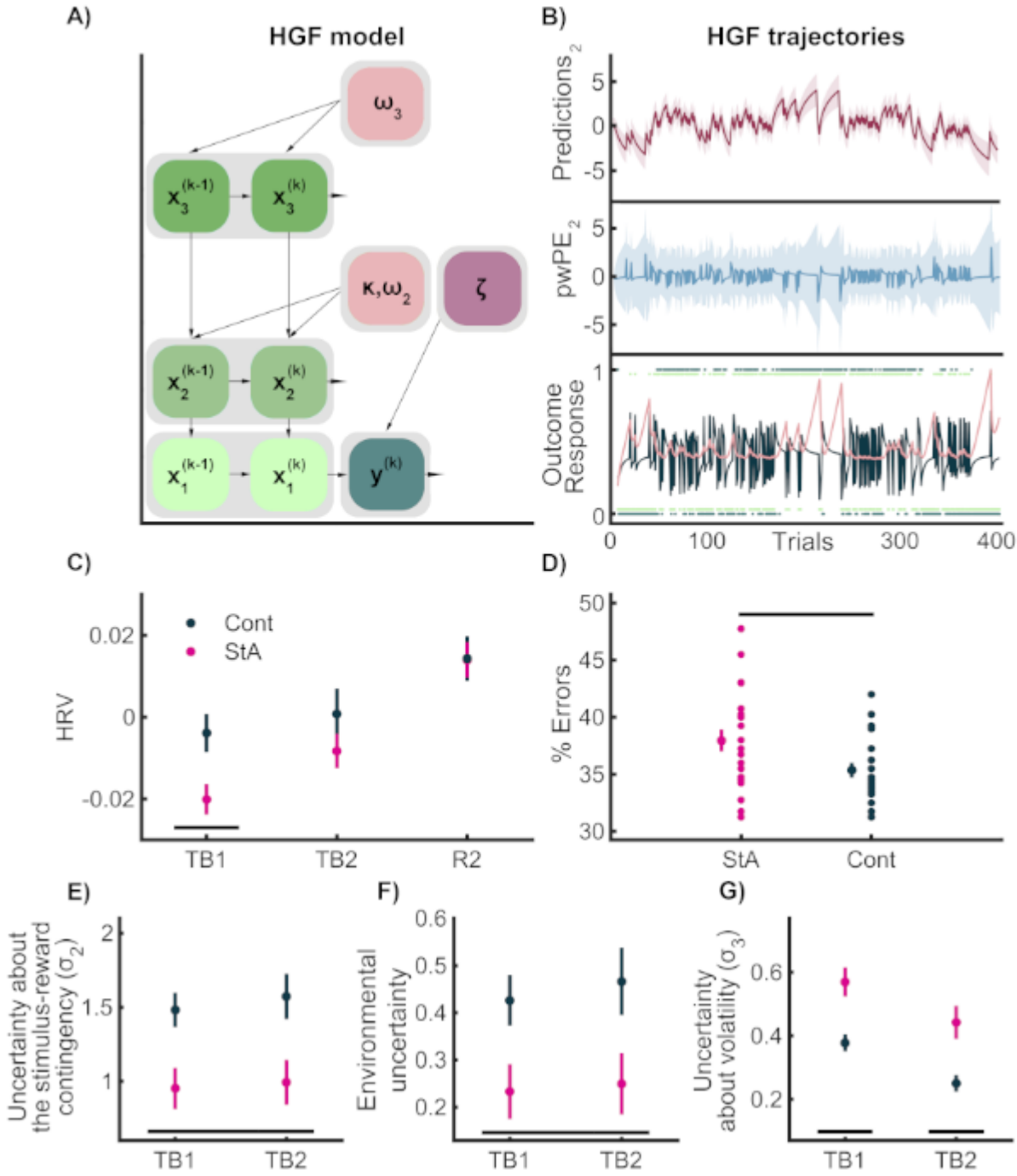


Figure 1. **HGF model, HGF trajectory estimates and HRV, HGF and behavioural results. A)** Schematic model of 3-level HGF used in Hein et al. (2021). The free perceptual model parameters ω_2, ω_3 and the response parameter ζ were estimated by fitting the HGF to observed inputs (u) and individual responses (y). **B)** HGF trajectories of the computational quantities used to form our GLM convolution regressors, from one participant. The lowest level shows the sequence of outcomes (green dots: 1 = blue win, 0 = orange win) and the participant's responses (dark blue dots) on each trial. The black line indicates the series of prediction errors (PE) about the stimulus outcome, and the pink line the precision weight on level 2. The middle layer of B) shows the trial-wise HGF estimate of pwPE about stimulus outcomes (pwPE updating level 2, termed $pwPE_2$ in the graphic, ε_2 the main text; blue). For our GLM convolution analysis, we used unsigned values of ε_2 as the first parametric regressor. The precision ratio included in the $pwPE_2$ term, in succession, weights the influence of prediction errors about stimulus

outcomes on the expectation of beliefs on level 2. Predictions about the tendency towards a stimulus-reward contingency on level 2 are displayed on the top level (maroon). We took the absolute values of this quantity as our second parametric regressor (labelled Predictions$_2$ in the graphic). **C)** In Hein et al. (2021) a significant drop in heart rate variability (HRV, a metric of anxiety using the coefficient of variation of the inter-beat-interval of the recorded heart beats), was observed in the StA group (pink) relative to Cont (black). Panel C) shows the mean HRV (with vertical SEM bars) over the experimental task blocks 1 and 2 (TB1, TB2) and the final resting state block (R2). These blocks (TB1, TB2, R1) were normalised to the average HRV value of the first resting state block (R1: baseline). A significant effect of group and block was discovered using non-parametric 2×2 factorial tests with synchronised rearrangements. After control of the FDR at level q = 0.05, planned comparisons showed a significant between groups result (black bar) in TB1. **D)** State anxiety impeded the overall reward-based learning performance as given by the percentage of errors. In the above, the mean of each group (StA, pink, Cont, black) is provided with SEM bars extending vertically. On the right of the group mean are the individual values depicting the population dispersion. State anxiety significantly increased the error rate relative to Controls. **E-G)** HGF modelling results. Hein et al. (2021) reported significantly lower ω_2 in StA relative to Cont. Simulations in that study showed that a lower ω_2 is associated with reduced estimation (informational) uncertainty on level 2, σ_2. **E)** In our StA group, the block average of estimation uncertainty about the stimulus-reward contingency (σ_2) was significantly smaller than in Cont (main effect of group; StA, pink; Cont, black). **F)** We observed significantly lower environmental uncertainty in StA relative to Cont (main effect of group). **G)** State anxiety increased uncertainty about volatility (σ_3, main effect of block and group). Planned between-group comparisons additionally revealed a significantly higher σ_3 in StA relative to Cont in each task block separately (TB1, TB2, black bars).

2.5 EEG and EKG acquisition and analysis

EEG, EKG and EOG signals were recorded continuously throughout the study using the BioSemi ActiveTwo system (64 electrodes, extended international 10–20, sampling rate 512 Hz). External electrodes were placed on the left and right earlobes to use as references upon importing the EEG data in the analysis software. EKG and EOG signals were recorded using bipolar configurations. For EOG, we used two external electrodes to acquire vertical and horizontal eye movements, one on top of the zygomatic bone by the right eye, and one between both eyes, on the glabella. For EKG we used two external electrodes in a two-lead configuration (Moody & Mark, 1982). Please refer to Hein et al. (2021) for further details on the electrophysiology acquisition.

EEG data were preprocessed in the EEGLAB toolbox (Delorme & Makeig, 2004). The continuous EEG data were first filtered using a high-pass filter at 0.5 Hz (with a hamming windowed sinc finite impulse response filter with order 3380) and notch-filtered at 48–52 Hz (filter order 846). Next, independent component analysis (ICA, runICA method) was implemented to remove artefacts related to eye blinks, saccades and heartbeats (2.3 components were removed on average [SEM 0.16]), as detailed in Hein et al. (2021). Continuous EEG data were then segmented into epochs centred around the outcome event

(win, lose, no response) from −200 to 1000 ms. Noisy data epochs defined as exceeding a threshold set to ± 100 μV were marked as artefactual (and were excluded during convolution modelling, see next section). Further to this, a stricter requirement was placed on the artefact rejection process to achieve higher quality time-frequency decomposition, as proposed for the gamma band (see Hassler et al., 2011; Keren et al., 2010). Data epochs exceeding an additional threshold set to the 75th percentile+1.5·IQR (the interquartile range, summed over all channels) were marked to be rejected (Carling, 2000; Schwertman et al., 2004; Tukey, 1977). The two rejection criteria resulted in an average of 22.37 (SEM 2.4) rejected events, with a participant minimum of 80% of the total 400 events available for convolution modelling. Following preprocessing, EEG continuous data were converted to SPM 12 (http://www.fil.ion.ucl.ac.uk/spm/ version 7487) downsampled to 256 Hz and time-frequency analysis was performed (Litvak et al., 2011).

Preprocessed EEG and behavioural data files are available in the Open Science Framework Data Repository: https://osf.io/b4qkp/. All subsequent results shown here are based on these data.

2.6 Spectral Analysis

Prior to assessing the effect of HGF predictors on “phasic” changes in the time-frequency representations, we determined whether the average spectral power differed between state anxiety and control participants during task performance. To achieve this, we extracted the standard power spectral density (in mV^2/Hz) of the raw data within 1–90 Hz and during task blocks TB1 and TB2 (fast Fourier transform, Welch method, Hanning window of 1 s, 75% overlap) and converted it into decibels (dB: $10*\log_{10}$).

Standard time-frequency (TF) representations of the continuous EEG data were estimated by convolving the time series with Morlet wavelets. TF spectral power was estimated in the range 4 to 80 Hz, using a higher number of wavelet cycles for higher frequencies. For alpha (8–12 Hz) and beta (13–30 Hz) frequency ranges, we sampled the range 8–30 Hz in bins of 2 Hz, using 5–cycle wavelets shifted every sampled point (Kilner et al., 2005)—achieving a good compromise between high temporal and spectral resolution (Litvak et al., 2011; Ruiz et al., 2009). Gamma band activity (31–80 Hz) was also sampled in steps of 2 Hz, using 7-cycle wavelets.

Following the time-frequency transformation, we modelled the time series using a linear convolution model for oscillatory responses (Litvak et al., 2013). This convolution model was

introduced to adapt the classical general linear model (GLM) approach of fMRI analysis to time-frequency data (Litvak et al., 2013). The main advantage of this approach is that it allows assessing the modulation of neural oscillatory responses on a trial-by-trial basis by one specific explanatory regressor while controlling for the effect of the other regressors included in the model. This control is particularly relevant in the case of stimuli or response events with variable timing on each trial. Convolution modelling of oscillatory responses has been successfully used in EEG (Litvak et al., 2013; Spitzer et al., 2016) and MEG research (Auksztulewicz et al., 2017).

In brief, the convolution GLM approach is an adaptation of the classical GLM, which aims to explain measured signals (BOLD for fMRI or time-domain EEG signals) across time as a linear combination of explanatory variables (regressors) and residual noise (Litvak et al., 2013). In convolution modelling for oscillatory responses, the measured signals are the time-frequency transformation (power or amplitude) of the continuous time series, denoted by matrix *Y* in the following expression:

$$Y = X \cdot \beta + \varepsilon,$$

Here $Y \in \mathbb{R}^{txf}$ is defined over *t* time bins and *f* frequencies. These signals are explained by a linear combination of *n* explanatory variables or regressors in matrix $X \in \mathbb{R}^{txn}$, modulated by the regression coefficients $\beta \in \mathbb{R}^{nxf}$. The coefficients β must be estimated for each regressor and frequency, using ordinary or weighted least squares.

The convolution modelling approach developed by Litvak et al. (2013) redefines this problem into the problem of finding time-frequency images R_i for a specific type of event *i* (e.g. outcome or response event type):

$$R_i = B\beta_i ,$$

Here, *B* denotes a family of *m* basis functions (e.g. sines, cosines) used to create the regressor variables *X* by convolving the basis functions *B* with *k* input functions *U* representing the events of interest at their onset latencies, and thus $X = UB$. The time-frequency response images $R_i \in \mathbb{R}^{pxf}$ have dimensions *p* (peri-event interval of interest) and *f,* and are therefore interpreted as deconvolved time-frequency responses to the event types and associated parametric regressors. It is the images R_i that are used for subsequent standard group-level statistical analysis. For a visual depiction of the convolution modelling of time-frequency responses, see **Figure 2**.

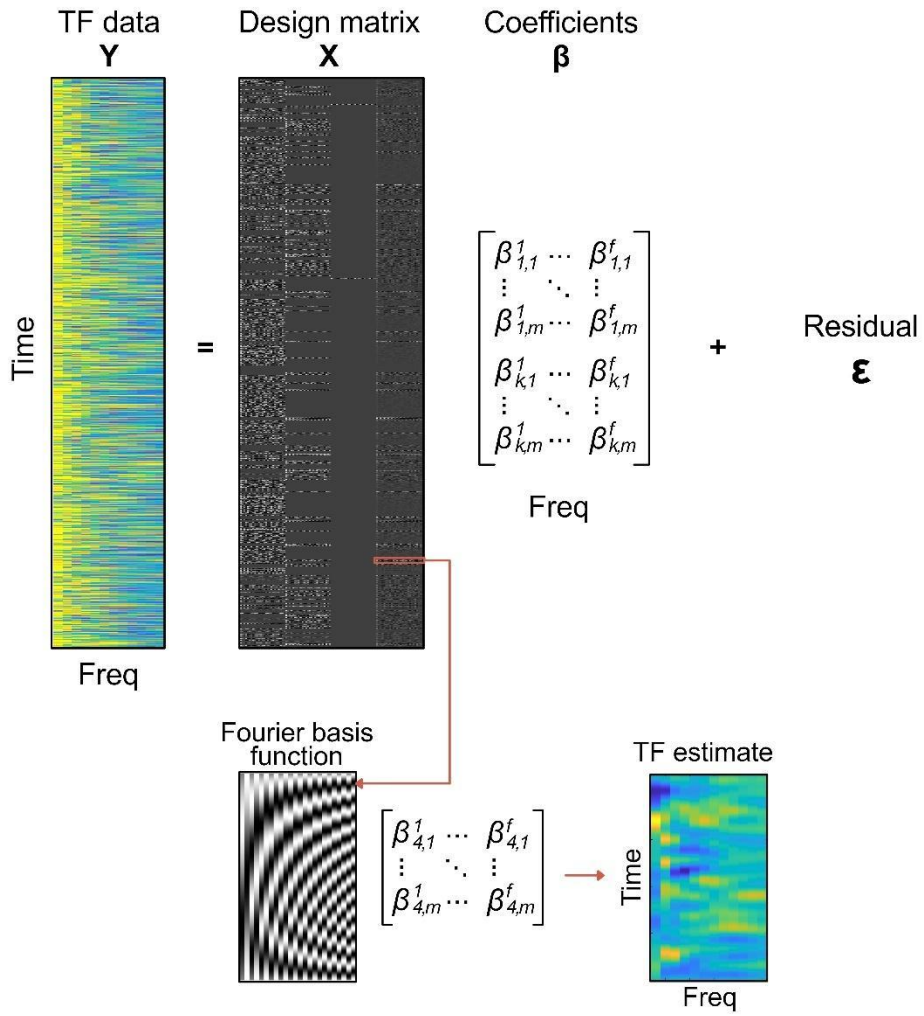


Figure 2. Convolution general linear model. Standard continuous time-frequency (TF) representations of the EEG signal (Y) were estimated using Morlet wavelets. In GLM, signals Y are explained by a linear combination of explanatory variables or regressors in matrix X, modulated by the regression coefficients β, and with an added noise term (ε). Our design matrix X in this example included the following regressors (columns left to right): Outcome Win, Outcome Lose, Outcome No Response, and absolute pwPE on level 2, which were defined over time. Matrix X was specified as the convolution of an impulse response function, encoding the presence and value of discrete or parametric events for each regressor and time bin, and a Fourier basis function (left inset at the bottom). Solving a convolution GLM provides response images (TF estimate in the figure) that are the combination of the basis functions and the regression coefficients β_i for a particular regressor type i. Thus, convolution GLM effectively estimates deconvolved time-frequency responses (TF estimate, rightmost image at the bottom) to the event types and associated parametric regressors.

In our study, we were particularly interested in assessing parametric effects of computational quantities, such as pwPEs and predictions, on the time-frequency representations of the EEG data in each electrode. We implemented convolution modelling by adapting code developed by Spitzer et al. (2016) freely available at https://github.com/bernspitz/convolution–models–MEEG. The total spectral power was first converted to amplitude using a square-root

transformation to conform with the GLM error assumptions (Kiebel et al., 2005; Litvak et al., 2013). Our trial-wise explanatory variables included discrete regressors coding for stimuli (blue image, orange image), responses (right, left, no response), outcome (win, lose, no response) and relevant parametric HGF regressors: unsigned HGF model estimates of predictions about the tendency towards a stimulus-reward contingency on level 2 ($|\hat{\mu}_2|$, hereinafter termed 'predictions') and precision-weighted prediction errors (pwPEs) on that level encoding the magnitude of the update in the beliefs about the reward contingency ($|\varepsilon_2|$, hereinafter termed 'pwPEs'; see **Figure 1B**). We selected the absolute value of predictions and pwPEs on level 2 because the sign in these HGF variables is arbitrary: a positive or negative value in pwPEs or predictions does not denote a win or a lose trial (see other HGF work using unsigned HGF variables as regressors, for instance, Auksztulewicz et al., 2017; Stefanics et al., 2018). The absolute values of predictions do, however, represent a prediction about the tendency towards a particular stimulus-reward contingency, and thus the greater the value of $|\hat{\mu}_2|$ the stronger the expectation that given the correct stimulus choice a reward will be received.

As in our previous work, pwPE on level 3 (ε_3) updating the log-volatility estimates were excluded from this analysis due to multicollinearity: high linear correlation between pwPEs and ε_3 (for further detail, see Hein et al., 2021). Likewise, trial-wise HGF estimates of predictions about stimulus outcomes were highly linearly correlated with predictions on the third level about volatility $\hat{\mu}_3$ (Pearson correlation coefficients ranging from −0.97 to −0.03 across all 42 participants, mean −0.7). As such, we also excluded $\hat{\mu}_3$ from the analysis. (For details on the impact of multicollinearity of regressors on GLMs see Mumford et al. [2015] and Vanhove [2020]). Another factor informing our decision to choose level 2 over level 3 regressors was that, as shown in simulations in Hein et al. (2021), in the winning model ω_2 can be estimated well, whereas ω_3 is not (see also Reed et al., 2020). The chosen HGF pwPE and prediction regressors were consistently uncorrelated, below 0.25 in line with previous work using HGF quantities as regressors (Auksztulewicz et al., 2017; Iglesias et al., 2013; Vossel et al., 2015). Our primary convolution GLM analysis introduced regressor values for pwPEs at the latency of the outcome regressor. This allowed us to assess the parametric effect of pwPEs about stimulus outcomes on the time-frequency responses in a relevant peri-event time interval. Although previous work analysed the effect of pwPEs on neural responses up to 1000 ms, we showed in Sporn et al. (2020) that pwPEs during reward-based learning can modulate neural oscillatory responses in the beta band up to 1600 ms, and these responses are dissociated between anxiety and control groups. The recent studies by Bauer et al. (2014) and Palmer et al. (2019) also showed that the latency of PE and pwPE effects on neural activity can extend up to 2 seconds. Accordingly, the pwPE convolution model was estimated using a window

from −200 to 2000 ms relative to the outcome event, and the statistical analysis focused on the 100-1600 ms interval (see next section).

Concerning the prediction regressor, we considered different time intervals in which we could capture neural oscillatory responses to predictions. This is a challenging task acknowledged before (Diaconescu et al., 2017), as the neural representation of predictions likely evolves gradually from the outcome on the previous trial to the outcome on the current trial. It is thus not expected to be locked to a specific event. This explains why most of the previous work using the HGF framework excluded predictions as a regressor for GLM analysis. Here we followed Auksztulewicz et al. (2017), who analysed predictions locked to the cue, and Palmer et al. (2019), who assessed a wide interval surrounding the movement (response); note that in the Palmer et al. (2019) study, the motor response was the last event in each trial (i.e. there was no additional response feedback). We thus hypothesised that the neural representation of predictions on the reward outcome contingencies could be captured by focusing on two complementary windows of analysis: (i) an interval following the stimulus presentation (stimulus-locked); (ii) an interval preceding the outcome on the current trial (outcome-locked). Unlike the targeted pwPE analysis described above, the analysis of the prediction regressor was exploratory as we did not have a strong hypothesis regarding which of both time windows would preferentially reflect prediction-related neural modulations.

To assess the stimulus-locked parametric effect of predictions on the time-frequency responses, we run a convolution GLM in a time interval from −200 to 2000 ms. For the outcome-locked parametric effect of predictions, the convolution GLM was run from −2500 to 0 ms. This later interval extended to −2500 to allow for the presence of a baseline interval in every trial prior to the preceding stimulus—which we used exclusively for within-subject analyses (see below). Thus, two separate convolution GLMs were run with the prediction regressor modulating neural activity locked to either the stimulus or outcome events. These broad windows were further refined in our statistical analysis (see next section).

In all alpha-beta convolution GLM analyses, discrete and parametric regressors were convolved with a 12th-order Fourier basis set (24 basis functions, 12 sines and 12 cosines), as in Litvak et al. (2013). For convolution models run from −200 to 2000 ms locked to an event type, using a 12th-order basis functions set allowed the GLM to resolve modulations in the TF responses up to ~ 5.5 Hz (12 cycles / 2.2 seconds; or 183 ms). For the outcome-locked GLM run from −2500 to 0 ms, the 12th-order Fourier basis set resolves frequencies up to ~5 Hz. Our choice of a 12th order set was compatible with the temporal extent of the pwPE and prediction effects on alpha-beta oscillatory activity reported in previous work (200–400 ms-

long effects in Auksztulewicz et al., 2017) up to 2000 ms-long effects in Palmer et al. (2019). In the case of gamma oscillations modulating pwPEs, we considered a higher order basis function set to allow for potentially faster gamma effects to be resolved. Using a 20th-order Fourier basis set on the gamma-band convolution GLM within −200 to 2000 ms enabled resolving modulations in the TF responses up to ~ 9 Hz (20 cycles / 2.2 seconds; or 110 ms).

2.7 Statistical analysis

The time-frequency images (in arbitrary units, a.u.) from the convolution model were subsequently converted to data structures compatible with the FieldTrip Toolbox for statistical analysis (Oostenveld et al., 2011). We used permutation tests with a cluster-based threshold correction to control the family-wise error rate (FWER) at level 0.05 (5000 iterations; Maris & Oostenveld, 2007; Oostenveld et al., 2011). These analyses were conducted with spatio-spectral-temporal data, after averaging the time-frequency responses within each frequency band (alpha, beta and gamma ranges). We thus run the cluster-based permutation tests along the spatial (64 channels), frequency-band (3) and temporal dimensions (FWER-controlled). Importantly, in convolution modelling for oscillatory responses the TF images are usually not baseline corrected as in standard TF analyses (no subtraction or division by the average baseline level). Instead, the baseline activity is estimated—similarly to the post-event activity—taking into account the latency variation of different events in the continuous recording (Litvak et al., 2013). Thus, TF images are not centred at 0 amplitude during the baseline period.

The statistics approach consisted of investigating separately within and between-group effects. The within-group level analysis used dependent-samples two-sided tests and aimed to assess whether the neural oscillatory responses to the HGF regressors were larger or smaller during a window of interest as compared to a reference (baseline) interval. Next, we separately evaluated between-group effects of HGF regressors on oscillatory responses using one-sided tests (N = 21 Cont, 21 StA). This allowed us to test our hypothesis of increased alpha and beta activity and reduced gamma activity in StA compared to Cont. In the case of two-sided tests, the cluster-based test statistic used as threshold the 2.5-th and the 97.5-th quantiles of the t-distribution, whereas we used the 95th quantile of the permutation distribution as critical value in one-sided tests.

Analysis of the pwPE regressor

At the within-group level, we assessed the changes in time-frequency activity during the window of interest relative to a baseline period (given independently below) separately in StA

and Cont groups (N = 21 each). For the within-group analysis of the pwPE regressor, we contrasted the time-frequency images between an interval from 100 to 1600 ms post-outcome and a baseline level averaged from −200 to 0 ms, separately in each group. The 100–1600 ms time window of analysis encompasses the effects from our previous single-trial ERP study (Hein et al., 2021) and our work on the modulation of beta oscillatory responses by pwPEs during motor learning in state anxiety, which revealed effects between 400–1600 ms (Sporn et al., 2020). Between-group differences in TF representations of pwPE were separately assessed. This analysis was also conducted within 100–1600 ms. Overall, we controlled the FWER at level 0.05 to deal with the issue of multiple comparisons emerging from the spatial (64) x spectral (3) x temporal dimensions.

Analysis of the predictions regressor

Within-group level statistical analysis of the stimulus-locked time-frequency images of the prediction regressor focused on the range 100–1000 ms, and relative to an average pre-stimulus baseline level from −200 to 0 ms. This target window for statistical analysis balanced the evidence from previous work (Auksztulewicz et al., 2017; Palmer et al., 2019) and aimed to exclude an overlap with the outcome events, which appeared 1000 ms after the response. In the current study, participants' reaction time was 598 ms on average (SEM 130 ms; minimum RT was ~300 ms). However, the effects of the response were factored out from the prediction-related oscillatory activity by including the response regressor in the convolution GLM. This was validated in a control analysis that assessed the effect of the response regressor in the same time window, between 100–1000 ms stimulus-locked, to confirm independent changes in sensorimotor regions.

Within-group statistical analysis of the outcome-locked effects of predictions was conducted in a similar window 100–1000 ms preceding the outcome event (that is, from −1000 to −100 ms before the outcome). Activation in this interval was contrasted to a baseline level of 200 ms, from −2300 to −2100 ms. This baseline period was calculated to safely precede stimuli presentation across all trials, during which participants were fixating on a central point on the monitor. As mentioned above, to confirm independent changes in sensorimotor regions in response to the response regressor, we used an identical window in an additional control analysis.

The between-group level stimulus-locked analysis of predictions was conducted within 100–1000 ms. The outcome-locked analysis targeted the interval from −1000 to −100 ms, as mentioned above. In all GLM analyses, the FWER was controlled at level 0.05.

3. Results

3.1 Previous results: Biases of state anxiety on processing uncertainty

In Hein et al. (2021) we showed state anxiety (StA) significantly reduced HRV and HF-HRV (0.15–0.40 Hz) relative to the control group (Cont, **Figure 1C**). This outcome suggested that our state anxiety manipulation had successfully modulated physiological responses in a manner consistent with changes in state anxiety (Friedman, 2007; Fuller, 1992; Klein et al., 1995; Miu et al., 2009; Pittig et al., 2013). We further showed that state anxiety significantly increased the percentage of errors made during reward-based learning when compared to the control group (**Figure 1D**). In parallel to the cardiovascular and behavioural changes induced by the anxiety manipulation, by modelling decisions with the HGF, we found that state anxiety impaired learning. First, we found significantly reduced *estimation* uncertainty (σ_2) in StA relative to Cont (**Figure 1E**). This bias in StA indicates that new information has a smaller impact on the update of beliefs about the tendency towards a stimulus-reward contingency (level 2). State anxious individuals also exhibited an underestimation of environmental uncertainty when compared with controls (**Figure 1F**). However, uncertainty on volatility (σ_3) increased in StA relative to Cont (**Figure 1G**). StA also had a lower ω_2 parameter than control participants, which in Hein et al (2021) was associated in a simulation analysis with the reduced estimation uncertainty in this group. Other model parameters (ω_3, ζ) did not differ between groups. These HGF model-based results were aligned with the results of our separate standard behavioural analysis as mentioned above, demonstrating a significantly higher error rate in StA during reward-based learning performance (**Figure 1D**).

3.2 Time-frequency responses

3.2.1 General modulation of spectral power

The average raw spectral power during task performance did not differ between state anxiety and control participants ($P > 0.05$, cluster-based permutation test; **Supplementary Figure 1**). Thus, the state anxiety manipulation did not significantly modulate the general spectral profile of oscillatory activity during task performance, as we showed in a recent study (Sporn et al., 2020).

3.2.2 Precision-weighted prediction errors about stimulus outcomes

The overall time course of the parametric modulation of alpha (8–12 Hz) and beta (13-30 Hz) oscillatory activity by pwPEs about stimulus outcomes is displayed in **Figure 3A** and **Figure 4A**, respectively. On the within-subject level, there was a significant decrease relative to baseline in alpha and beta activity in the control group (one negative cluster, $P = 0.0002$, two-sided test, FWER-controlled. Effect within 600-1400 ms for alpha, and 400–1120 ms for beta activity). No significant clusters were found in the gamma band. The alpha-band effect originated in centro-parietal electrodes and later spread across the whole scalp (**Figure 3B**). The beta-band modulation, on the other hand, had a widespread topography and started earlier than the alpha-band effect (at 400 ms; **Figure 4B**). In the StA group, a negative cluster was also found, corresponding to a decrease from baseline in alpha and beta activity ($P = 0.0054$, two-sided test; 600–1000 ms for alpha, 440–1000 ms for beta). The StA alpha-band effect also emerged in centro-parietal electrodes but later shifted to frontocentral electrodes (**Figure 3C**). In the beta range, the negative modulation of oscillatory activity in StA had a right frontocentral and left centro-parietal distribution (**Figure 4C**).

Complementing the within-subject results, between-group statistical analysis across the alpha, beta and gamma ranges revealed one significant positive cluster in the beta range (between 1200–1570 ms, $P = 0.027$, one-sided test; FWER-controlled). This effect was associated with higher beta activity at left sensorimotor and frontocentral electrodes in StA relative to Cont (**Figure 5AB)**. The individual average of beta-band activity in the significant cluster is shown in **Figure 5C.** Of note, in StA, a *qualitative* comparison of the sensorimotor and frontocentral beta activity associated with the significant cluster of the between-group statistical analysis revealed a greater activity increase in the sensorimotor than in the frontocentral electrode region (**Figure 5D**). In the control group, the beta response to pwPE decreased in both electrode regions, but the reduction was more pronounced in frontocentral electrodes (**Figure 5E**). There were no additional significant clusters associated with between-group differences in the alpha or gamma ranges (see illustration of gamma responses to pwPE in **Supplementary Figure 2**).

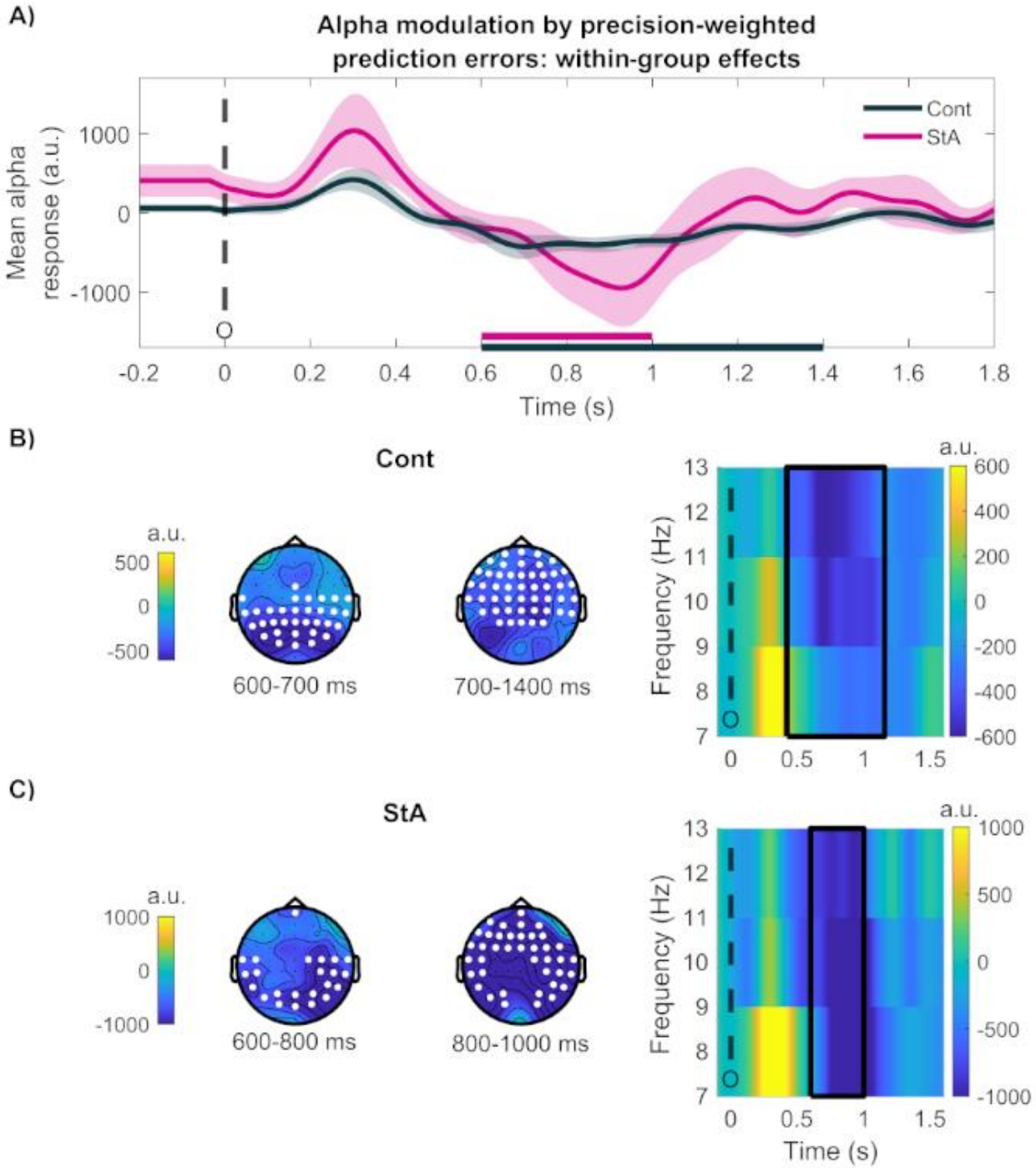


Figure 3 Alpha activity is modulated by precision-weighted prediction errors about stimulus outcomes: within-group effects. A) Time course of the average alpha response (8–12 Hz) to pwPEs in each group (Controls, black; StA, pink), given in arbitrary units (a.u). The time intervals correspond to the dependent-samples significant clusters (B, C) in each group, and are denoted by horizontal bars on the x-axis. **B)** Within-group effect of the $pwPE_2$ regressor modulating alpha oscillations relative to baseline in the Cont group (one negative cluster within 600–1400 ms, $P = 0.0002$). Left: The topographic distribution of this effect starts in posterior centroparietal regions and expands across widespread frontal and central regions. Right: Time-frequency images for pwPE on level 2, averaged across the cluster electrodes. The black dashed line marks the onset of the outcome, and black squares indicate the time-frequency range of the significant cluster. **C)** Same as (B) but in the StA group. We found a significant negative cluster also in the alpha and beta-band ranges. The alpha-band effect was found between 600–1000 ms ($P = 0.0054$), starting in posterior central electrodes and spreading to frontocentral

electrodes later. Dashed and continuous black lines denote outcome onset and the extension of the significant cluster in the time-frequency range, as in (B).

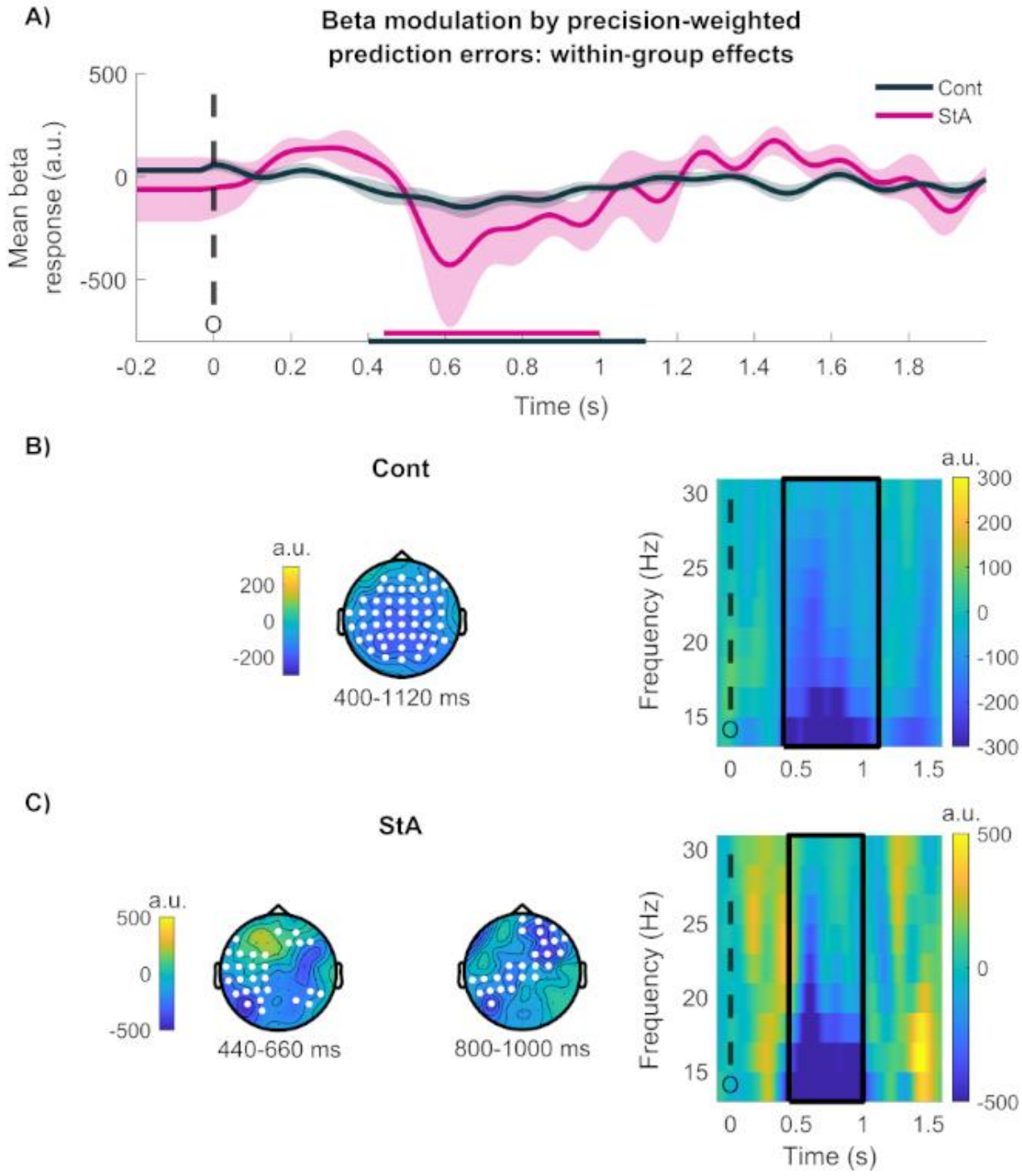


Figure 4 Beta activity is modulated by precision-weighted prediction errors about stimulus outcomes: within-group effects. A) Time course of the average beta response (12–30 Hz) to pwPEs in each group (Controls, black; StA, pink), given in arbitrary units [a.u.]). The time intervals corresponding to the dependent-samples significant clusters (B, C) in each group are denoted by horizontal bars on the x-axis. **B)** In Cont participants, beta-band oscillations were significantly modulated relative to a baseline level in one negative cluster spanning 400–1120 ms (P = 0.0002, FWER-controlled due to multiple comparisons arising from testing across space x frequency-band x

time dimensions) Left: The topographic distribution of the beta-band effect is widespread across the entire scalp. Right: Time-frequency images for pwPE on level 2, averaged across the significant cluster electrodes. The black dashed line marks the onset of the outcome, and black squares indicate the time-frequency range of the significant cluster. **C)** Same as (B) but in the StA group. As reported in Figure 3, we found a significant negative cluster across the alpha and beta band, with a latency of 440–1000 ms for beta activity (P = 0.0054, FWER-controlled). The beta modulation started in posterior central electrodes and later spread to frontocentral electrodes. Dashed and continuous black lines denote outcome onset and the extension of the significant cluster in the time-frequency range, as in (B).

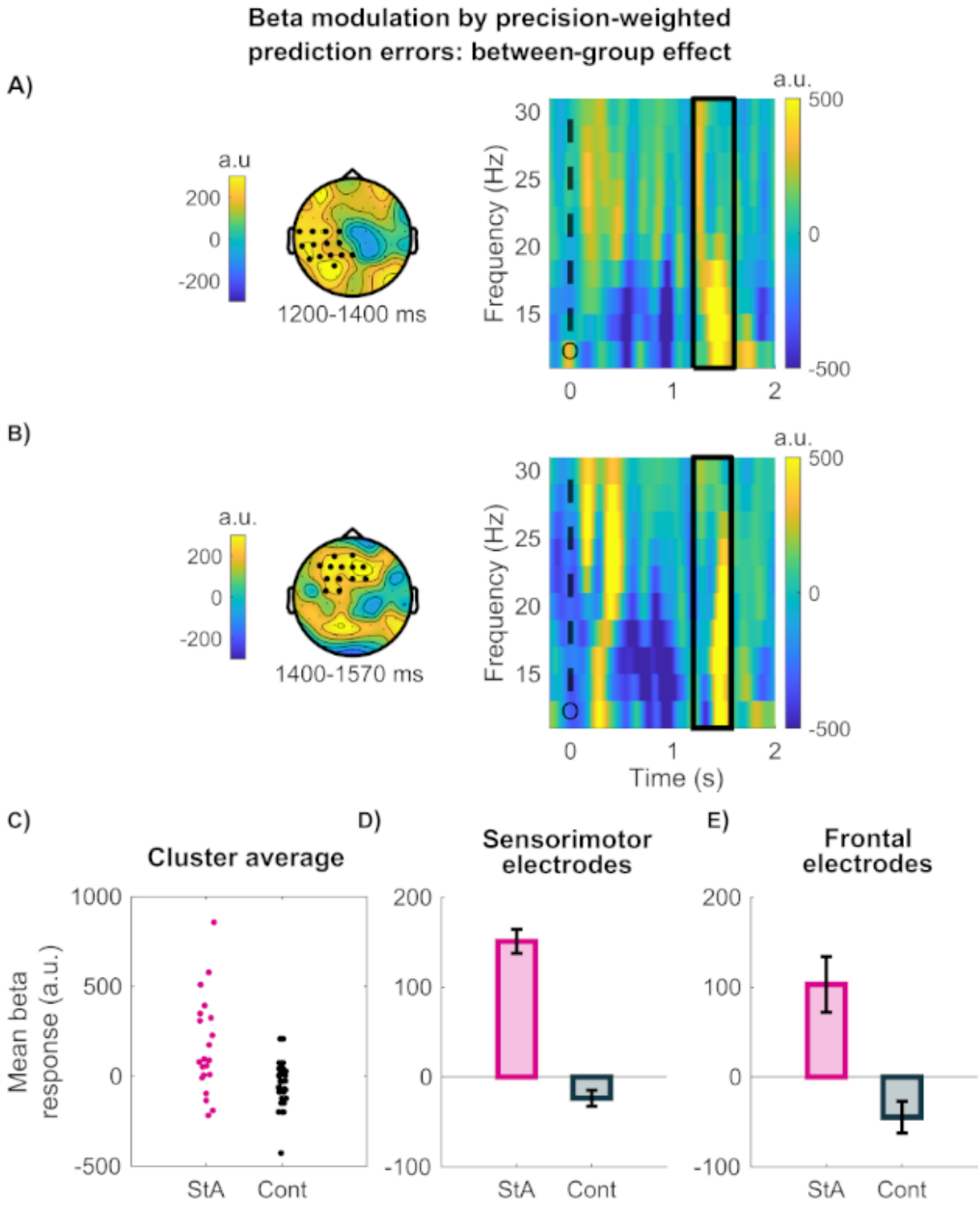


Figure 5. Between-group effects of pwPEs on beta oscillatory activity. A-B) Between-group differences in the oscillatory activity (alpha 8–12 Hz, beta 13–30 Hz, and gamma 31–90 Hz) modulations by pwPEs about stimulus outcomes were found exclusively in the beta band (one significant positive cluster between 1200–1570 ms; *P* = 0.0270, FWER-controlled). **A)** The topography of this effect evolved in time from a left sensorimotor distribution to a **B)** frontocentral electrode distribution. The time-frequency images on the right panels are averaged across the electrode selection on the left topographic panels and are given in arbitrary units (a. u.). Note that there was one single significant cluster between 1200–1570 ms, represented by the solid black rectangle in each TFR plot.

The dashed black line 'O' represents the time of the outcome. **C)** The average beta response (a. u.) to pwPE for individuals in each group (StA, pink; Cont, Black) in the significant positive cluster. The modulation of time-frequency responses to pwPEs in beta is displayed separately in **D)** sensorimotor and **E)** frontal electrodes pertaining to the significant positive cluster. Pink bars represent results in the state anxiety group (StA), whereas the control group (Cont) is denoted by black bars. Black "error" bars indicate the standard error of the mean (SEM).

Because the within-subject and between-group modulation of TF images by the pwPE regressor were limited to the alpha and beta frequency ranges, we performed an additional control analysis to determine the separate effect of the precision weight (σ_2) and PE (abs[δ_1]) regressors. Of note, the absolute value of PEs (abs[δ_1]) is often termed surprise (see e.g. de Berker et al., 2016). Like for pwPE about stimulus outcomes, the sign in δ_1 is not informative and thus a sensible choice is to use the unsigned values (de Berker et al., 2016; Auksztulewicz et al., 2017; Stefanics et al., 2018). This control analysis could determine whether the alpha and beta pwPE effects primarily stem from precision weights modulating lower frequency activity, or rather from a modulation by the surprise experienced by the participants. Moreover, similarly to PEs, surprise about inputs has been shown to correlate with gamma oscillations (Bauer et al., 2014). Thus, the analysis of the abs[δ_1] regressor could identify gamma modulation effects that may not be observable in the pwPE analysis. This convolution GLM model included both continuous regressors σ_2 and abs(δ_1) as well as the discrete regressors coding for outcomes.

At the between-subject level, we observed a significant increase in beta-band oscillatory responses to surprise about stimulus outcome in the StA group relative to Cont (one positive cluster within 1380–1600 ms, P = 0.01, one-sided test, FWER-controlled). This effect was distributed across frontocentral and left sensorimotor electrodes (**Figure 6**), similarly to the pwPE effects. There was no significant difference between groups in alpha or gamma-band modulation by surprise. Within-subject effects also demonstrated that the absolute PE regressor alone modulated alpha and beta oscillatory activity in each group separately (see details in **Supplementary Results** and **Supplementary Figures 3** and **4**).

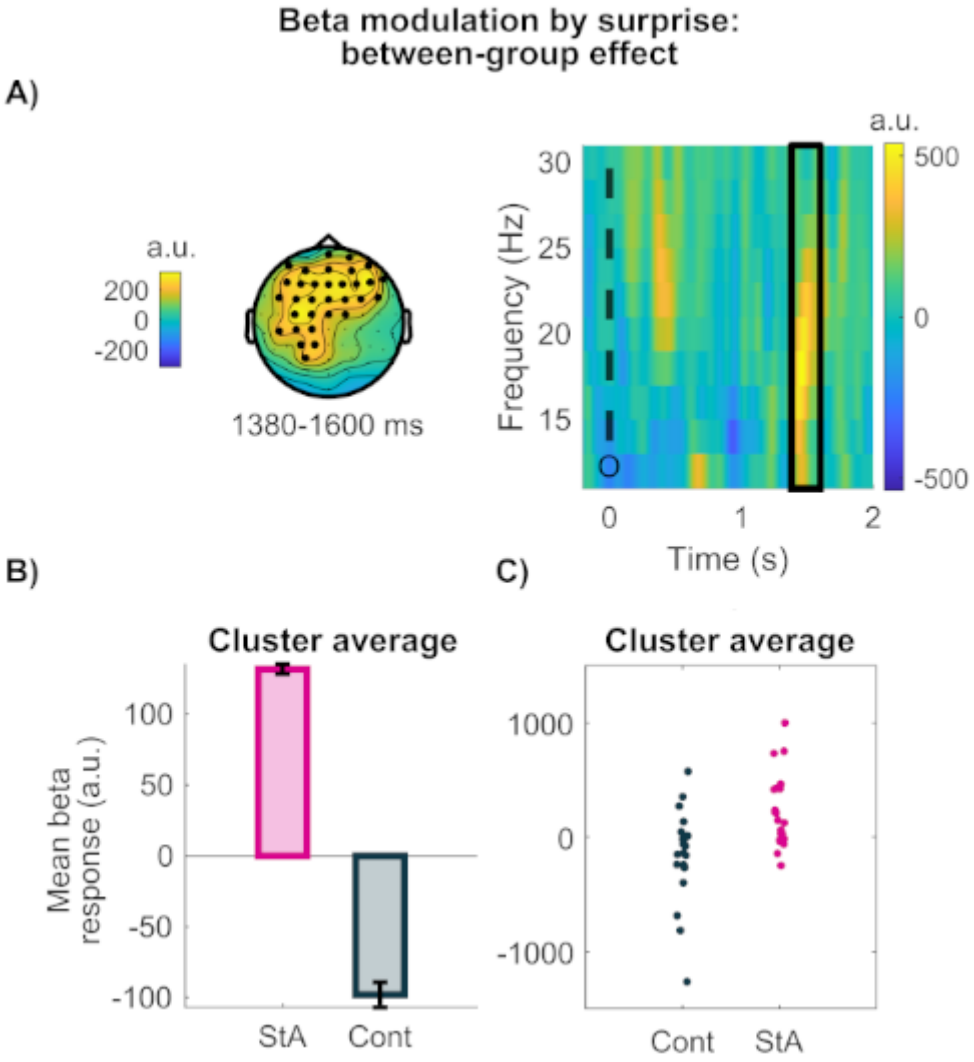


Figure 6. Between-group effects of surprise (absolute PEs) on beta oscillatory activity. The time-frequency images representing modulation by the absolute value of PE about stimulus outcomes (abs[δ_1]) were estimated in a control convolution GLM using continuous (abs[δ_1], σ_2) and discrete regressors coding for outcome events (win, lose, no response). **A-B)** Between-group differences in beta oscillatory activity (13–30 Hz) modulations by surprise about stimulus outcomes (one significant positive cluster between 1380–1600 ms; *P* = 0.01, FWER-controlled). **A)** This effect was topographically distributed across frontocentral and left sensorimotor electrodes. The right panel shows the average TF image in the electrodes pertaining to the significant cluster, shown in the left topographic panel. TF images are presented in arbitrary units (a. u.). The solid black rectangle denotes the range spanned by the significant cluster; the dashed black line 'O' represents the time of the outcome. **B)** The average beta response (a. u.) to surprise in each group (StA, pink; Cont, Black) within the frontocentral positive cluster. Pink bars represent results in the state anxiety group (StA), whereas the control group (Cont) is denoted by black bars. Black "error" bars indicate the standard error of the mean (SEM) **C)** The average beta response to surprise for individuals in each group (StA, pink; Cont, Black) in the frontocentral positive cluster.

Using the precision weights term (σ_2) as a regressor, the comparison between groups demonstrated a positive significant cluster exclusively in the alpha frequency range (within

1200–1600 ms, *P* = 0.01, FWER-controlled). The positive cluster was associated with higher alpha activity primarily at central electrodes but also at fronto-central and temporal electrodes in StA relative to Cont (**Figure 7**). At the within-subject level we only observed that in Cont participants there was a negative change in alpha activity to the precision weight regressor (one negative cluster within 1270–1530 ms; *P* = 0.024 FWER-controlled, see **Supplementary Figure 5**).

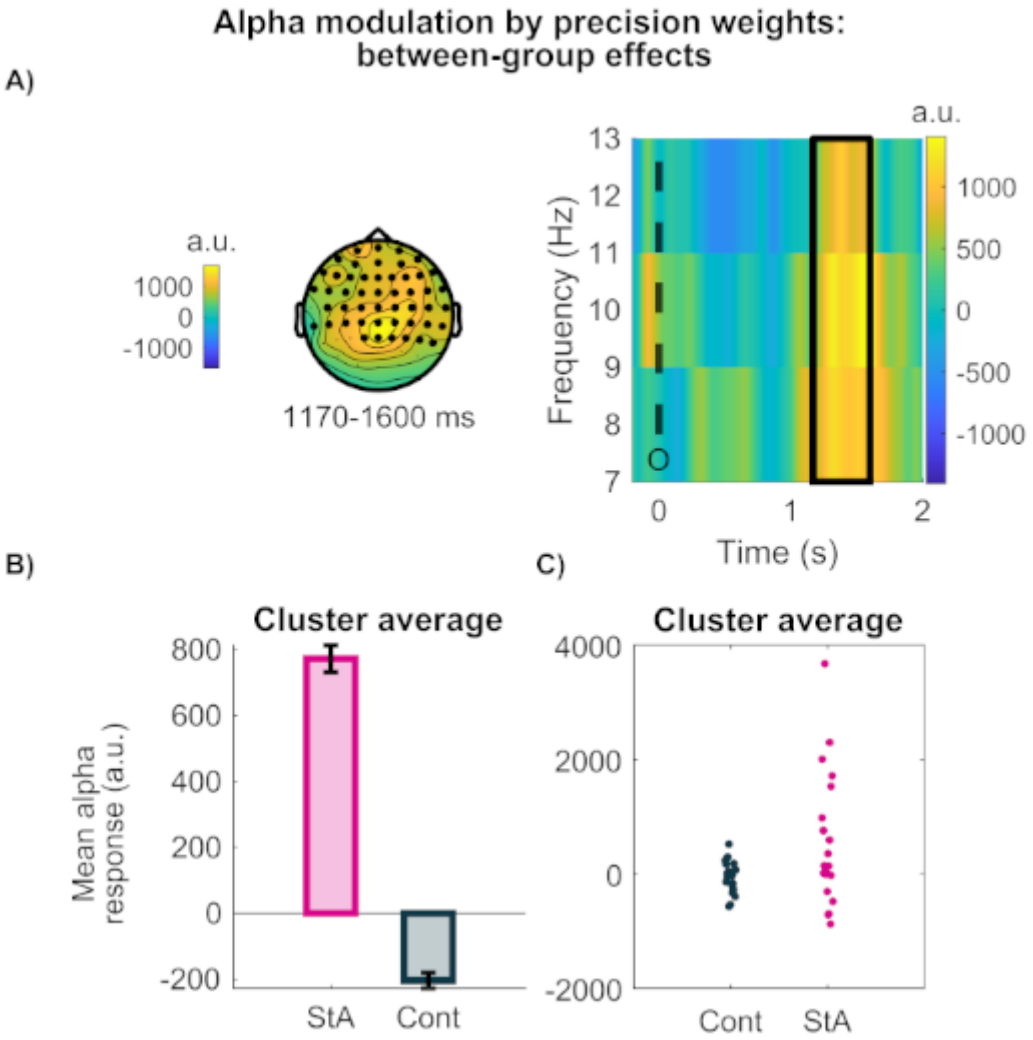


Figure 7. Between-group effects of precision weights on beta oscillatory activity. A-B) Between-group differences in alpha oscillatory responses (8–12 Hz) modulated by precision weights (σ_2; one significant positive cluster between 1170–1600 ms; *P* = 0.01, FWER-controlled). **A)** In the topographic map this effect can be seen over central, fronto-central and temporal electrodes. The right panel displays the time-frequency responses in the significant cluster, after averaging the TF images across the central/fronto-central electrodes in the cluster. The TF image is given in arbitrary units (a. u.). The solid black rectangle shows the significant cluster, while the dashed black line 'O' represents the time of the outcome. **B)** The average beta response (a. u.) to precision weights in each group (StA, pink; Cont, Black) within the central/fronto-central positive effect. State anxiety group (StA, pink); Control group (Cont, black). Black "error" bars indicate the standard error of the mean (SEM) **C)** The average

beta response to precision weights for individuals in each group (StA, pink; Cont, Black) in the central/fronto-central positive cluster.

Lastly, to exclude the possibility that our high-pass filter settings (0.5 Hz) explained the lack of significant modulation effects in the gamma band, we reanalysed the data in four representative participants after applying a 0.1 Hz high-pass filter during pre-processing. This analysis was motivated by studies showing that higher cutoff frequencies for high-pass filters can impact the signal-to-noise ratio (SNR) in general and gamma activity in particular (Bénar et al., 2010; Jas et al., 2018). In brief, using a 0.1 Hz cutoff as opposed to our choice of 0.5 Hz for high-pass filtering did not reveal any prominent gamma modulation by pwPE or surprise/PE regressors (**Supplementary Figures 6–7**), and did not substantially affect the general SNR level in the power spectral density (**Supplementary Figures 8)**.

3.2.3 Predictions about the stimulus-reward contingency

Stimulus-locked

When assessing within-group level modulations in oscillatory activity by the prediction regressor, there were no significant effects, neither in the Cont or StA group ($P > 0.05$, FWER-controlled). Between-group statistical analysis revealed that predictions about the tendency towards a certain stimulus-reward contingency are associated with significantly higher levels of beta activity in StA than in Cont across frontocentral and parietal electrodes (one positive cluster in the beta band only, from 200 to 640 ms, $P = 0.04$, one-sided test, FWER-controlled; **Figure 8AB**). There were no additional significant clusters extending to the alpha range.
The between-group effect of predictions on beta activity was not confounded by any concomitant effect of motor responses on the neural oscillatory responses, as we had included a response regressor in this analysis. A control analysis on this between-group effect of the response regressor on beta activity showed no significant difference between the two groups (see **Supplementary Figure 9A**).

Outcome-locked

Figure 8C displays the time course of the parametric effects of predictions on outcome-locked beta activity. State anxious participants exhibited a significant increase from baseline in beta oscillatory activity (one significant positive cluster from −1000 to −468 ms, $P = 0.0106$, two-sided test, FWER-controlled). This effect peaked at central parietal and left frontocentral

electrodes (see **Figure 8D**). There were no significant changes from baseline in alpha or beta oscillatory activity for the control group participants. Neither did we find significant between-group differences in outcome-locked alpha or beta activity. Like in our stimulus-locked results, the significant outcome-locked increase from baseline in beta oscillatory activity in the StA group was not confounded by motor modulation, as this was included as a separate regressor in the convolution model. A control analysis of the effect of the response regressor on beta activity yielded non-significant changes from baseline in StA (**Supplementary Figure 9B**).

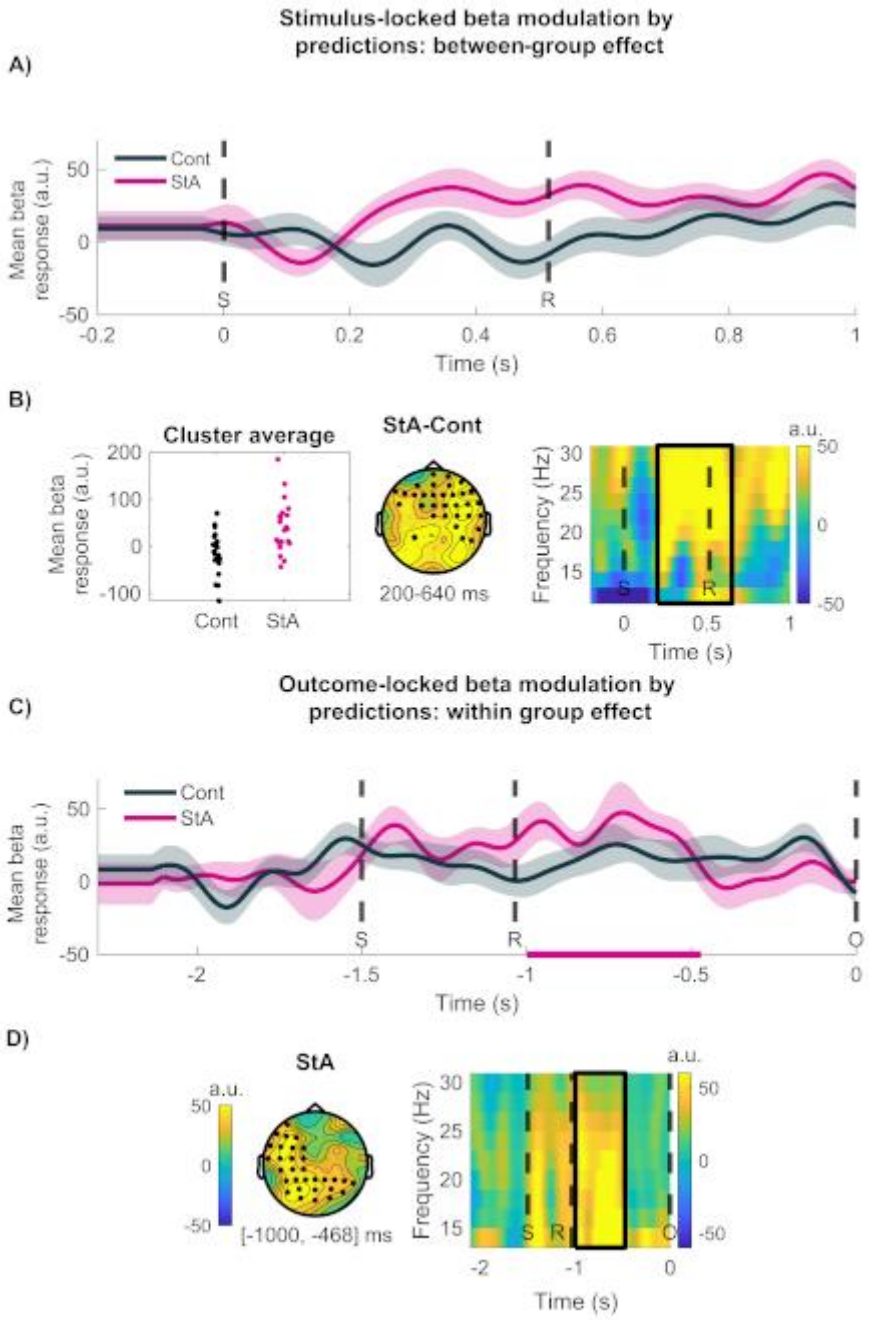


Figure 8. Stimulus-locked and outcome-locked modulation of beta activity by predictions about the reward tendency. A) Average time course of stimulus-locked beta activity modulated by predictions about the tendency towards a stimulus-reward contingency in Cont (black) and StA (pink). Modulation of time-frequency images by a regressor is dimensionless, and thus given in arbitrary units (a.u). **B)** In the leftmost column is the average beta activity (a. u.) in the cluster to pwPE for individuals in each group (StA, pink; Cont, Black). An independent-samples test on beta activity revealed a significant increase in StA relative to Cont from 200 to 640 ms across parietal and frontocentral electrodes (one significant positive cluster, $P = 0.04$, FWER-controlled). **C)** Outcome-locked modulation of beta activity by predictions about the reward tendency. Average time course in a.u. of outcome-locked beta activity reflecting modulation by predictions in Cont (black) and StA (pink). The significant within-group effect in StA is shown by the pink horizontal bar on the x-axis. **D)** Within-group statistical analysis with dependent-samples cluster-based permutation tests revealed one positive cluster in the state anxious group ([−1000, −468] ms, $P = 0.0106$, two-sided test, FWER-controlled), reflecting increased beta activity in centroparietal and left frontal electrodes during processing predictions. Solid black lines

represent the time and frequency of the significant cluster. Dashed black lines represent the average time of the stimuli presentation ‘S’, participant’s response ‘R’ and the outcome ‘O’.

4. Discussion

This study investigated how anxiety states modulate the oscillatory correlates of predictions and prediction errors during the learning of stimulus-reward associations in a volatile environment. The analysis focused on low-level predictions about the tendency of stimulus-outcome contingencies and prediction errors about stimulus outcomes. Because in generative models of the external world precision weights regulate the influence that PEs have on updating predictions (Feldman & Friston, 2010; Friston, 2010), we assessed the neural oscillatory responses to precision-weighted PEs (pwPEs), similarly to Auksztulewicz et al. (2017). We tested this by re-analysing data from our previous study, which investigated Bayesian predictive coding (PC) in state anxiety (Hein et al., 2021). That study showed that anxious individuals overestimate how precise their belief about the stimulus-reward contingency is, attenuating pwPEs on that level and decreasing learning. In the current study, trial-wise model estimates of predictions and pwPEs were used as parametric regressors in a convolution model to explain modulations in the amplitude of oscillatory EEG activity (Litvak et al., 2013).

Consistent with our hypotheses, we found that state anxiety alters the spectral correlates of pwPE and prediction signalling. While pwPEs did not significantly modulate gamma activity as a function of anxiety, they enhanced the amplitude of beta oscillations in state anxiety relative to control participants. This outcome is aligned with our recent findings in temporary anxiety during reward-based motor learning (Sporn et al., 2020). Below we discuss whether this result can be reconciled with hypotheses from generalised PC (Brown & Friston, 2013; Feldman and Friston, 2010) in which attention modulates precision weights on PEs through changes in synaptic gains and lower frequency oscillations (Bauer et al., 2014; Sedley et al., 2016). Our exploratory analysis of the neural representation of predictions suggested that anxiety states enhance beta oscillations during the generation of predictions about the stimulus-reward contingency. This finding should be taken with care as a between-group difference was observed exclusively in the stimulus-locked analysis, not in the outcome-locked analysis. If validated in future work, this outcome could be an indication that state anxious individuals exhibit a stronger reliance on prior beliefs (Bauer et al., 2014; Sedley et al., 2016), down weighting the role of PEs in updating predictions and suppressing gamma responses (Bauer et al., 2014). Overall, our results extend computational work on maladaptive learning in anxiety, suggesting that altered beta frequency oscillations may explain impeded reward-based learning in anxiety, particularly in volatile environments (Browning et al., 2015; Piray, Ly, et al., 2019; Pulcu & Browning, 2019).

Oscillatory correlates of precision-weighted prediction errors in state anxiety

In Hein et al. (2021), a 3-level HGF model best explained learning behaviour. Key findings were that state anxiety decreased the overall learning rate and led to an underestimation of environmental uncertainty and estimation uncertainty about the tendency towards a stimulus-reward contingency. As lower estimation uncertainty (greater precision) drove smaller pwPEs on that level, decreasing learning rates, here we predicted lower gamma activity during processing pwPEs in the state anxiety group. Given that both enhanced gamma and suppressed beta (and alpha) activity have been associated with pwPE during perceptual learning (Auksztulewicz et al., 2017) and with processing unexpected stimuli (Bastos et al., 2020), we also hypothesised concurrent higher alpha and beta modulation in state anxiety during pwPE signalling. More generally, gamma oscillations are anticorrelated with beta (and alpha) oscillations across the cortex, as shown for sensorimotor processing and working memory (Hoogenboom et al., 2006; Lundqvist et al., 2016, 2018, 2020; Miller et al., 2018; Potes et al., 2014).

Our results provide novel insight into how rhythm-based formulations of (Bayesian) PC—initially proposed for sensory processing—can be extended to learning about changing stimulus-reward associations. Our findings show that unsigned pwPEs about stimulus outcomes decreased alpha and beta activity 400–1000 ms post-outcome, separately in each group, suggesting that attenuation of lower frequency responses is associated with processing pwPEs independently of anxiety. Similar findings were observed when analysing separately the unsigned PEs about stimulus outcome—representing the surprise experienced by the participants—and after controlling for the concomitant effect of precision weights on the update of beliefs. Subsequently, during 1200–1570 ms, state anxiety relative to controls increased beta responses to pwPEs (and similarly for surprise) in sensorimotor and frontocentral electrode regions. This effect is closely aligned with the effects of state anxiety on beta activity (power and burst events) during processing pwPEs in reward-based motor learning (Sporn et al., 2020). Reduced alpha-beta activity was linked to pwPEs in Auksztulewicz et al. (2017). In addition, beta oscillations have also been previously shown to be involved in updating the content of sensory predictions in auditory processing and visuomotor learning paradigms (Sedley et al., 2016; Tan et al., 2016). This is also in line with our results, as the update steps of beliefs about the tendency of the stimulus-outcome contingency in the HGF are a function of the pwPEs on level 2. Accordingly, the increased beta activity in anxiety during the encoding of pwPEs could reflect smaller updates to predictions, explaining poorer learning in this group. The frontal and sensorimotor distribution of the beta effects, however, should be validated in

future work combining EEG/MEG with individual MRI scans to conduct convolution modelling in the individual source space.

While recent studies observed an attenuation of low frequency activity during encoding PEs/pwPEs in perceptual tasks, this effect was paralleled by increased gamma oscillatory activity (Auksztulewicz et al. 2017; Bastos et al., 2020)—in line with PC hypotheses. We failed to find any effects of pwPEs or unsigned PEs (surprise) on gamma activity, limiting the interpretation of the results. We outline below different accounts that could partially explain the lack of gamma-band effects in this study.

Hierarchical models of sensory information message-passing propose that suppression of PEs conveyed by gamma oscillations can occur through two main mechanisms: (1) the inhibitory effects of top-down predictions, and (2) postsynaptic gain regulation (Bauer et al., 2014; Brown & Friston, 2013; Larkum et al., 2004). Both mechanisms could partly account for our findings, yet not exclusively. On the one hand, the greater beta activity associated with predictions in state anxiety would convey inhibitory input to superficial pyramidal neurons encoding PEs, decreasing gamma (Bastos et al., 2012; Sedley et al., 2016). On the other hand, the estimation uncertainty σ_2 is the term modulating PEs about stimulus outcomes: $\varepsilon_2 = \sigma_2 \delta_2$ (**Eq. 2**). Accordingly, the lower σ_2 in state anxiety would attenuate pwPEs, while the putative associated gamma activity would decline.

Mechanistically, precision is thought encoded via postsynaptic gain, modulated by neurotransmitters and attentional processes (Bauer et al., 2014; Feldman & Friston, 2010; Friston & Kiebel, 2009a; Moran et al., 2013). Empirical investigations of sensory PEs implicate alpha and beta oscillations in the encoding of the precision of predictions about upcoming sensory input (Bauer et al., 2014; Palmer et al., 2019; Sedley et al., 2016). Because we investigated biases in learning about stimulus-reward contingencies in anxiety, the relevant precision term in our computational model was $\pi_2 (1/\sigma_2)$: the precision of the posterior belief about the tendency towards a stimulus-reward contingency. Increased precision π_2, or reduced estimation uncertainty σ_2, as we observed in state anxiety, was associated in our control GLM analysis with increases in alpha activity. One possible interpretation of our results is that the enhanced alpha modulation by σ_2 in StA could decrease synaptic gain, as proposed for attentional alpha (Bauer et al., 2014), thereby dampening the transmission of prediction errors about stimulus outcomes and the associated gamma oscillations.

Importantly, however, our results do not show that state anxiety attenuates gamma oscillatory activity during encoding pwPE or surprise (absolute PEs). Rather, our analysis suggests that, in our paradigm, even in a normative population such as our control group, encoding pwPEs and surprise about stimulus outcomes is not associated with gamma modulation. This outcome was unexpected as growing evidence indicates that cortical gamma activity is modulated by reward information in different domains beyond perception. Earlier work demonstrated a prominent gamma-band coupling between the frontal cortex and striatum in rats during reward processing and under pharmacological manipulation of dopamine (Berke, 2009). More recently, optogenetic stimulation of dopamine neurons in the rodent ventral tegmental area was shown to increase gamma activity in the medial prefrontal cortex (mPFC, Lohani et al., 2019). The effects were larger on sustained relative to phasic gamma and therefore it remains unclear whether dopamine in the PFC can provide transient teaching signals about stimulus-outcome contingencies (Ellwood et al., 2017). Yet a recent study in humans demonstrated a role of dmPFC gamma oscillations in the encoding of unsigned reward prediction errors during an exploration-exploitation dilemma (Domenech et al., 2020). Using invasive local field potential (LFP) recordings across the dmPFC and ventromedial PFC, this latter study provided compelling evidence that the rhythm-based PC mechanism proposed for sensory processing can account for decision making during exploration-exploitation behaviour. Because the pwPE and surprise regressors in our model are not directly coding reward PEs, it is possible that the lack of gamma effects in our study is due to our choice of experimental task and modelling approach. On the other hand, the reduced sensitivity of EEG (unlike invasive LFPs) to gamma oscillations may also account for the lack of gamma activity correlates of pwPEs during reward-based learning in our study. Using invasive LFP recordings in humans, when available, could be particularly relevant in future work to inform an extension of rhythm-based proposals of Bayesian PC to more general learning contexts.

More generally, EEG/MEG studies consistently show that frontocentral beta oscillations are modulated by positive reward feedback or predicting cues (Bunzeck et al., 2011; Cunillera et al., 2012; Marco-Pallares et al., 2008). The effects seem to stem from cortical structures linked to the reward-related fronto-subcortical network, such as the PFC (HajiHosseini et al., 2012; Mas-Herrero et al., 2015; O'Doherty, 2004). These studies, however, did not directly model the update of predictions about the stimulus-reward contingency via PEs. Beyond the Bayesian PC interpretations, a common view is that reduced beta activity in the prefrontal, somatosensory, and sensorimotor territories facilitates the encoding of relevant information to shape ongoing task performance (Engel & Fries, 2010; Schmidt et al., 2019; Shin et al., 2017). Accordingly, state anxiety could be more broadly associated with disrupting processing of relevant information through changes in beta oscillations, in line with some of the evidence on

EEG markers of social anxiety disorders (Al-Ezzi et al., 2020) and subclinical state anxiety (Sporn et al., 2020). This can also account for the lack of anxiety-related effects on the modulation of EEG signals in the time domain in our previous work (Hein et al., 2021). In that study we observed that pwPEs about the stimulus tendencies modulated the event-related potentials (ERP) exclusively in the control group during ~400-600 ms. This effect had a similar latency and topography to the P300-ERP components that had been associated with Bayesian surprise or precision in previous computational studies using EEG (Kolossa et al., 2015 ; Mars et al., 2008; Ostwald et al.,2012). Although not directly comparable, given that the amplitude of the P300 decreases with increased beta power (Enriquez-Geppert & Barceló, 2018; Polich, 2007), it is possible that the abnormally enhanced amplitude of beta oscillations in state anxiety during encoding pwPE may be paralleled by a reduced pwPE-ERP amplitude, explaining the null results in Hein et al. (2021). Overall, the current results suggest beta oscillations as a candidate marker of biased learning and attenuated belief updating in state anxiety and, as such, could be used as an intervention target in non-invasive brain stimulation, neurofeedback or pharmacological studies.

Biased predictions in state anxiety are associated with enhanced beta oscillations

Capturing neural modulations by predictions is challenging (Diaconescu et al., 2017). The neural representation of predictions could develop anywhere between the previous and current trial's outcome. To address this, we separately analysed oscillatory correlates of predictions about the tendency towards a certain stimulus-reward contingency, both post-stimulus and pre-outcome. Between-group effects were obtained exclusively in the stimulus-locked analysis, corresponding with an increase in beta activity between 200–640 ms in the StA relative to the control group, with a widespread topography. This effect was paralleled by a significant beta activity increase in the state anxiety group, yet exclusively in the outcome-locked representation, from −1000 to −500 ms prior to the outcome. The topography of this effect extended across central, parietal, and frontal electrode regions. Our analysis focusing on two different yet dependent windows was exploratory; we did not have a strong hypothesis concerning which time interval would be best suited to assess the effect of anxiety on neural oscillatory correlates of predictions, given the gradual modulation of predictions argued before (Diaconescu, Litvak, et al., 2017). The results are, accordingly, interesting yet preliminary and require validation in future work. Previous work associated alpha and beta oscillatory power to encoding predictions—potentially down-modulating precision weights (Auksztulewicz et al., 2017; Bauer et al., 2014; Sedley et al., 2016). Extant work, however, focused on sensory predictions and healthy control participants, which leaves open the question of how aberrant affective states may interact with oscillatory correlates of prediction signals. In our study,

interpretation of results in healthy controls is limited given the lack of a significant modulation by prediction in this group.

Further investigation is needed to identify the oscillatory responses to prediction and PE signalling in healthy controls, opening up rhythm-based accounts of Bayesian PC to learning stimulus-reward contingencies in volatile environments. Above all, our findings extend recent computational work on learning difficulties in anxiety (Browning et al., 2015; de Visser et al., 2010; Huang et al., 2017; Lamba et al., 2020; Miu et al., 2008; Piray, Ly, et al., 2019). We propose amplified beta oscillations as one neurophysiological marker associated with impaired reward-based learning and attenuated belief updating in state anxiety.

5. Supplementary Materials

Table of priors

Supplementary Table 1 provides the prior values of the free and fixed parameters in the winning model (3-level HGF) from Hein et al. (2021). In that study, we chose to estimate the parameters ω_2, ω_3, and ζ as in previous work (de Berker et al. 2016). Following also de Berker et al (2016), we chose to fix the coupling constant κ, the initial mean values of the belief trajectories and their initial variances: $\mu_2^{(0)}$, $\mu_3^{(0)}$, $\sigma_2^{(0)}$, $\sigma_3^{(0)}$.

Prior	Mean	Variance
κ	1	0
ω_2	-4	16
ω_3	-7	16
$\mu_2^{(0)}$	0	0
$\sigma_2^{(0)}$	0.1	0
$\mu_3^{(0)}$	1	0
$\sigma_3^{(0)}$	1	0
ζ	48	1

Supplementary Table 1. HGF perceptual parameter priors and initial values of the beliefs of the winning 3-level HGF model (see Hein et al., 2021). Means for the prior values of beliefs $\mu^{(0)}$ and variances $\sigma^{(0)}$ (provided in the space in which parameters are estimated). The table presents the perceptual parameters κ (coupling parameter) and ω_2, ω_3 (tonic volatility estimates on levels 2 and 3), and the response model parameter ζ. In Hein et al. (2021), $\sigma_2^{(0)}$, $\sigma_3^{(0)}$, and ζ were estimated in the log space, while $\mu_2^{(0)}$, $\mu_3^{(0)}$ and κ are estimated in the logit-space.

The prior values on these parameters were partially based on previous work. For instance, de Berker et al. (2016) used a prior variance on ω_2 and ω_3 of 16, as we did. The prior mean value on ω_2 was taken from Iglesias et al. (2013), and equal to −4. In that study, however, the winning model did not have a volatility level and therefore the study does not have ω_3 as a free parameter. We chose a prior mean value on $\omega_3 = -7$ (lower than the default on the HGF toolbox) because higher values did not provide plausible (i.e., properly regularized, reasonably smooth) belief trajectories. After discussing with Christoph Mathys, he suggested we reduce ω_3 until we obtain plausible belief trajectories (i.e. not jumpy trajectories). This suggestion is also included in the HGF toolbox scripts.

Concerning the decision noise parameter, ζ, de Berker et al (2016) did not report the prior value they used, so we followed Diaconescu et al. (2014).

Regarding fixed parameters, following de Berker et al. (2016), we set the starting values of the belief trajectories (mean [variance]) to $\mu_3^{(0)} = 1(0)$ and $\mu_2^{(0)} = 0$ (0). We also fixed the coupling parameter κ to 1 (0). The initial uncertainties of these beliefs ($\sigma_2^{(0)} = 0.1$ [0] and $\sigma_3^{(0)} = 1$ [0]) corresponded to the default settings of the toolbox, and we verified that small changes in these values (as in de Berker et al., 2016) had a negligible impact on the estimated belief trajectories.

Supplementary Results

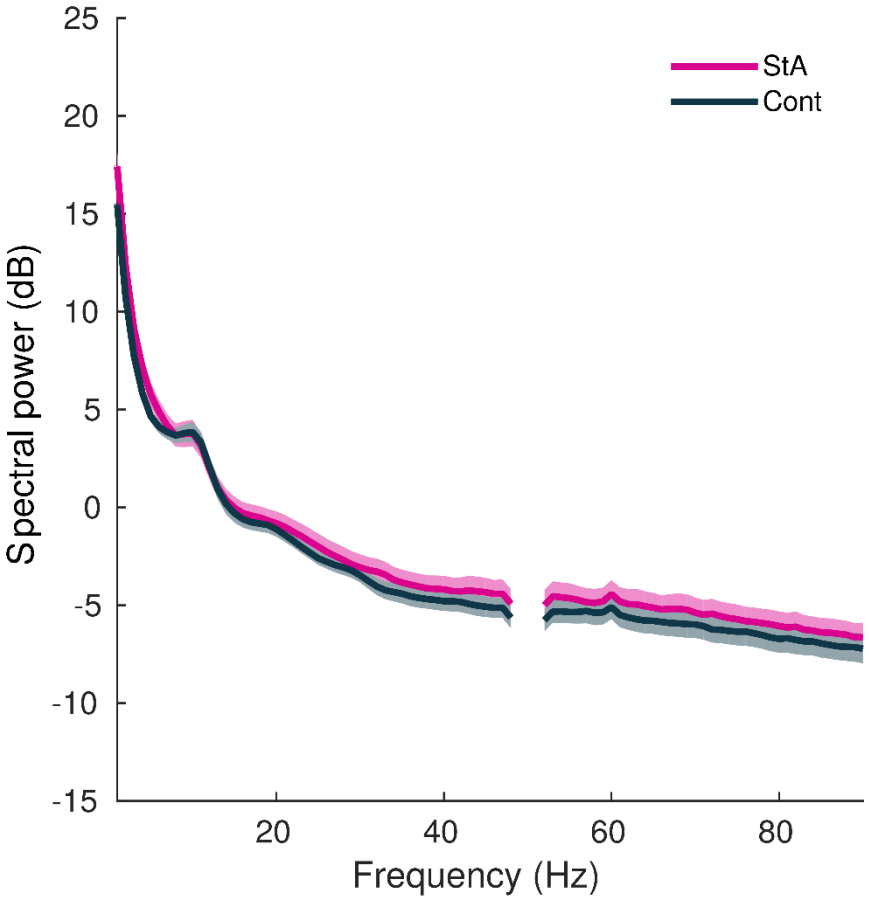
Control convolution GLM analysis

A control convolution model for oscillatory responses was conducted to assess the separate effect of σ_2 and abs(δ_1) regressors on the oscillatory activity. In this model, both continuous regressors σ_2 and abs(δ_1) as well as the discrete regressors coding for outcomes were included. The absolute value of PEs is usually termed surprise (e.g. de Berker et al., 2016).

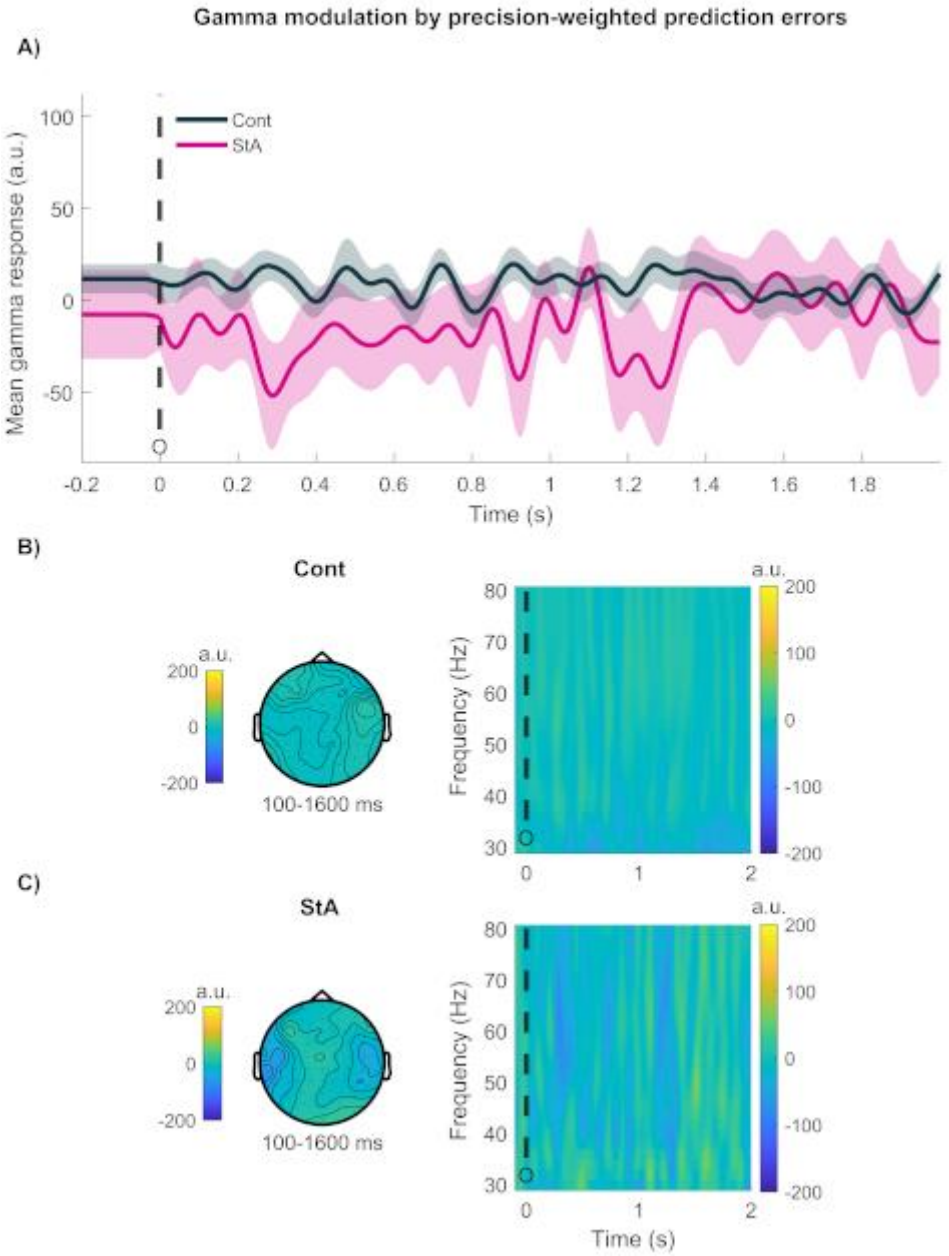
Within-subject effects to surprise (absolute PE) about stimulus outcomes: In both groups of participants we found that the surprise regressor significantly modulated the alpha oscillatory responses in the post-outcome interval relative to a reference (baseline) level (Cont: one positive cluster, P = 0.01 within 100–370 ms; StA: one positive cluster, P = 0.03 within 150–440 ms post-outcome, FWER-controlled, see **Supplementary Figure 3**). For both groups, this increase in alpha activity had a frontocentral distribution. In addition, beta activity was significantly modulated by PE regressor also in each group separately (one negative cluster in each group, P = 0.001, FWER-controlled, see **Supplementary Figure 4**). In control participants, the reduction in alpha activity was observed within 380–1140 ms and had a posterior centroparietal distribution. In StA, alpha TF responses dropped within 460–1000 ms and spread across centroparietal and right frontal electrodes

Within-subject effects to precision weights σ_2**:** There was a significant negative modulation of alpha oscillatory responses by the precision weights regressor in the Cont group relative to baseline interval (one negative cluster within 1270–1530 ms; P = 0.024, FWER-controlled, see **Supplementary Figure 5**). This cluster consisted of frontocentral and parieto-central electrodes. No further effects of this regressor were found in the beta or gamma band for control participants, or in any frequency range for the StA group.

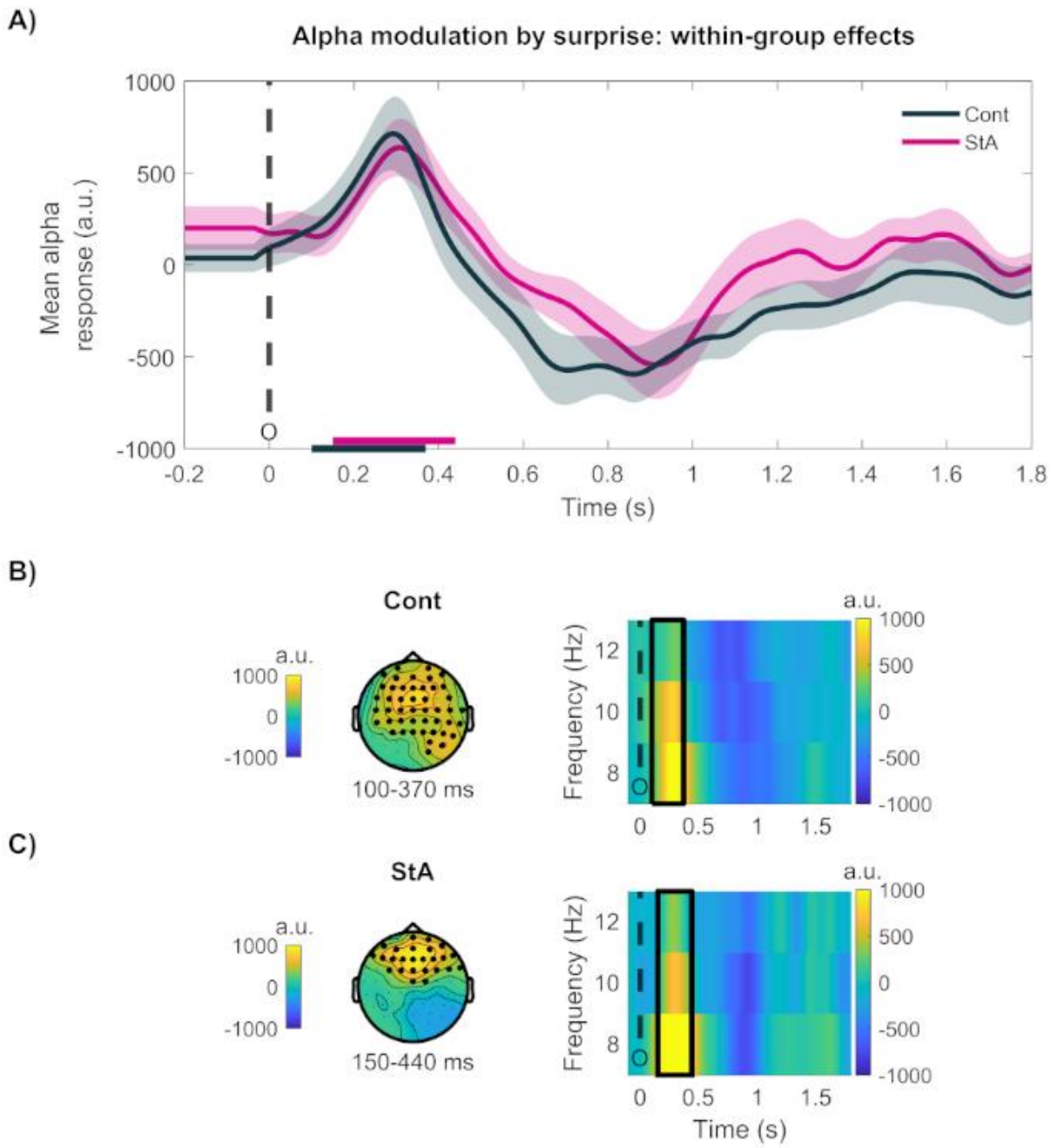
Supplementary Figures



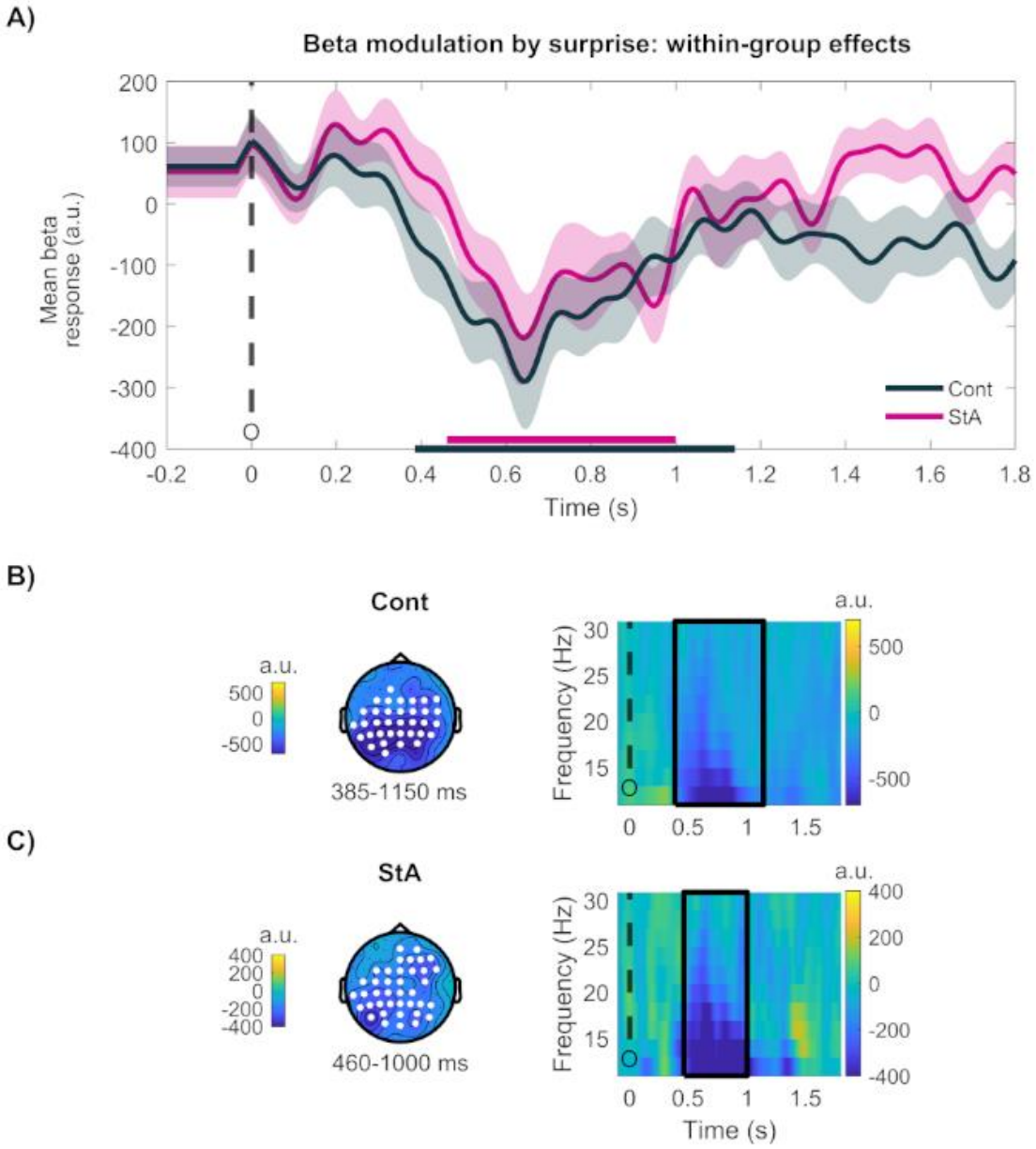
Supplementary Figure 1. Grand-average of the raw spectral power during task performance. The raw power spectral density during task performance was converted into decibels (dB: $10*\log_{10}$), and grand-averaged separately in state anxious (pink) and control (black) participants. Shaded areas denote the standard error of the mean (SEM). There was no significant difference in raw power between groups ($P > 0.05$, cluster-based permutation test).



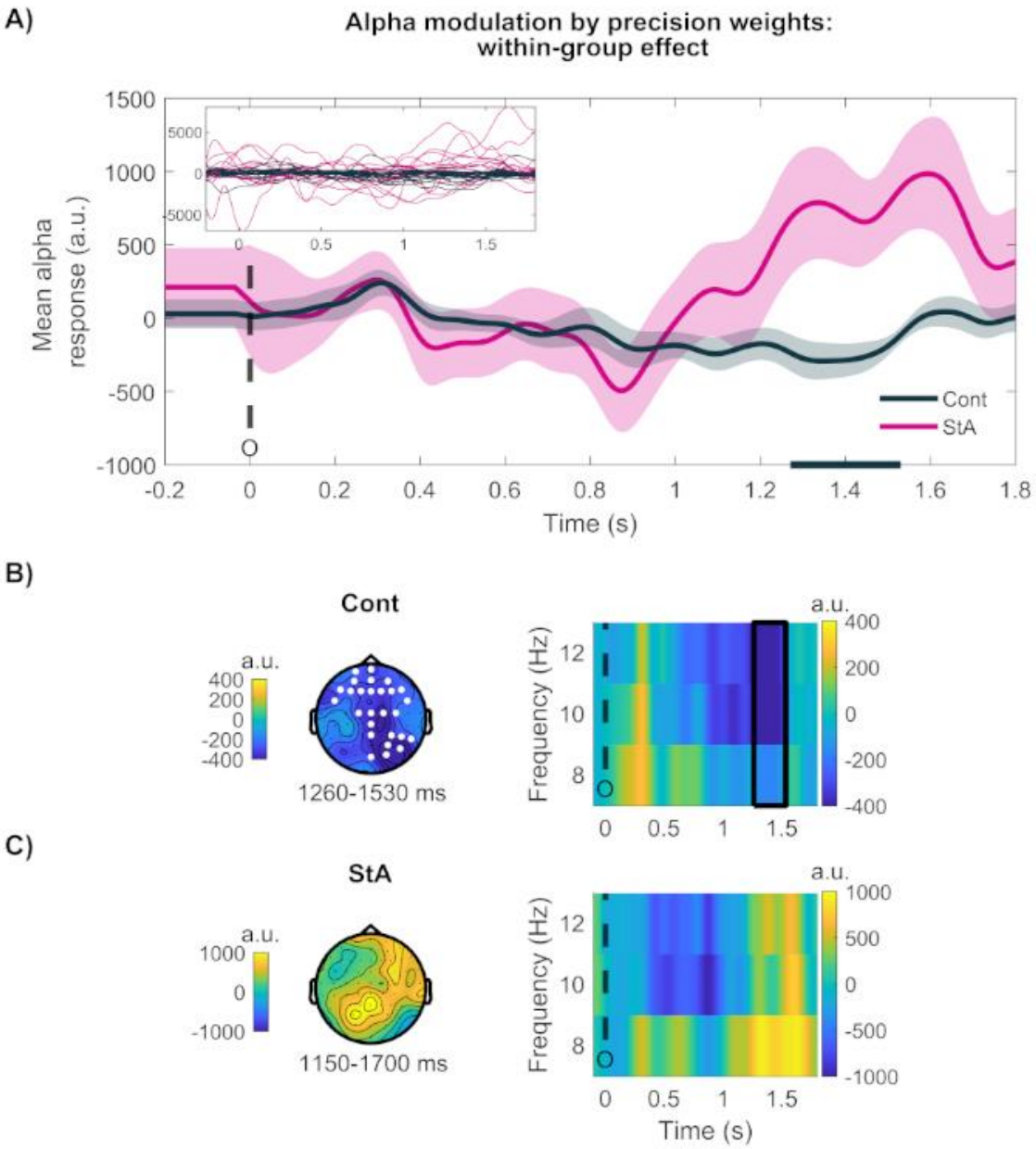
Supplementary Figure 2. Gamma activity modulated by precision-weighted prediction errors updating belief estimates about the stimulus-reward contingency. A) The average gamma response (31–80 Hz) in arbitrary units (a.u.) to pwPEs on level 2 ($|\varepsilon_2|$) in each group (Controls, black; StA, pink), with time in seconds (s) on the x-axis. **B)** The correlates of pwPEs $|\varepsilon_2|$ in gamma activity in the Cont group. The left topographic distribution shows activity between 100–1600 ms; the right time frequency image is for $|\varepsilon_2|$ in gamma activity in all electrodes presented 0–2 s from the outcome (black dashed line, 'O'). **C)** StA topographic representation of $|\varepsilon_2|$ in gamma activity (left), with the time frequency image (right) between 0–2 s (outcome given by the black dashed line, 'O').



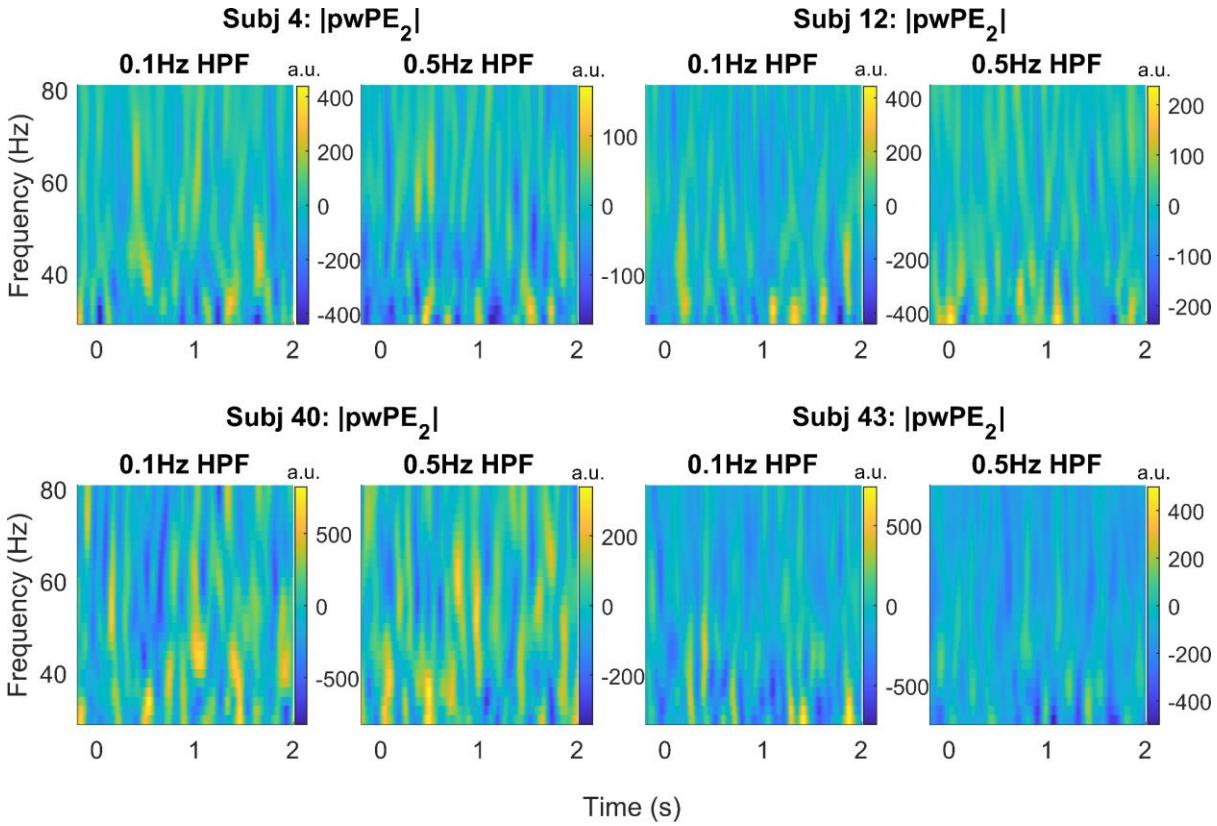
Supplementary Figure 3. Within-group effects in alpha oscillatory activity modulated by surprise (absolute PEs) about stimulus outcomes. A) The time course of the average oscillatory response (8–12 Hz) to absolute PEs about stimulus outcomes (abs[δ_1]) in each group (StA, pink; Cont, black), given in arbitrary units (a.u.). The time range of the significant clusters in each group are denoted by the horizontal bars on the x-axis in their respective colours. **B)** In the Cont group the alpha activity was significantly modulated by the abs[δ_1] regressor (one positive cluster spanning 100–370 ms relative to baseline; *P* = 0.01, FWER-controlled). The left panel displays the topographic distribution of this effect in arbitrary units (a.u.), which spread primarily over frontocentral electrodes (denoted by the black dots). On the right is the time-frequency image (a.u.) for the cluster averaged across the electrodes pertaining to the cluster (the outcome latency is marked by the black dashed line; the black square denotes the significant cluster). **C)** Follows the same format as (B) but showing the within-group effect in the StA group. We found a significant positive difference from baseline level in alpha activity between 150–440 ms (*P* = 0.03, FWER-controlled) in frontocentral electrodes.



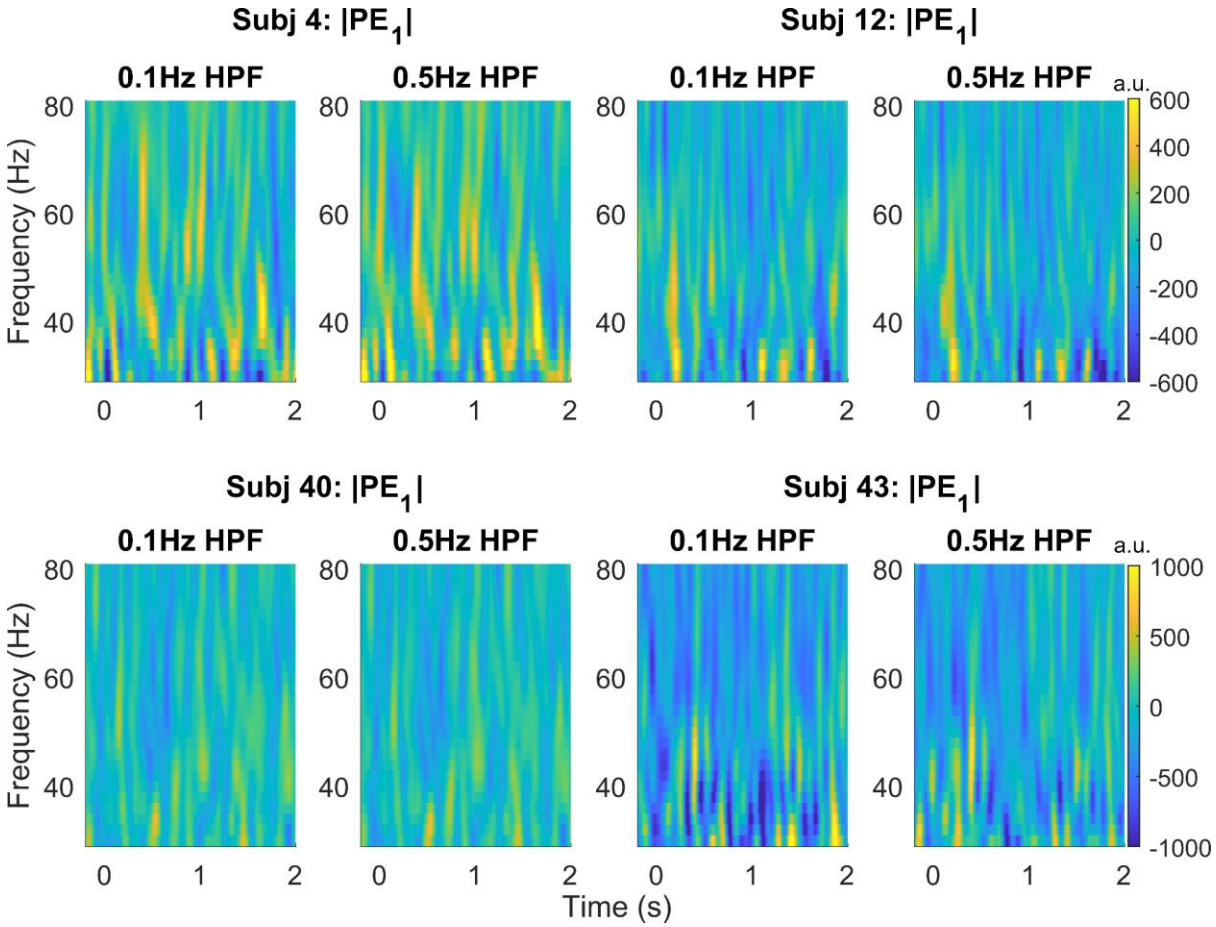
Supplementary Figure 4. Within-group effects in beta oscillatory activity modulated by the absolute value of prediction errors (surprise) about stimulus outcomes. **A)** Time course in arbitrary units (a.u.) of the average oscillatory response between 13–30 Hz by the absolute PE regressor (abs[δ_1]) in each group (Cont, black; StA, pink). The latency of the dependent-samples significant cluster obtained for each group is denoted by a horizontal bar on the x-axis in the respective colour. Dashed black lines mark the onset of the outcome ('O'). **B)** In Cont participants, beta-band oscillations were significantly within 380–1140 ms post-stimulus, relative to a baseline level (one negative cluster, P = 0.001, FWER-controlled). Left: Topographic distribution of this effect, which spread across posterior centroparietal electrodes. Right: Time-frequency image for absolute PEs about the stimulus outcomes averaged across the cluster electrodes (dashed black lines mark the onset of the outcome; the black square denotes the significant cluster). **C) Similar to** (B) but presenting the within-group effect of abs[δ_1] on beta activity in the StA group. We observed a significant drop in beta activity between 460–1000 ms relative to baseline (one negative cluster, P = 0.001, FWER-controlled), with a centroparietal and right frontal distribution.



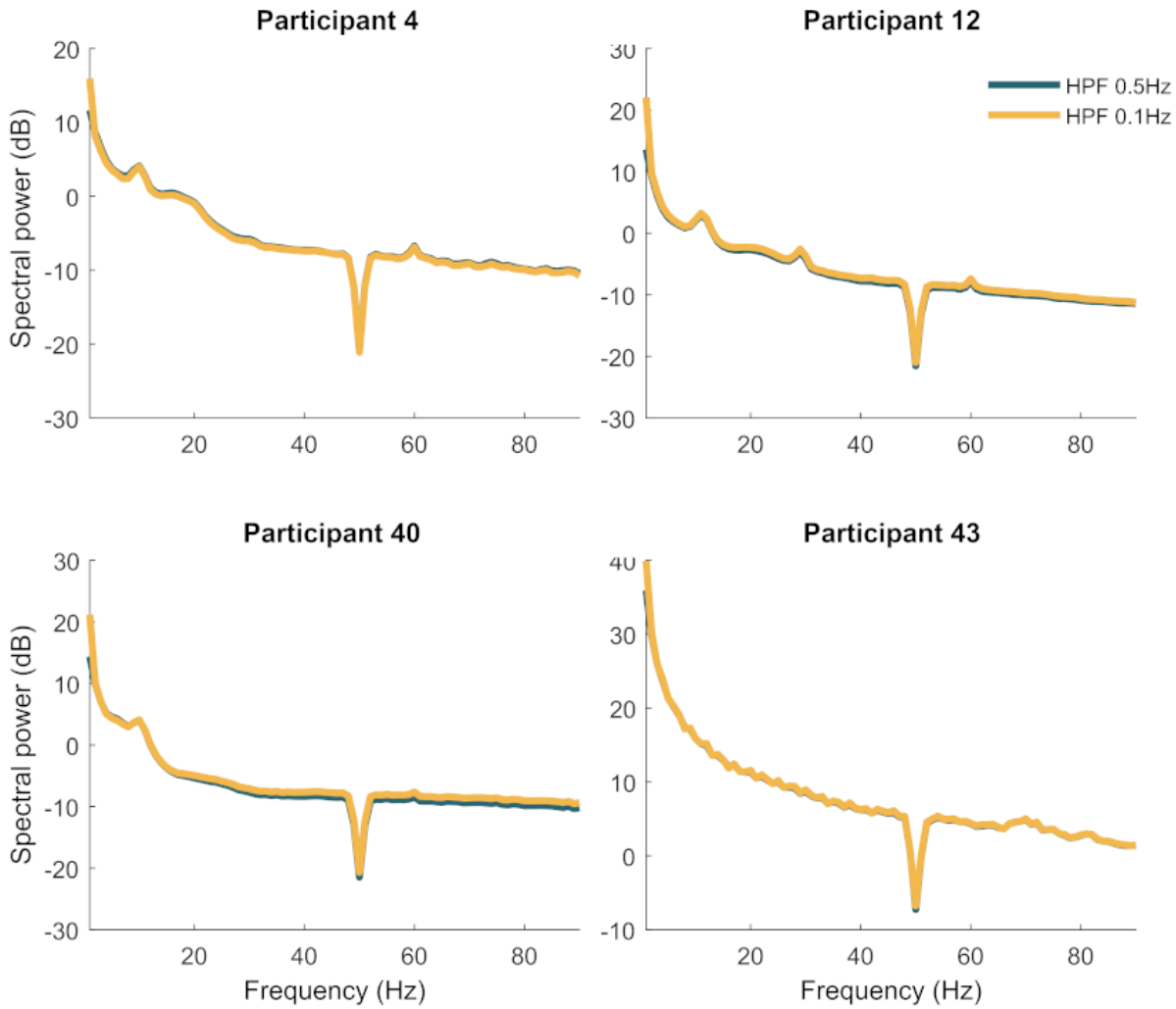
Supplementary Figure 5. Within-subject effects on the modulation of alpha (8–12 Hz) activity by precision weights. **A)** Time course in arbitrary units (a.u.) of the average alpha response (8–12 Hz) to precision weights (σ_2) on level 2 in each group (Controls, black; StA, pink). The x-axis displays the time interval of the dependent-samples significant cluster found in the Cont group only (black horizontal line). **B)** Alpha activity was significantly negatively modulated by the precision weight regressor at the within-group level in the Cont group (one significant cluster relative to baseline within 1270–1530 ms, P = 0.024, FWER-controlled). On the left is the topographic distribution of this effect, in frontocentral and parieto-central electrodes. On the right is a time-frequency image for the precision weight regressor, averaged across the cluster electrodes. The black dashed line provides the onset of the outcome, and the black square shows the time-frequency range of the significant cluster. **C)** Follows the same format as (B), however we did not observe a significant difference in alpha oscillations from baseline in the StA group. We present the topographic distribution of activity in 1150–1700 ms to show the observed increase in alpha oscillations at a similar time interval to the Cont group. We also provide on the right the time-frequency response image for StA alpha activity from the precision weight regressor in all electrodes.



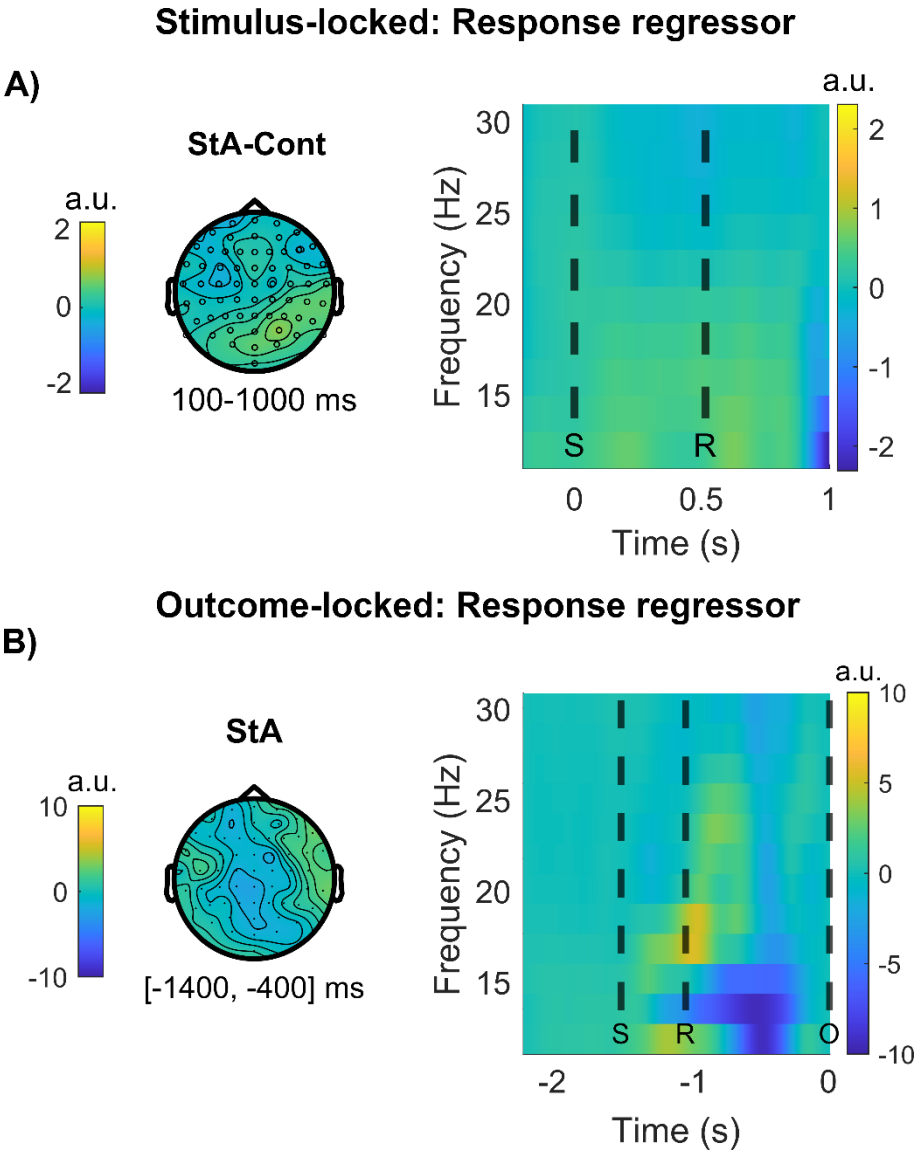
Supplementary Figure 6. Influence of the high-pass filter cutoff on the time frequency images representing modulation by the pwPE regressor. The TF images **(** arbitrary units, a.u.) are displayed in four representative participants after using a 0.1 Hz or 0.5 Hz cutoff as high-pass filter during pre-processing. The figure shows that in each participant, the TF images have a generally higher signal-to-noise ratio (SNR) after applying a 0.1 Hz cutoff; however, neither the 0.1 Hz or 0.5 Hz high-pass filter settings demonstrate a prominent gamma modulation by the pwPE regressor.



Supplementary Figure 7. Similar to Supplementary Figure 6 but using the absolute PE about stimulus outcomes abs(PE_1), termed abs[δ_1] in the main text, as regressor.



Supplementary Figure 8. Raw power spectral density during task performance (TB1, TB2 blocks) after using a 0.1 Hz or 0.5 Hz cutoff as high pass filter in four representative participants. The power spectral density (PSD) obtained at 1-90 Hz after applying a 0.1 Hz high-pass filter is shown in yellow, while the PSD after a 0.5 Hz cutoff is displayed in dark blue. The raw PSD was converted into decibels (dB: $10*\log_{10}$). As shown in the figure panels, the choice of a 0.5 Hz cutoff used in our study only minimally affected the signal-to-noise ratio in the general PSD level.



Supplementary Figure 6. Stimulus and outcome-locked beta (13–30 Hz) changes to response regressor (related to Figure 8). All responses were provided using the right hand. **A)** An independent-samples statistical analysis on stimulus-locked beta activity by the response regressor—in arbitrary units (a.u.)—using cluster-based permutations showed no significant differences between StA and Cont between 100–1000 ms post-stimulus presentation. The right column shows the time-frequency image response in beta activity for the between-group data averaged over all electrodes (0–1 s, stimulus-locked). Dashed black lines represent the average time of the stimuli presentation 'S' and the response 'R'. **B)** For the state anxiety group (StA), we tested the within-group effect of the response regressor on the oscillatory activity locked to the outcome (between −1000 to −100 ms). This analysis showed no significant change in beta activity from baseline. The corresponding time-frequency image to the right shows beta activity averaged across all electrodes (−2 to 0 s, outcome-locked, dashed black lines for the average time of the stimuli presentation 'S' and the response 'R'.

Chapter 4: Linking anxiety and motivation to computations of uncertainty during reward-based learning

A portion of the data reported in this chapter was collected by Fahima Ahmed, to whom I am sincerely grateful. [9]

[9] Author contributions M.H.R. and T.P.H. designed the experiment, F.A and T.P.H collected the data, M.H.R. and T.P.H. analysed the data, M.H.R., and T.P.H. wrote code for data analysis, T.P.H. wrote the manuscript. M.H.R reviewed the manuscript. Sample size was estimated by M.H.R.

Abstract

Neuropsychiatric conditions such as depression and bipolar disorder exhibit motivational deficits and interfere with reward processing. How negative emotional states like anxiety modulates motivation to learn from reward is unclear. Transient states of anxiety can interfere with reward prediction error signals vital for learning in computational theories of adaptive behaviour. Here we coupled learning about reward to reducing anxiety in an uncertain environment. We recruited 45 participants to complete a binary probabilistic reward-learning task in a volatile environment to study behavioural group differences. Participants needed to estimate the probability of images being either rewarded or unrewarded while we recorded continuous electrocardiographic (ECG) signals and self-reports of anxiety. To manipulate motivation, we induced anxiety by threatening a secondary psychosocial stress task in one group of 15 participants. Importantly, reward in that group was operationalised to reduce the time of the subsequent psychosocial stress task. The motivation group was compared with a state anxiety control group (where anxiety was uncoupled from reward) and a control group (where no anxiety was induced)—each consisting of 15 participants. Computational modelling with a hierarchical Bayesian filter was used to succinctly describe participants' behaviour and estimates of uncertainty. There were no differences between the three groups' self-reported anxiety and physiological responses from the heart, or overall performance of the reward-learning task. Computational modelling demonstrated that participants' responses were best described by a model that is informed by dynamic estimates of task volatility. However, there were no group differences in posterior estimates of volatility, uncertainty about the stimulus-outcomes and their changes, or environmental uncertainty. Our results suggest that anxiety was not successfully induced. In contrast to our previous work, behavioural modelling results indicated that our probabilistic binary reward task using a stable degree of volatility can produce decisions informed by trial-wise estimates of changes to the probabilistic contingencies. These findings indicate that our experimental manipulation did not sufficiently induce anxiety and was not sensitive enough to detect changes in motivation. Future work may benefit from using lab-based shock to more reliably drive and control anxiety levels.

1. Introduction

The impact of anxiety disorders on public health is an increasingly high-priority issue (Grillon et al., 2019; Stein & Craske, 2017). Clinical anxiety notwithstanding, everyday experiences with anxiety can also be distressing, unpleasant, and lead to a lower quality of health (Carleton, 2016; Chisholm et al., 2016; Dugas et al., 2005). In spite of this, we are only just beginning to understand how anxiety influences learning and decision making in the brain, and in turn, how learning biases maintain anxiety. Elegant work experimentally inducing anxiety has unexpectedly shown how anxiety can both impede and benefit cognition (Robinson et al., 2013)—an outcome strongly conditional upon the context and nature of the task performed. More recently, evidence of performance impairments have been extended to a computational reward-learning context (see **Chapter 2** and Hein et al., 2021; Huang et al., 2017; Sporn et al., 2020). Yet these studies did not focus on the potential benefits and adaptive value of anxiety on reward-based learning. Here we seek to address this by investigating whether motivation induced by anxiety can improve reward-based learning in an uncertain and volatile learning condition. We take inspiration from previous studies detailing the role of motivation in invigorating ongoing instrumental learning behaviour (Niv, 2009; Niv et al., 2006, 2007) and from work indicating that anxiety activates the implicit goal of uncertainty reduction (Raghunathan & Pham, 1999). Our study focuses on an unmedicated population for three reasons: 1) to connect similar work in alternative learning contexts to the reward-learning domain 2) to understand how anxiety may transiently motivate healthy humans in response to everyday encounters with anxiety, and 3) to make important bridges in understanding between healthy and clinical populations and work on non-human animals (Grillon et al., 2019; Valton et al., 2019). Of prime consideration is uncovering how motivation and its effect on learning about reward might inform future treatments of anxiety and anxiety disorders (Moutoussis et al., 2018).

Anxiety produces cognitive, physiological, and affective changes driven by the disproportionate worry over upcoming uncertain events (Grupe & Nitschke, 2013; Raymond et al., 2017; Robinson et al., 2013). Anxiety induces a state of apprehensive avoidance with simultaneously increased sensory vigilance (Robinson et al., 2013). This may be an adaptive response to prepare actions for potential negative outcomes in uncertain environments—a response consistent across multiple species (Sokolowska & Hovatta, 2013). But as with many adaptive responses in animals, anxiety can carve both favourable, and if commissioned too frequently, unfavourable consequences—turning adaptive to maladaptive (Robinson et al., 2013). Here we focus on anxiety acting as an extrinsic motivator to learn about reward (for a

review on the behavioural, neural, and pharmacological effects of anxiety, see Grupe & Nitschke, 2013; Hartley & Phelps, 2012; Slee et al., 2019).

Studies using the threat of shock in a lab have demonstrated that anxiety—in line with adaptive responses—enhances early sensory processing, producing a hypervigilant state which can benefit performance on perceptual tasks (Cornwell et al., 2017; Grillon et al., 2019). However, in anxiety disorders, this is exaggerated and overapplied to everyday unthreatening contexts. Shock-elicited anxiety also facilitates the detection of threatening signals (Grillon & Charney, 2011; Robinson et al., 2011) and attention to task-relevant information (Cornwell et al., 2011; Edwards et al., 2010; Kim et al., 2020). In the context of these tasks, anxiety facilitates performance. By contrast, anxiety impedes probabilistic learning in economic decision-making tasks (de Visser et al., 2010; Hartley & Phelps, 2012; Jiang et al., 2018; Miu, Heilman, et al., 2008; Remmers & Zander, 2018) and interferes with the optimal adaption of learning rates to changes in the learning environment (Browning et al., 2015; Pulcu & Browning, 2017).

As anxiety is consistently shown to influence cognition, recent research has turned to ask to what extent anxiety alters the computational processes driving perception, learning, and action in the brain (Bishop & Gagne, 2018; Pine, 2017; Raymond et al., 2017). Work on healthy humans has investigated the role of computational and neural mechanisms critical to successful perception and learning (Behrens et al., 2007; Bland & Schaefer, 2012; Doya, 2008; Iglesias et al., 2013; Payzan-LeNestour et al., 2013; Summerfield et al., 2011; Yu, 2007; Yu & Dayan, 2005). These studies highlight the brain's use of hierarchically related error discrepancy signals, prediction errors (PEs), and their uncertainty weighted counterpart, precision-weighted prediction errors (pwPE), in updating an internal model of the environment (Friston, 2005, 2010; Iglesias et al., 2013; Knill & Pouget, 2004; Mathys et al., 2011; Moran et al., 2013; Mumford, 1992; Rao & Ballard, 1999; Srinivasan et al., 1982). These accounts are couched in the Bayesian predictive coding (PC) process theory (Aitchison & Lengyel, 2017; Shipp, 2016), which emphasises how humans seek to refine a hierarchical generative model by using approximate Bayesian inference (Friston, 2005; Knill & Pouget, 2004). According to Bayesian PC, we seek to reduce uncertainty and the associated error signals across all levels in the hierarchical cortical structure of the brain (Friston & Kiebel, 2009a).

One of the many triumphs of the computational approach in neuroscience and psychology is its ability to bring together different levels of analysis, forming a bridged understanding of learning and conditioning (Friston et al., 2014; Pezzulo et al., 2018; Piray, Dezfouli, et al., 2019; Piray & Daw, 2020b). More specifically, this technique has provided good neural foundations for the normative evaluation of the challenges organisms confront through

adaptable mechanisms of PE updating (Courville et al., 2006; Daunizeau, den Ouden, Pessiglione, Kiebel, Friston, et al., 2010; Daunizeau, den Ouden, Pessiglione, Kiebel, Stephan, et al., 2010; Dayan et al., 2000; Dayan & Long, 1998; Gershman et al., 2010). As touched on above, these describe learning as statistical inference, blending experience with some goal with responses from cues and action outcomes. This recasting of learning as statistical inference has driven a compelling agenda of enquiry into the brain's approach to identifying, tracking, and estimating uncertainty concerning its beliefs and how these shape learning (Behrens et al., 2007; Iglesias et al., 2013; Mathys et al., 2011; McGuire et al., 2014; Nassar et al., 2010; Piray & Daw, 2020b; Soltani & Izquierdo, 2019).

The emphasis of Bayesian PC on uncertainty is an important one, as recent computational work highlights how affective disorders may be driven by misestimation of uncertainty (Paulus & Yu, 2012; Pulcu & Browning, 2019). In fact, anxiety is strongly coupled to uncertainty, with uncertainty often experienced as distressing and harmful (Carleton, 2016; Dugas et al., 2005; Dugas et al., 1998; Grupe & Nitschke, 2013; Hartley & Phelps, 2012; Paulus & Yu, 2012; Pulcu & Browning, 2019). To understand better how uncertainty shapes learning and decision making, recent work has classified uncertainty into three principal forms: *expected*, *estimation*, and *unexpected* (*environmental*) uncertainty (Bland & Schaefer, 2012; de Berker et al., 2016; O'Reilly, 2013; Soltani & Izquierdo, 2019; Yu & Dayan, 2005). *Expected* uncertainty represents the intrinsic and *irreducible* uncertainty produced from any complex probabilistic environment (de Berker et al., 2016; Yu & Dayan, 2005). *Estimation* uncertainty, by contrast, reflects our insufficient knowledge about the environmental statistics that create our sensory impressions, which we can reduce through learning (Bland & Schaefer, 2012). And when the world around us changes (*unexpected* environmental uncertainty: with the true change rate denoted by the term *volatility*), our prior beliefs about those environmental statistics may become outdated (Behrens et al., 2007; Dayan & Yu, 2003; O'Reilly, 2013).

Consider the following illustration of uncertainty and volatility using an everyday example. Imagine an atypically cold day in late summer. Should you readily update your belief, donning thermals and a jumper the following day? This response indicates adapting to the information and inferring a change to the climatic environment. Or, did that drop in temperature represent an atypical solitary event, merely noise? If so, you might infer that the atypical climatic event was not indicative of environmental change but rather unpredictability in temperature, donning a tee-shirt the next day. Here it would be critical to represent our prior belief and uncertainty about general seasonal patterns and offset them against our estimates of sensory information and uncertainty. This example also highlights the distinctly Bayesian technique of combining multiple sources of information to inform more accurate predictions. Perhaps you seek advice

from the weather forecasting news, with high uncertainty. Alternatively, take to the thermometer and measure the temperature continuously throughout the day, with lower uncertainty.

Importantly, previous work has linked volatility with pathological decision making in depressive and anxiety disorders (Hartley & Phelps, 2012; Huys et al., 2015; Paulus & Yu, 2012). And further research has shown that anxiety can alter the way we estimate uncertainty and bias learning from uncertain and changing environments (Browning et al., 2015; Gagne et al., 2018; Hein et al., 2021; Huang et al., 2017; Miu, Heilman, et al., 2008; Pulcu & Browning, 2019). This body of work about the atypical processing of uncertainty and abnormal inference are thought to contribute to a wide range of mental health disorders, particularly in anxiety and depression, but also in other neuropsychiatric conditions such as schizophrenia and autism (Brazil et al., 2017; Browning et al., 2015; Cole et al., 2020; Deserno et al., 2020; Diaconescu et al., 2020; Katthagen et al., 2018; Lawson et al., 2017; Paliwal et al., 2019; Piray, Ly, et al., 2019; Powers et al., 2017; Pulcu & Browning, 2017).

High levels of trait anxiety have been shown to leverage learning adjustments during decision making. Specifically, high trait anxiety leads to difficulties adapting the rate of learning (the extent to which previous and current sensory inputs inform belief updates) in changing environments in both aversive (Browning et al., 2015) and reward-learning contexts (Huang et al., 2017). These insensitivities to task volatility express themselves in changes to pupil dilation (Browning et al., 2015) and neural responses in the dorsal anterior cingulate cortex (dACC) linked with the processing of learning rates (Piray, Ly, et al., 2019). Other computationally based studies have demonstrated how anxiety is connected to biased inhibition responses and impaired extinction of learned fear associations (Duits et al., 2015; Grillon et al., 2017; Christian Grillon et al., 2017; Mkrtchian, Aylward, et al., 2017; Pittig et al., 2018; Pulcu & Browning, 2019; Robinson, Krimsky, et al., 2013). And more recent empirical work has shown biases to evidence accumulation (Kim et al., 2020). Our recent studies in **Chapters 2 and 3** have shown unbalanced uncertainty estimates but in the reward-learning domain. We demonstrated how transient experiences with anxiety in otherwise healthy volunteers can lower the overall learning rate and produce underestimates of *environmental* and *estimation* uncertainty about the reward tendency. Each provides evidence for a degree of fixity in beliefs or biased learning when experiencing anxiety, especially when interacting with uncertainty and volatility (Browning et al., 2015; Grupe & Nitschke, 2013; Pulcu & Browning, 2019). However, more work needs to be done linking different findings from the different forms of anxiety (state, trait, clinical disorders) and types of cognition (attention,

perceptual learning, reward learning, motor learning) using the available suite of computational and cognitive methods.

Here, in a similar study on state anxiety and reward-based learning that directly builds on **Chapter 2**, we use an identical state anxiety group and control group. However, in addition, we also tested a state anxiety group that could explicitly reduce the time required to perform a threatened secondary anxiety-inducing task by performing well during the preceding reward-learning task. We thereby coupled the threat of an anxious event to the reward learning performance. We define this coupling as an extrinsic external motivator. Using this manipulation and design, we aimed to test whether anxiety could facilitate learning about reward, potentially tapping into a theorised form of adaptive uncertainty reduction (Raghunathan & Pham, 1999).

In our previous **Chapter 2**, the threat of anxiety remained until it was revoked after completing the second task block of reward learning. The anxiety-inducing event was set to occur after the second task block irrespective of how the participants performed. We thus hypothesise that if reward learning performance is coupled to our anxiety manipulation, we may observe a motivational improvement in overall reward learning performance, with estimates of uncertainty predicted to be similar with controls. Our prime rationale for investigating the motivational component to reward learning in anxiety is to provide an experimental model based on non-clinically anxious individuals that might aid in the potential treatment of anxiety disorders. The coupling of anxiety to reward may prove a fruitful approach to shift feelings of anxiety away from apprehension, and on to viewing anxiety as an adaptive tool to resolve uncertainty about the world. Similarly to how anxiety has been shown to benefit early sensory processing, anxiety may drive faster learning about rewards in changing environments if learning to obtain reward reduces the likelihood of an anxiety inducing event. And this resolution of uncertainty may, in turn, lessen feelings of anxiety.

Motivation is, however, a nebulous field of research (Berridge, 2004). Broadly, motivation is commonly defined as object-oriented behaviour, orienting and pushing an animal to act in line with its goal and the goal value (Pessiglione et al., 2017). Oftentimes, this is toward maximising pleasure or minimising pain (Berridge, 2004; Hassin et al., 2009; Madan, 2017). Motivation shapes the salience of goals, calibrates attention, and drives behavioural control to achieve some outcome (such as food or water). These can be of instrumental value, to obtain some distinct outcome from the task at hand. Or these can be a value assigned for the inherent satisfaction of working on some goal (Oudeyer & Kaplan, 2009). Others have cast motivation as either specific to an outcome (the effect of directing) or independent of some outcome (an

energising effect: see Niv et al., (2006) for a normative treatment of motivation). These two core attributes of motivation have led to psychologists defining motivation using two broad categories: *intrinsic* and *extrinsic*.

Animals, and in particular, humans, display a predilection to general motivations that drive exploration, curiosity, investigation of environments, and playful engagement in novel activities (Friston, Lin, et al., 2017; Oudeyer & Kaplan, 2009; Schwartenbeck et al., 2013). These forms of motivation are typically described as intrinsic, as the activity was performed simply for fun, internal satisfaction, or the challenge (Ryan & Deci, 2000). They are not, however, classified as homeostatic, as they do not address, for example, tissue deficits brought on by hunger or pain states that motivate organisms to regain homeostasis (Oudeyer & Kaplan, 2009). For instance, we can paint pictures to enjoy painting and have fun (intrinsic motivation), but while painting pictures could be utilised to make money and pay for food, painting can also be enjoyed independently of that homeostatic reward. Psychologists believe these intrinsic motivations are critical for cognitive and sensorimotor development (Deci & Ryan, 2000, 2010). Computational work has further shown these intrinsically motivated behaviours facilitate the successful prediction of the consequences of our actions (Friston, Lin, et al., 2017). Psychologists and neuroscientists both argue that active learning driven by intrinsic interest is indispensable for providing understanding about the world that can be exploited for development in the absence of external reward (Friston, Lin, et al., 2017).

On the other hand, we may value action as it may gain us other valuable outcomes (Pessiglione et al., 2017). In our painting illustration above, we might paint to increase our skill to later study art and pass an exam (extrinsic motivation). Here, succeeding in the examination is the extrinsic motivator and the behaviour instrumental in achieving that goal. As such, extrinsic motivation is defined by actions executed to obtain separable rewards or avoid negative outcomes (Oudeyer & Kaplan, 2009). These 'goal-directed' actions can be multidimensional, and they might consist of positive and negative parts simultaneously, such as rewards and punishments (Pessiglione et al., 2017). As an example, to motivate behaviour, the negative aspects of reaching the pinnacle of a career (as say the chief executive officer of a global company with ultimate responsibility for results and performance) should not outweigh the positive values of the set goal (money, power, and potentially self-esteem, Pessiglione et al., 2017).

In this study, our definition of motivation is influenced by the reinforcement learning literature (Sutton et al., 1998). We define motivation here as the mapping between the time required to perform a threatened social stress task and the reduction of that time. As such, this temporal

discount function represents a valued utility, instrumental to drive behaviour to achieve the goal of avoiding anxiety (Niv et al., 2006). Our motivator is external and extrinsic, as the reward-learning task is not performed for the sake of being the best at reward learning or for the fun of obtaining points, but for the entirely independent outcome of avoiding a potentially unpleasant secondary stressful task. Thus, the question we ask is: how does eliciting anxiety by means of the threat of a social stressor modify the coupled reward-learning behaviour and model estimates of uncertainty in a dynamic environment?

In animal studies, motivation might be operationalised by measuring how much effort an animal is willing to expend for some reward (Dickinson & Balleine, 2002). Task vigour can be represented by how fast an animal responds to receive a reward (or more plainly, how hard they are willing to work, Niv et al., 2005). As such, when the net reward rate and value for some outcome are higher (for example a hungry rat), the requisite behavioural responses to perform optimally to gain reward should be faster (Niv, 2009; Niv, Daw, et al., 2005; Niv et al., 2006, 2007). For our task, optimal task performance would translate to a decrease in the overall percentage of errors performed across the reward-learning task.

In terms of neuromodulation, motivation (Mogenson et al., 1980; Wise & Rompre, 1989; Wise, 2004) and reward learning (Berridge & Robinson, 1998; Berridge, 2007) are associated with mesolimbic and mesocortical dopaminergic functioning. In comparison, motor function is associated with nigrostriatal dopaminergic function (Jenkinson & Brown, 2011). The dopamine hypothesis of reward states that rewarding outcomes reinforce memory traces of behavioural responses (Wise, 2004), having a drive-like impact that amplifies the likelihood of an agent responding to the reward-linked stimuli. This refers to either internal tissue-related needs like hunger or external stimuli linked with previously experienced rewards or 'incentive motivational' stimuli (Gallistel et al., 1974; Wetzel, 1963; Wise, 2004). Importantly, Schultz and colleagues found phasic dopamine bursts increase as a function of increases to predicted value, decrease with unpredicted reductions to value, and more generally represent error signals between predicted and received outcomes (Hollerman & Schultz, 1998; Schultz, 1998; Schultz et al., 1998, 2000; Schultz & Dickinson, 2000; Tremblay et al., 1998; Tremblay & Schultz, 1999). The dopamine-related invigoration of ongoing behaviour may modulate the learning process about reward outcome tendencies, as prior work has shown that increased tonic dopamine in the striatum invigorates Pavlovian and instrumental behaviour (Ikemoto & Panksepp, 1999; Salamone & Correa, 2002).

Tonic dopamine is thought to change and develop slowly, over the course of minutes, linking behaviour over the trials or potentially the blocks of an experiment (Niv, 2009). Thus a

motivation such as reducing anxiety might drive a higher level of contextual neuromodulator that pushes both phasic PE responses and increases the residual average dopamine expressed by tonic dopamine levels (Niv, 2009). We cannot measure dopamine levels in this experiment, but coupling the anxiety manipulation to the reward-learning task may drive incentive-motivational value to the outcomes of the learning task that we can measure using behavioural responses; outcomes that are otherwise neutral in our control groups. Others, however, argue that potential modulation by motivation on task performance is more complex and that different rewards might have distinct outcomes on specific cognitive processes (Ivanov et al., 2012; Padmala & Pessoa, 2011; Savine & Braver, 2010). Despite not measuring dopamine levels in this study, the dopamine hypothesis is an important theoretical consideration to explain how extrinsic motivators might invigorate task behaviour. Our understanding of reward learning and motivation grew in the past decade. Yet the interaction between affective states (like anxiety), motivation, reward PEs, and predictions of predictability (i.e. precision) ~~driven by the amplitude of PEs~~ remain understudied.

Typically in studies on motivation, binary behavioural tasks may be operationalised to produce two effort-based choices: 1) to exert little effort for a small reward, or 2) more effort for a larger reward (Pessiglione et al., 2017). Using this approach, one can quantify subjective reward value. This has been established in experiments manipulating effort for obtaining food in rats (Salamone et al., 1994, 2003; Salamone, Yohn, et al., 2016; Walton et al., 2006) and in humans asked to actively eliminate stimuli for monetary rewards (Croxson et al., 2009; O'Doherty, 2016). These studies show that prior to a response, the reward processing dopaminergic midbrain represents both the expected reward and the required degree of effort to obtain it. Studies on humans typically utilise external rewards like money and points as they are commonly thought to engage identical neural reward processing regions as rewards like food (Levy & Glimcher, 2012; Pessiglione et al., 2017). To illustrate, during socially motivating circumstances, conducting a binary reward-learning task, Le Bouc and Pessiglione (2013) discovered motivation was chiefly fuelled by personal utility, expressed neurally by increased BOLD activity in the above mentioned dopaminergic reward processing regions of the brain.

For the current experiment, we focused on manipulating motivation and studying its effect on reward-based learning using the threat of a secondary stressor task, the Trier Social Stress Test (Birkett, 2011; Kirschbaum et al., 1993; Labuschagne et al., 2019). By these means, we tapped into the more complex experience of human anxiety (or, anticipatory worry) to operationalise motivation. The rationale was twofold. Firstly, that anxiety about performing a social stress test would produce motivation to perform well (to reduce anxiety) through an aversion to aversive experiences. Secondly, similar to physical effort, mental effort through

difficult executive tasks (like social performance) also incentivise avoidance motivation. This has been firmly established in decision-based tasks in healthy humans, who consistently favour avoiding difficult tasks involving executive functions (Apps et al., 2015; Kool et al., 2010; Shenhav et al., 2017; Westbrook et al., 2013). Moreover, it has been suggested that avoiding potential aversive experiences is equivalent to gaining rewards (Bach & Dayan, 2017; Peter Dayan, 2012a; Lloyd & Dayan, 2016; Maia, 2010; Mowrer, 1951). This suggests that reward learning performance in our task could be driven by a personal utility (motivation, producing greater effort) to avoid an upcoming stressful task, the TSST, that involves both a difficult public mental arithmetic test and a public oral presentation.

Little is known about how extrinsic motivation to avoid an anxiety-inducing event may influence learning performance. To obtain a proxy measure of changes in anxiety from our psychosocial stressor manipulation we administered self-reported assessments and recorded continuous electrocardiographic data. We then devised an experiment where anxiety was coupled to reward. We used a time discount (TD) function defined as a positive reward yielding a reduction to the amount of time needed to perform the upcoming TSST. As an example, if a participant correctly predicts the rewarding stimulus in a trial, they are rewarded with 3 seconds to subtract from the total time of the upcoming TSST. An unrewarded 'lose' trial was defined as 0 seconds subtracted from the TSST. Using this manipulation, we coupled anxiety reduction (motivation) to learning about rewards. For this experiment, we used a combination of the threat of socially assessed difficult mental arithmetic and an oral presentation about an unspecified artwork (like in our previous work Hein et al. [2021] and **Chapter 2**). Importantly, however, our study only threatens psychosocial stress tests, as opposed to tasks where the psychosocial stress test itself represents the independent and dependent variables (Allen et al., 2014; Boesch et al., 2014; Dedovic et al., 2005; Kudielka et al., 2004).

We found that participants did not experience physiological or subjective psychological responses to our TSST experimental anxiety procedure, unlike in our previous experiments (Hein et al., 2021; Sporn et al., 2020). Because of that, our anxiety induction procedure did not alter model-free responses (the percentage of overall errors while reward learning or reaction times) or model estimates of uncertainty ~~Bayesian predictive coding~~ from the Hierarchical Gaussian Filter (HGF) model (Mathys et al., 2011, 2014). We discuss in further detail the reasons contributing to the results in our discussion, offering solutions and direction for future studies looking to investigate anxiety and motivation.

2. Methods

2.1 Participants

Healthy volunteers (N = 45, 30 female, 15 male) with an age range between 19–39 (mean 26.3, standard error of the mean [SEM] 0.78) took part in this reward-learning task. The study was approved by the local ethics committee (Goldsmiths University of London ethical review board). The first 22 participants' data was collected by F.H, with the remainder by T.P.H. All participants provided informed written consent and all were healthy, reporting no past neurological or psychiatric condition. Our sample size was estimated using the behavioural and modelling data from our previous study (Hein et al., 2021) using MATLAB (MathWorks, 2012, The Math-Works, Inc., MA, USA: function *sampsizepwr*).

Each participant was pseudo-randomly designated into one of three experimental groups (N = 15 each) after measuring trait anxiety using Spielberger's Trait Anxiety Inventory (STAI, Spielberger, 1983b). These groups were a control (Cont), state anxiety with a time-discount (StA_{TD}), and a state anxiety group with no time discount ($StA_{TD\text{-}Cont}$)—for further details see the experimental manipulation **Section 2.3**. Crucially, by screening trait anxiety levels prior to assigning groups we assured equivalent average subclinical trait anxiety scores in each group (Cont, mean 40, SEM 1.6; StA_{TD}, mean 47, SEM 2.2; $StA_{TD\text{-}Cont}$, mean 41, SEM 2.2; anxiety disorders typically report values exceeding a score of 46, see Fisher & Durham, 1999). Moreover, two important confounding factors in measuring state anxiety are age and sex (Voss et al., 2015). Accordingly, each group were homogeneous in age (Cont, mean 28.7, SEM 1.4; StA_{TD}, mean 24, SEM 1.3; $StA_{TD\text{-}Cont}$, mean 26, SEM 1.2) and sex (Cont, male 7, female 8; StA_{TD}, male 6, female 9; $StA_{TD\text{-}Cont}$, male 4, female 11).

2.2 Protocol: reward-learning task

All groups (Cont, StA_{TD}, $StA_{TD\text{-}Cont}$) were instructed following completion of practice trials (but prior to the first experimental task block) that this experiment was composed of two parts (see **Supplementary Figure 1**). Each group was informed that the first half tests their performance on a reward-learning task, while the second assesses oral presentation and mental arithmetic skills. However, only the experimental state anxiety groups (StA_{TD}, $StA_{TD\text{-}Cont}$) were informed that the secondary tasks would be performed in front of an audience—aimed to induce anxiety (see next sections).

Participants were seated in an isolated room in front of a computer with their right hand resting on a keyboard. Each participant performed an experimental task across four blocks: an initial resting state block (R1: baseline), task block 1 of reward learning (TB1), task block 2 of reward learning (TB2), and a final resting state block (R2). The experimental reward-learning task was presented using MATLAB 2012a and custom code using Cogent 2000, and the design was adapted from similar tasks used previously (de Berker et al., 2016; Hein et al., 2021).

Each resting state block consisted of five minutes of sitting relaxed with eyes open focused on a fixation mark while we recorded electrocardiographic (ECG) responses continuously. Following the first resting state block, participants received information on the reward-learning task and performed 15 practice trials. All were informed that this experiment involved binary choices where the aim was to predict which of two images (blue, orange) was going to reward them in each trial. Participants were informed that the probability of reward assigned to the two images (that were reciprocally related: p,1-p) would change throughout the two task blocks, similar to the procedure implemented in prior studies (Behrens et al., 2007; de Berker et al., 2016; Iglesias et al., 2013).

The stimuli were displayed randomly to either the left or right of a central fixation mark in each trial (see **Figure 1A**). Stimuli were either removed from the display when the time allowed to make a prediction expired (2200 ms ± 200 ms) or when a response (left or right arrow key) was provided. On receipt of a response, the chosen stimulus was highlighted in bright green. The selection remained on screen for 1200 ms (±200 ms). After, the trial outcome (win, lose, no response) was shown in the centre of the screen for 1200 ms (±200 ms) in either green (win) or red (lose, no response). Each trial finished with a white fixation cross (inter-trial interval, 1250 ms ± 250 ms).

Critically, in contrast to previous work, the reward in this experiment was designated by a reduction to either points or the time required to perform the anxiety-inducing secondary tasks (public presentation and mental arithmetic). The state anxiety with time discount group (StA_{TD}) were rewarded for correct predictions with a reduction to the amount of time needed to perform the two secondary anxiety-inducing tasks. For StA_{TD}, each rewarded win was −3 seconds from those tasks, while an unrewarded lose and no response was 0 seconds deducted. The control group (Cont) experienced no anxiety and no time discount. And in contrast to StA_{TD}, the time discount control group ($StA_{TD\text{-}Cont}$) experienced anxiety but with no time discount. Both Cont and $StA_{TD\text{-}Cont}$ groups were rewarded for correct predictions by a reduction to a total 1200 '*corrupted*' points assigned at the start of the task (win, −3 corrupted points). An unrewarded

lose outcome or no response in Cont or $StA_{TD\text{-}Cont}$ represented the loss of an opportunity to decrease those corrupted points (0 corrupted points, **Figure 1A**).

Both TB1 and TB2 consisted of 200 trials. Across the total 400 trials, 10 probabilistic reward contingency mappings were randomly arranged for each participant. The possible contingency mappings for one image (for example, blue) were either highly biased (0.9/0.1), moderately biased (0.7/0.3), or unbiased (0.5/0.5) probabilities. These relationships were then reversed for the alternative image (orange: 0.1/0.9; 0.3/0.7, see de Berker et al., 2016). In addition, each contingency mapping was randomly assigned a duration between 34 and 46 trials (See **Figure 1B**).

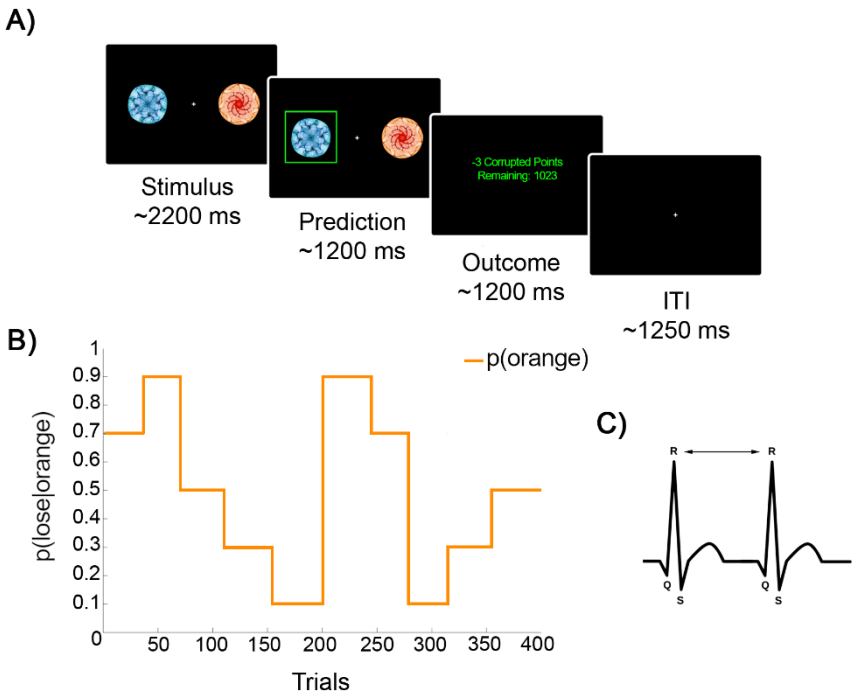


Figure 1. Reward-learning task structure and physiological measure. **A)** Each trial presented participants with visual icons (blue, orange) randomly to either the left or right of a central fixation cross for 2200 ms (± 200 ms). Once a response using either the left or right key was pressed, participants were immediately shown their choice highlighted in bright green for 1200 ms (±200 ms). Subsequently, the trial outcome was revealed. The rewarded win outcome displayed "*−3 p*" representing corrupted points for the Cont and $StA_{TD\text{-}Cont}$ groups, as shown in the above outcome screen in the colour green. The StA_{TD} group saw "*−3 s"* instead, representing seconds. Correspondingly, the unrewarded lose outcome displayed "*0 p/s*" while no response resulted in the message "*No response! 0 p/s*"—both in the colour red. Each outcome was displayed. Each trial ended with a fixation cross and an inter-trial interval of 1250 ms (±250 ms). **B)** The task took place over two task blocks (TB1, TB2, 200 trials each). The probabilistic structure of the task was randomly generated for each participant using the probabilities 0.9, 0.7, 0.5, 0.3, and 0.1, varying in length (34–46 trials). In the above example, the experiment starts with a 0.7 probability of orange rewarding (p(win|orange) with a reciprocal 0.3 probability for the blue stimulus: p(win|blue) = 1 − p(win|orange). After 36 trials the contingency mapping changes to a 0.9 probability of orange rewarding and a 0.1 probability of blue rewarding. **C)** During the experiment, we

continuously recorded ECG signals, using the R-peaks to calculate heart-rate variability (HRV) and to estimate the high-frequency spectral power (0.15–0.4 Hz) in HRV (HF-HRV).

Each element of the experiment (instructions, all questionnaires, and the reward-learning task) were conducted using the computer. Instructions were presented on screen and brief questions testing comprehension of the task were included to confirm understanding. For example, to test understanding of the reciprocal probabilities of the task, we asked: “If the probability of the blue image rewarding you is 70%, what is the probability of the orange image rewarding you?”. For the questionnaires, written instructions were provided on screen and responded to using the keyboard numbers.

Each group was instructed about the secondary oral presentation and mental arithmetic tasks. However, the Cont group were informed that the secondary task will be performed privately, and thus lacked the anxiety component. One final procedural difference between the experimental groups and control group is that we included 5-minutes of anticipatory waiting time just after being instructed on the secondary anxiety-inducing tasks but before the beginning of the first experimental block (see **Supplementary Figure 1**). The rationale was to strengthen the anxiety effect, in line with recent instructions on the practical implementation of the Trier Social Stress Test (see Birkett, 2011; Kirschbaum et al., 1993; Labuschagne et al., 2019). During this time, participants in StA_{TD} and $StA_{TD\text{-}Cont}$ were informed the experimenter was organising the three examiners for their second task.

All participants were paid £10 for their participation. In contrast to the recent implementation of similar tasks (de Berker et al., 2016; Hein et al., 2021), in this experiment, participants did not receive additional payment for correct predictions. The reason for this discrepancy was to isolate the motivational reward during the dynamic reward learning in the StA_{TD} group only. All participants were thus paid a base rate for conducting the experiment; but only the StA_{TD} group experienced the additional reward motivation of reducing an upcoming anxiety-inducing event. Participants were permitted to take a self-timed break between the two task blocks.

2.3 Experimental manipulation

The instructions concerning the additional arithmetic and oral presentation tasks that followed the reward-learning task were different between groups to induce state anxiety only during TB1 and TB2 in the two experimental groups (StA_{TD}, $StA_{TD\text{-}Cont}$). Following previous work, the two experimental groups were instructed that a random draw selected them for a secondary

public arithmetic test and oral presentation, known to induce anxiety (Cumming & Harris, 2001; Kirschbaum et al., 1993; Wolf et al., 2015). The public oral presentation consisted of presenting an abstract piece of artwork to three examiners for 10 minutes, after a 3-minute preparation time (Feldman et al., 2004; Lang et al., 2015). The public mental arithmetic task consisted of 10 minutes of timed subtracting from large numbers (for example, “What is 673−391?”). Participants would be asked these questions by the same panel of examiners, and responses to each question were to be given verbally within 5 seconds (Labuschagne et al., 2019). Critically, the state anxiety group with a time discount (StA_{TD}) could reduce the time they were required to perform these anxiety-inducing tasks by correctly predicting rewarded outcomes in the reward-learning task (a rewarded outcome, −3 seconds from the total 20 minutes). The StA_{TD} group were informed that the sum of their correctly predicted trials would be subtracted from the total 20 minutes performing the public arithmetic and oral presentation tasks—with the intention of manipulating motivation.

The $StA_{TD\text{-}Cont}$ group, however, experienced anxiety but with no time discount, reducing only the corrupted points they were assigned at the start of the experiment. Likewise, the Cont group reduced corrupted points but Cont importantly had no anxiety manipulation. Those in Cont were instead told they would need to perform both tasks privately to themselves. After finishing the reward-based learning task the threat of a social stress test was countermanded in both $StA_{TD\text{-}Cont}$ and StA_{TD}. For $StA_{TD\text{-}Cont}$ participants were instructed that the assessment panel was suddenly unavailable and that they were to ultimately describe only the artwork privately in line with the Cont group. StA_{TD} was ultimately informed that they, in fact, deducted sufficient time from the secondary arithmetic task to not perform it, but that they still needed to describe the artwork privately as the examiners were not present—aligning with the Cont and $StA_{TD\text{-}Cont}$ groups.

2.4 ECG Recording, pre-processing and heart rate variability analysis

Electrophysiological signals were recorded across each task block (R1, TB1, TB2, R2) using the BioSemi ActiveTwo system with a sampling rate of 512 Hz. Two electrodes were used to record the ECG and were placed in a two-lead configuration (Moody & Mark, 1982). Both electrodes were affixed using surgical tape and the signal was improved using highly conductive bacteriostatic Signa gel (by Parker Laboratories, Inc., 4 Sperry Road. Fairfield, NJ 07004 USA). Event markers recorded in the ECG BioSemi file denoted task blocks.

Analysis of the ECG data was performed using MATLAB (The MathWorks, Inc., MA, USA). First, in the EEGLAB toolbox (Delorme & Makeig, 2004), ECG data were preprocessed by

notch-filtering between 48-52 Hz (847 points) to remove power line noise. After, in the FieldTrip toolbox (Oostenveld et al., 2011) with their recommended function for identifying cardiac events,[10] we determined the QRS-complex and the R wave peak, deriving the latency of the R-peak to calculate the coefficient of variation (CV = standard deviation/mean) of the difference intervals between successive R-peaks (inter-beat interval: IBI). This CV of the IBI was one metric of heart rate variability (HRV) used as a proxy for physiological changes consistent with anxiety (Chalmers et al., 2014). The second physiological index of our anxiety manipulation was the spectral characteristics of the IBI time series. We interpolated the IBI time series at 1 Hz using a spline function (order 3), with spectral power estimated using Welch's periodogram method (Hanning window: see Rebollo et al., 2018). After, the extracted power estimates were normalised to the average power of R1 and converted to decibels (dB) for statistical analysis (see **Section 2.8** below).

2.5 Assessing state anxiety

To assess state anxiety, we analysed the continuous ECG signal and self-reports of anxiety (see below). Our primary indicator of anxiety was the CV of the IBI time series as a proxy of HRV, a metric that has consistently been reported to drop during anxious states (Chalmers et al., 2014). Prior work has established that anxious physiological responses, like stress responses, reduce HRV—linked with the reduction of complexity in physiological systems (Friedman & Thayer, 1998; Friedman, 2007; Goldberger et al., 2002; Gorman & Sloan, 2000). We also validated the use of this CV approach to HRV in our recent studies that showed lower HRV under anxiety using a similar experimental manipulation (Hein et al., 2021; Sporn et al., 2020). In addition to the HRV index, we also analysed the spectral profile of the IBI data to show lower high-frequency HRV (0.15–0.40 Hz) content. Reduced HF-HRV is a known marker across several anxious states (from high levels of trait anxiety, worry, and clinical anxiety disorders) that relates to parasympathetic vagal control of the heart and autonomic modulation (Aikins & Craske, 2010; Fuller, 1992; Miu et al., 2009; Mujica-Parodi et al., 2009; Pittig et al., 2013; Thayer et al., 1996). In our previous study in **Chapter 2**, we observed lower HRV with lower HF-HRV in response to state anxiety, validating the use of HF-HRV as an additional proxy index of a physiological change accordant to an anxious state.

---

10 http://www.fieldtriptoolbox.org/example/use_independent_component_analysis_ica_to_remove_ecg_artifacts

Complementing the HRV analysis we obtained dynamic estimates of anxiety levels using the Hospital Anxiety Depression scale (HAD, Zigmond & Snaith, 1983). Our motivation to use the HAD questionnaire was that the subscale for anxiety (HAD-A) consists of only seven questions, and is reported to perform well in assessing anxiety (Bjelland et al., 2002). We acquired a subset of two HAD-A responses recorded 13 times throughout the experiment using a coded version (participants were instructed to respond using the keyboard number buttons). By using a subset we could quickly and unobtrusively obtain a time-varying estimate of anxiety levels throughout the experiment (see **Supplementary Figure 1** for a visual presentation of the task structure and anxiety assessments). The two questions selected from the HAD-A were: i) I feel tense or 'wound up' ii) Worrying thoughts go through my mind. We also added two questions to serve as control responses: i) I feel tired ii) I feel bored. Our self-report questionnaire (hereafter termed HAD-S for 'Hospital Anxiety Depression Subset') thus consisted of 4 questions: 2 concerning anxiety and 2 control questions. Participants were instructed to respond with 1 indicating 'Not at all', 2 'Somewhat', and 3 'Very much so'.

The first HAD-S assessment was taken after R1 but before the practice trials (see **Supplementary Figure 1**). The second HAD-S was completed after the practice but prior to informing participants of the secondary tasks (experimental manipulation: public speaking and mental arithmetic). The average over the first and second HAD-S responses was used as a baseline to normalise HAD-S scores recorded throughout the experimental task blocks. As mentioned above, after the anxiety induction, the StA_{TD} and $StA_{TD\text{-}Cont}$ groups waited for 5 minutes to strengthen the anxiety manipulation following recent instructions (see Birkett, 2011; Kirschbaum et al., 1993; Labuschagne et al., 2019). During this time, the StA_{TD} and $StA_{TD\text{-}Cont}$ participants filled out 3 HAD-S scales separated by 1.5 minutes. After, all groups started the reward-learning blocks TB1 and TB2 where HAD-S reports were taken every 40 trials across the total 400, making 10 HAD-S measures in TB1 and TB2. After completing TB2, one final HAD-S was completed (see **Supplementary Figure 1**).

2.6 Computational model

In this paper, we assess how a motivation to reduce anxiety alters model estimates of belief updates about probabilistic rewards in a volatile environment. All probabilistic models use explicit assumptions concerning the generation of observations and approximation strategies on inference; with recent models using variational approaches that break difficult inference problems into smaller more pliant ones (Mathys et al., 2011, 2014; Piray & Daw, 2020a). We use the Hierarchical Gaussian Filter (HGF, version 6.0.0), a generic hierarchical Bayesian model of learning under uncertainty that affords inference on a learner's beliefs about the

learning environment using their responses (Mathys et al., 2011, 2014). Recently, it has seen widespread application in clinical (Cole et al., 2020; Deserno et al., 2020; Reed et al., 2020), pharmacological manipulation (Marshall et al., 2016; Vossel et al., 2014; Weber et al., 2020), and non-clinical settings (Diaconescu et al., 2014; Iglesias et al., 2013; Lawson et al., 2020; Palmer et al., 2019; Weilnhammer et al., 2018).

The model adheres to Bayesian brain theories (Doya et al., 2007; Friston, 2010) which assert that the brain approximates a statistically optimum generative model of the world, emphasising the use of uncertainty for updating a hierarchy of beliefs through the sequential use of pwPE signals (Diaconescu et al., 2014). The task used here is identical to **Chapter 2**; we used a 3-level HGF perceptual model combined with two softmax decision models operating the mapping between belief and response. One with a fixed decision noise parameter ζ (for further detail on the update equations, we refer the reader to **Eq. 21** in **Chapter 2** and the original paper by Mathys et al. [2011]). The other with a decision temperature that is the exponential of log-volatility (i.e., the inverse decision temperature (β) is the exponential of negative log-volatility, $\beta = e^{-\mu_3}$, termed HGFμ_3), which may vary in each trial with estimated volatility (Diaconescu et al., 2014).

In the HGFμ_3, ζ depends dynamically on the mapping between estimated beliefs about the volatility of the probabilistic reward environment (see Diaconescu et al., 2014). As trial-by-trial estimates of volatility increase, the sigmoid function decreases, leading to more decision noise (Cole et al., 2020; Diaconescu et al., 2014; Reed et al., 2020). Increased volatility thus leads to exploratory or 'noisier' response behaviour. By contrast, if the agent estimates lower environmental volatility, the sigmoid function becomes steeper, and their response choices will become less noisy and correspond more to their beliefs (Diaconescu et al., 2014).

Here as in **Chapter 2,** subject-specific parameters determine how the states of the task (stimulus outcome tendency, volatility) evolve in time. We set the phasic volatility parameter κ to 1 (as the scale of x_3 is arbitrary) and estimated ω_2 ω_3 and ζ. A further four free parameters control the initial values of an agent's beliefs at the beginning of the task ($\mu_2^{(0)}$, $\sigma_2^{(0)}$, $\mu_3^{(0)}$, $\sigma_3^{(0)}$)—for priors used on each HGF model used here, see **Supplementary Materials: Table 1**).

We fitted the HGF models to the trial-wise responses of all participants using Bayesian model inversion. Maximum-a-posteriori (MAP) estimates of model parameters were calculated with the priors on parameters (see **Supplementary Materials: Table 1**) and the sequence of

inputs, optimised with the quasi-Newton optimisation algorithm (Cole et al., 2020; Diaconescu et al., 2014; Reed et al., 2020).

2.7 Model selection

Our model space includes the binary 3-level HGF (HGF_3) featuring volatility estimates and a set decision noise parameter that maps beliefs to decisions (Mathys et al., 2011). Our additional 3-level HGF model ($HGF_{\mu 3}$) as detailed above is one where decisions depend dynamically on the estimated volatility of the probabilistic reward environment (Diaconescu et al., 2014). We also included an alternative HGF with two levels (HGF_2) that fixes volatility. As in previous research, we also included simpler reinforcement learning models (Cole et al., 2020; de Berker et al., 2016; Lawson et al., 2020). First, a Rescorla-Wagner (RW) model, which describes learners as acting to maximise the probability of upcoming rewards (Rescorla & Wagner, 1972). The RW does not quantify uncertainty or use an adaptive learning rate, but iteratively learns the values of outcomes using PEs to drive associative changes directly, with the sign and magnitude of the error determining the associative strength. And secondly, a Sutton K1 model (SK1), that can change the learning rate based on recent PEs (Sutton, 1992).

We extracted the log-model evidence (LME) for all participants (N = 45) and compared two families of models: a Bayesian family (HGF_2, HGF_3, $HGF_{\mu 3}$) and a reinforcement learning family (RW, SK1). We used the LME to compare each family of models by using random effects Bayesian model selection (BMS, see Stephan et al., 2009), utilising code from the MACS toolbox (Soch & Allefeld, 2018) as performed in **Chapter 2** and Hein et al. (2021). After, we ran BMS on the winning family of models to determine which model best explains participants' learning behaviour.

2.8 Statistical analysis

Statistical tests were applied to the listed dependent variables: *i)* behavioural indices (model-free measures, the percentage of errors and reaction time, RT—averaged in each task block) *ii)* Block averages of our HRV metric using the CV, normalised by subtracting the R1-baseline mean from each task block (TB1, TB2) and the spectral content of the IBI time series HF-HRV, baselined using an identical method. *iii)* the time-varying index of HAD-S self-reported anxiety normalised to a baseline level. *iv)* The HGF model estimates of the perceptual model parameters: ω_2 and ω_3 *v)* HGF quantities *a)* estimation (*informational*) uncertainty about the reward tendency x_2 (σ_2) *b)* estimates of belief on volatility (mean, μ_3, and variance, σ_3) *c)* environmental uncertainty: $\exp(\kappa\mu_3 + \omega_2)$.

To assess the main effects and interactions between our Group factor (StA_{TD}, $StA_{TD\text{-}Cont}$, Cont) and our Block factor (TB1, TB2) we used two-way repeated-measures analysis of variance (ANOVA) tests. The HGF parameters are normally distributed in the space in which they are estimated (for example, log-μ_3, ω_2, ω_3), satisfying the ANOVA assumption of normality. Where the factor Group or Interaction term is significant, factorial analyses are complemented with pair-wise comparisons using ANOVA. We used 3 x 2 ANOVAs instead of our preferred non-parametric factorial tests based on synchronised permutations because freely available code for N x 2 tests with N > 2 is at present unavailable. Note, however, that mathematical descriptions of the extension of 2 x 2 synchronised permutations to 2 X N or M x N more broadly exist (Anderson & Braak, 2003; Basso et al., 2007; Salmaso, 2003). Post-hoc pair-wise tests were carried out using permutation tests.

Where null results are found, we use Bayesian ANOVA using JASP (JASP Team, 2019) to test whether there is more evidence for the null hypothesis compared to the alternative hypothesis. This test yields a Bayes factor (BF): the relative evidence in the data (posterior odds) favouring one hypothesis among two competing hypotheses given equal priors (Kass & Raftery, 1995). BF_{10} is the standard BF for quantifying evidence for the alternative hypothesis (H_1) over H_0, written p(data|H_1) over p(data|H_0) (van Doorn et al., 2021). Here we use BF_{01}, which is typically used when quantifying the evidence in favour of H_0 over H_1, written p(data|H_0) over p(data|H_1). BF_{01} > 1 expresses evidence supporting H_0 to a different degree depending on the actual value of the BF (Kass & Raftery, 1995; Lavine & Schervish, 1999). For example, a BF_{01} = 4 can be interpreted as the data are 4 times more likely under H_0 than under H_1 (van den Bergh et al., 2020), with BF_{01} = 3–10 sometimes referred to as moderate or substantial evidence for H_0 (Wagenmakers et al., 2011). Alternatively, BF_{01} = 1–3 is anecdotal, BF_{01} = 10–30 is strong, BF_{01} = 30–100 is very strong, and BF_{01} > 100 is extreme evidence for H_0 (Wagenmakers et al., 2011).

After identifying the model that best explains response behaviour using BMS, we selected the trial by trial trajectories for each HGF model estimate being analysed (σ_2, σ_3, μ_3, and environmental uncertainty) and averaged across trials within each block (TB1, TB2). In doing so, we sought to evaluate the overall group-and block-related monotonic changes using the above mentioned 3 x 2 factorial analysis with the factors Group and Block, respectively. Planned pairwise comparisons were used to test differences between groups in the time-varying HAD-S measurements, and thus assess time-varying changes in anxiety. This was carried out using one-way ANOVA (10 independent tests). To control for multiple comparisons, the false discovery rate (FDR) was controlled for by applying an adaptive linear step-up

procedure set to a level of q = 0.05 yielding an adapted threshold p-value (P_{FDR}, Benjamini et al., 2006). We present the mean and standard error of the mean (SEM) as descriptive statistics for dependent variables. For pair-wise comparisons using permutation tests we provide non-parametric effect sizes and bootstrapped confidence intervals (Grissom & Kim, 2012; Ruscio & Mullen, 2012). As done in **Chapter 2**, we calculate the non-parametric effect sizes for within and between-group comparisons using the probability of superiority for dependent samples (Δ_{dep}) and probability of superiority (Δ), respectively.

3. Results

3.1 Hospital Anxiety Depression Scale

We first tested the averaged normalised HAD-S scores over task blocks (TB1, TB2) using a 3 x 2 repeated-measures ANOVA with Group the between-subject factor (Cont, StA_{TD}, $StA_{TD-Cont}$). The main effect of Group ($F(2,42) = 0.01$, $P = 0.99$) and the interaction effect was not significant ($F(2,42) = 0.82$, $P = 0.44$). There was a significant main effect of Block ($F(2,42) = 11.01$, $P = 0.002$, see **Figure 2A**). After, we tested planned pairwise comparisons between groups at each of the 10 consecutive HAD-S measurements independently using an ANOVA. This approach yielded no differences that survived FDR correction at any sampling point ($P_{FDR} > 0.05$, **Figure 2A**). To assess whether HAD-S responses were equivalent across groups, we ran a Bayesian ANOVA to test if there is more evidence for the null hypothesis (H_0: no effect of the factor Group). We found that the data are 10 times more likely under H_0 (strong evidence) compared with H_1 (the model with Group as a predictor, $BF_{01} = 10.04$), aiding in our interpretation that the data are diagnostic of supporting the null hypothesis over the alternative model hypotheses.

3.2 Heart-rate variability

The effect of Group and Block on the normalised HRV index was tested using repeated-measures ANOVA. The main effect of Group ($F(2,42) = 0.17$, $P = 0.85$) and Block was not significant ($F(2,42) = 2.67$, $P = 0.11$). The interaction effect was not significant ($F(2,42) = 0.57$, $P = 0.56$, see **Figure 2B**). An analysis of the normalised high-frequency content of the HRV (HF-HRV, 0.15 – 0.4 Hz) using a 3 x 2 ANOVA test revealed the interaction ($F(2,42) = 0.74$, $P = 0.48$) and main effect of Group ($F(2,42) = 1.9$, $P = 0.16$) and Block was not significant ($F(2,42) = 0.93$, $P = 0.34$, see **Figure 2C**). Subsequently, we tested if there was more evidence for H_0, that HRV and HF-HRV are equivalent across groups using a Bayesian ANOVA. The

data was around 9 times more likely under H_0 when compared with the model with the Group predictor (BF_{01} = 8.8), providing moderate evidence for H_0.

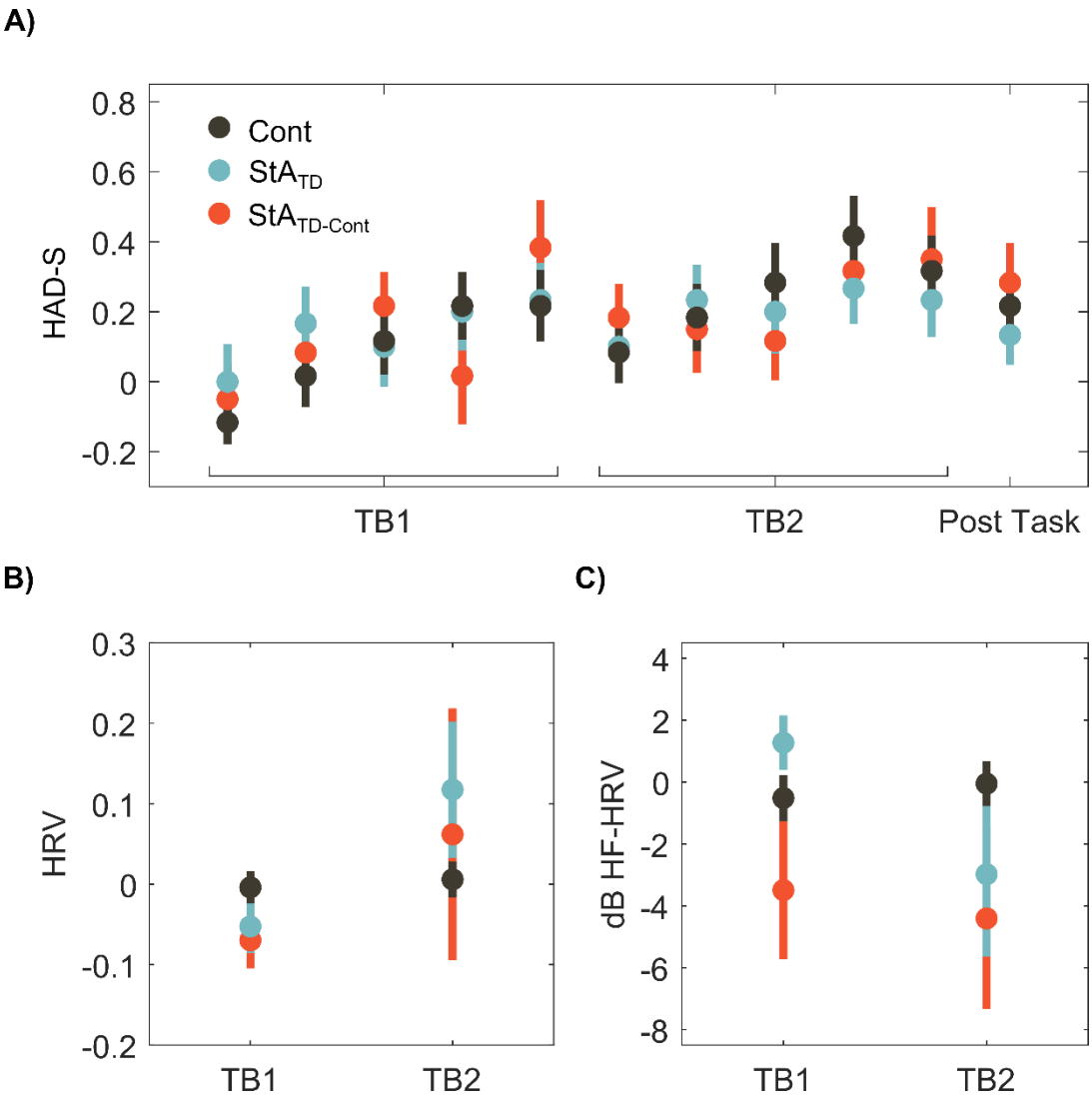


Figure 2. Anxiety assessments. **A)** Normalised Hospital Anxiety and Depression (HAD) scores using our subset HAD-S between groups (StA_{TD} [light blue], $StA_{TD\text{-}Cont}$ [orange], Cont [ink blue]) presented across our reward-learning task block 1 (TB1) and task block 2 (TB2). The final point named 'Post Task' was one final HAD-S assessment administered after completing the reward-learning task. The results showed no significant difference among the three groups **B)** Changes to our heart-rate variability (HRV) proxy measure of anxiety during the anxiety manipulation and reward-learning task. The normalised average HRV (coefficient of variation of the inter-beat-interval [IBI] of the ECG signal) is shown for each group (StA_{TD}, $StA_{TD\text{-}Cont}$, Cont) across TB1 and TB2. The normalised average was calculated by subtracting the resting state (R1: baseline) from each task block. There were no significant differences between the groups. **C)** Spectral analysis of the IBI time series data in high frequency (0.15 – 0.4 Hz) HRV range (HF-HRV). No tests were significant.

The above findings indicate that our combination of anxiety-inducing procedures, the threat of a TSST, did not induce anxiety. No changes were found in self-reported HAD-S scores or modulation to our two proxy measures of anxiety (HRV, HF-HRV)—as had been previously

reported (see **Chapter 2** and Chalmers et al., 2014; Feldman et al., 2004; Hein et al., 2021; Miu et al., 2009; Sporn et al., 2020).

3.3 Behavioural analysis

We used a repeated measures ANOVA to test the difference in the overall percentage of errors and found no significant main effect of Group ($F(2,42) = 0.94$, $P = 0.4$), Block ($F(2,42) = 0.66$, $P = 0.42$) or Interaction effect ($F(2,42) = 0.01$, $P = 0.99$, see **Figure 3A**). Afterwards, we tested whether there was more evidence for H_0 that reward learning performance is equivalent across all groups using a Bayesian ANOVA. We found there was only moderate evidence supporting H_0 ($BF_{01} = 3.5$).

Next, we used repeated measures ANOVA to test reaction times (RT, in milliseconds; averaged over all trials) and found the main effect of Group ($F(2,42) = 2.27$, $P = 0.12$) and Interaction term not significant ($F(2,42) = 2.27$, $P = 0.12$)—corresponding to our results in **Chapter 2** and prior work (Bishop, 2009). However, the Block factor was significant ($F(2,42) = 19.6$, $P < 0.01$, see **Figure 3C**). A follow-up Bayesian ANOVA revealed only moderate evidence in support of the null hypothesis of no effect of the Group factor ($BF_{01} = 3$).

Using the RT value from the current trial, we also compared the total-population (Cont, StA_{TD}, $StA_{TD\text{-}Cont}$) mean RT between losing and winning using non-parametric permutation tests. As found in **Chapter 2** and Hein et al. (2021), losing (mean 677.9, SEM 2.91) relative to winning (mean 645.5, SEM 2.24) on a trial was associated with significantly slower RT ($P = 0$, $\Delta = 0.87$, CI = [0.82, 0.91], **Figure 3C**). In addition, there were no significant differences found in RT between the three groups when comparing RT separately for either predictable (0.9–0.1 probabilistic relationship) or unpredictable contingency phases (0.5–0.5 probabilistic relationship, $P > 0.05$). This result confirms that there was no deficit in attention as shown in classical attention paradigms where attentional shifts lead to larger RTs to both predictive and uninformative cues (Prinzmetal et al., 2009).

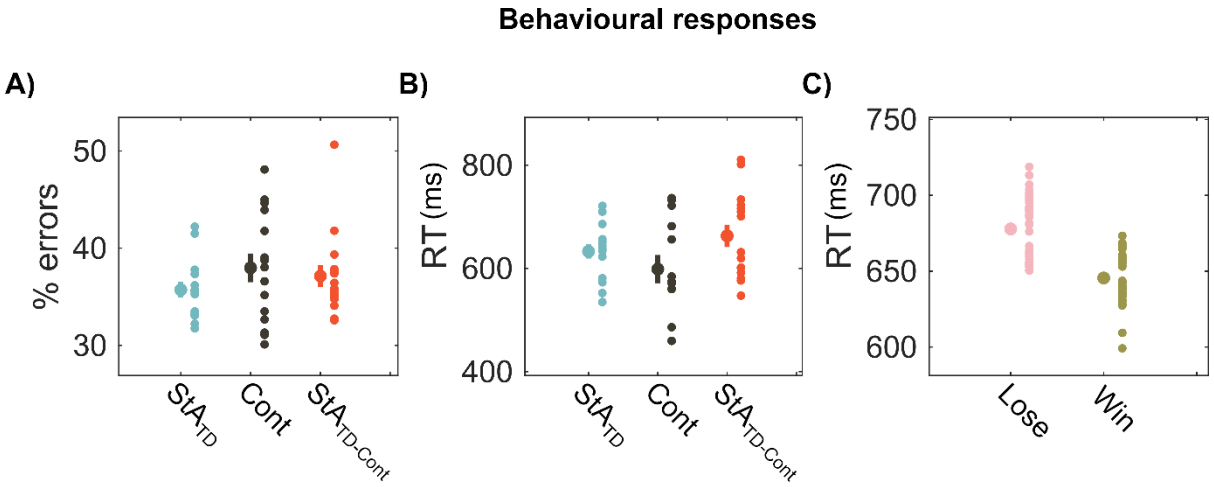


Figure 3. Model-free behavioural measures. **A)** The percentage of errors during reward-learning task blocks (TB1, TB2) between the $StA_{TD\text{-}Cont}$ (orange), StA_{TD} (blue), and Cont (ink blue) groups. The left dot is the group mean with SEM bars. On the right are individual scores comprising each group to present population dispersion. There were no significant differences between groups **B)** The average reaction time (RT) in milliseconds (ms) between groups is given by a coloured dot with SEM bars. On the right is respective group dispersion. No significant differences in the factor Group or Interaction were found. The Block factor was significant, with TB1 (mean 703.1, SEM 23.84) slower than TB2 (mean 646.1, SEM 24.41) **C)** RT presented by the trial outcome, with lose trials (pink) and win trials (green). The left is the group mean with SEM bars; the right, individual scores. Losing was associated with significantly RT relative to winning ($P = 0$).

3.4 Bayesian model selection

The family of Bayesian models (HGF_3, HGF_2, and $HGF_{\mu 3}$) was found to have stronger evidence than the reinforcement-learning models (RW, SK1)—determined by an exceedance probability of 1, and an expected frequency of 0.91 (**Figure 4A**). Subsequently, BMS within each Bayesian model (HGF_3, HGF_2, and $HGF_{\mu 3}$) indicated much stronger evidence for the $HGF_{\mu 3}$ model when compared with the HGF_3 and HGF_2 models (exceedance probability of 0.99 and an expected frequency of 0.67, see **Figure 4A**). As an additional check, we confirmed that the $HGF_{\mu 3}$ model was also the Bayesian model with the highest exceedance probability and expected frequency when performing BMS separately in each group (Cont, StA_{TD}, $StA_{TD\text{-}Cont}$).

Despite our previous experiment using a closely related task **Chapter 2** (Hein et al., 2021) finding the HGF_3 outperforms the $HGF_{\mu 3}$ version of the model, here we discovered that the $HGF_{\mu 3}$ as described in Diaconescu et al. (2014) yields stronger model evidence. This indicates that in this study the decision noise parameter is modulated by the trial-wise dynamics of μ_3.

3.5 Model-based results

By fitting the HGF, perceptual parameters for each participant were estimated that characterise their distinct learning style. In line with our previous work, we tested for differences between groups in the tonic volatility perceptual parameter ω_2 that contributes to the learning rate independently of the time-varying volatility estimate (Hein et al., 2021).

A one-way ANOVA was used to test the average ω_2 between the Cont (mean −2.3, SEM 0.68), StA_{TD} (mean −1.6, SEM 0.26), and $StA_{TD\text{-}Cont}$ (mean −2.1, SEM 0.50) groups. There was no significant effect on the Group factor ($F(2,42) = 2.2$, $P = 0.1$). Differences in the parameter ω_2 drive changes to estimates such as the learning rate on the lowest level (α), increasing or decreasing learning about the reward outcomes; however here, we observed no change in ω_2 between groups. We additionally tested the perceptual model parameter ω_3 finding no significant effect ($F(2,42) = 0.4$, $P = 0.7$) between the Cont (mean −7.0, SEM 0.24), StA_{TD} (mean −7.1, SEM 0.31), and $StA_{TD\text{-}Cont}$ (mean −7.1, SEM 0.09) groups

3.5.1 Posterior belief on environmental volatility

Using a 3 x 2 repeated measures ANOVA when testing the average posterior belief on environmental volatility (μ_3) revealed the Group factor was not significant ($F(2,42) = 1.61$, $P = 0.21$). Also, the main effect of Block ($F(2,42) = 0.57$, $P = 0.45$) and interaction was not significant ($F(2,42) = 0.61$, $P = 0.55$). Afterwards, a Bayesian ANOVA revealed only anecdotal evidence ($BF_{01} = 1.5$) for the null hypothesis that the average posterior belief on environmental volatility is equivalent across all groups.

3.5.2 Informational uncertainty about the reward tendency

Factorial tests using repeated measures ANOVA on estimation (belief) uncertainty about the outcome tendency (σ_2) revealed no significant main effect of Group ($F(2,42) = 0.09$, $P = 0.91$) or interaction effect ($F(2,42) = 0.26$, $P = 0.77$) during reward-based learning blocks. However, there was a significant effect in the factor Block ($F(2,42) = 5.1$, $P = 0.03$, **Figure 4B**). Similar levels of belief uncertainty about the reward tendency in $StA_{TD\text{-}Cont}$, StA_{TD}, and Cont groups suggests that new information has an equivalent impact on the update equations for beliefs about x_2. These results confirm a lack of difference between groups in the perceptual tonic volatility parameter ω_2, as belief uncertainty about the reward tendency tends to increase with higher ω_2. To empirically confirm the average σ_2 data are best explained by the null model,

we used Bayesian ANOVA. We found moderate evidence that the data are more likely under H_0 relative to H_1 (the model with the Group predictor, BF_{01} = 8.69).

3.5.3 Environmental uncertainty

We then asked whether the group factor modulates estimates of environmental uncertainty. We found no main effect of Group (F(2,42) = 0.38, P = 0.68) or Block (F(2,42) = 2.67, P = 0.11). The interaction effect was not significant (F(2,42) = 0.76, P = 0.48, see **Figure 4C**). A follow up Bayesian ANOVA showed moderate evidence for H_0 (that environmental uncertainty is equivalent across groups) relative to H_1 (BF_{01} = 7.05), with the data 7 times more likely under H_0.

3.5.4 Uncertainty about volatility

The main effect of Group (F(2,42) = 0.12, P = 0.89) and Block was not significant (F(2,42) = 2.67, P = 0.11) in the average belief uncertainty about volatility (σ_3) using 3 x 2 repeated measures ANOVA. We also found no interaction effect (F(2,42) = 0.59, P = 0.56, see **Figure 4D**). Subsequently, a Bayesian ANOVA showed moderate evidence for H_0 compared with the model with the Group predictor (H_1, BF_{01} = 8.4).

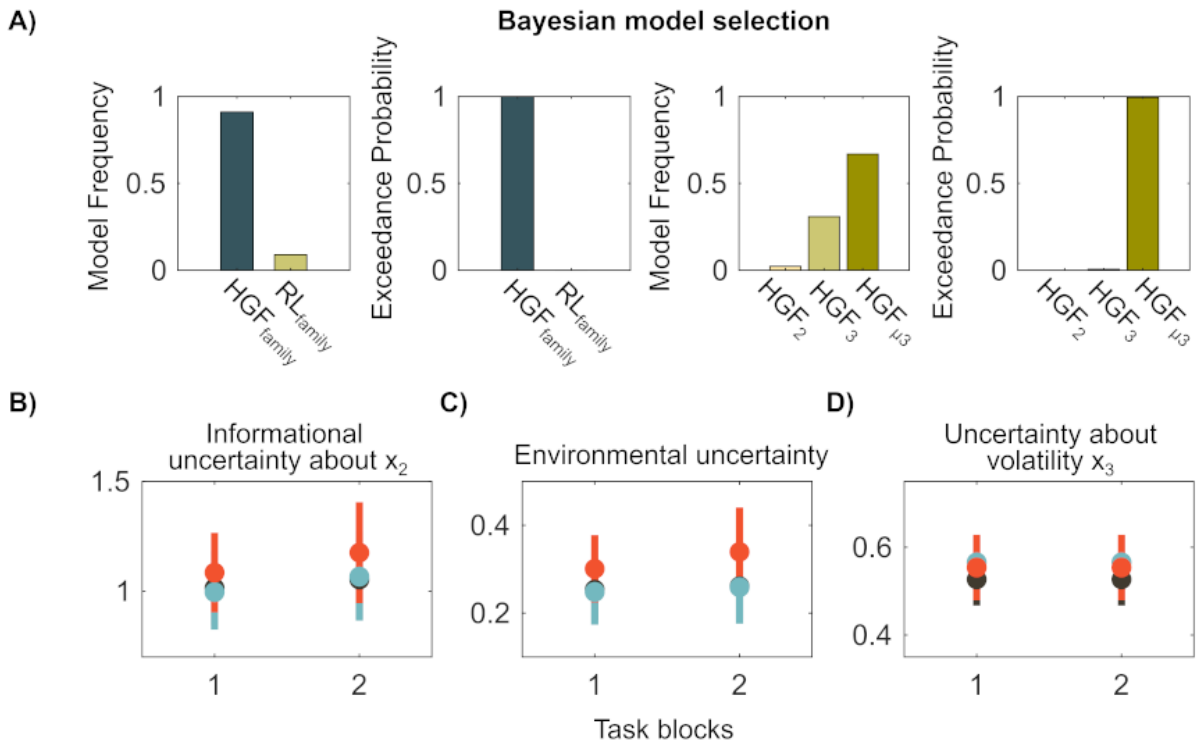


Figure 4. Bayesian model selection and model-based results. **A)** Our Bayesian model selection (BMS) procedure consisted of two stages: the comparison between the two families of models (two leftmost columns) and the comparison between HGF models (two rightmost columns). The family level BMS provided the model frequency and exceedance probability for each family of models: the HGF Bayesian models 'HGF family' (HGF_2, HGF_3 and $HGF_{\mu 3}$: dark blue) and the family of reinforcement learning models 'RL family' comprised of the Rescorla-Wagner (RW) and Sutton-Barto (SK1), represented by light green. The family of HGF models provided the best model evidence. In the two right panels is the comparison between the three HGF models (Two-level HGF [HGF_2: light green], 3-level HGF [HGF_3: green] and the HGF with decision parameter informed by estimates of volatility [$HGF_{\mu 3}$: dark green]). The $HGF_{\mu 3}$ provided stronger model evidence. **B)** Similar levels of estimation (informational) uncertainty about x_2 was found between the Cont (orange), $StA_{TD\text{-}Cont}$ (ink blue), and StA_{TD} (light blue) groups. A significant effect of the Block factor showed that on average σ_2 increased from TB1 to TB2. **C)** Environmental uncertainty, and **D)** uncertainty about volatility were not significantly different between the groups. No interaction effects or Group comparisons were significant. These results indicate that our anxiety manipulation did not significantly modulate uncertainty about the reward tendency, volatility, or environmental uncertainty.

4. Discussion

By combining a volatile probabilistic reward-based learning task and coupling it to an anxiety manipulation, we took a novel approach to test the motivational effect of anxiety on learning and decision-making. The key insight from this approach is that using the threat of a psychosocial stress test to induce anxiety in the lab is a nuanced and challenging procedure. We were unable to replicate the conditions that produced anxiety in **Chapter 2** (Hein et al., 2021) and in Sporn et al. (2020). Here, the threat of performing both a public speaking task and a difficult mental arithmetic task did not induce changes in either self-reported state anxiety levels or electrophysiological responses from the heart. By using both a model-free and behavioural modelling approach, we stated clearly and in detail the outcome on overall reward-learning performance and model estimates of uncertainty in three groups: a state anxious group, a motivated state anxious group, and a control group. Using factorial testing, we provided evidence that there were no significant changes to the percentage of errors, the reaction times, or the HGF model estimates. Bayesian ANOVA provided strong evidence for H_0 (no effect of factor Group) on self-reported (HAD-S) assessments, and moderate evidence for H_0 from HRV and model estimates of estimation uncertainty about the stimulus outcome tendency, volatility uncertainty, and environmental uncertainty. Bayesian ANOVA, however, only provided inconclusive anecdotal evidence of H_0 for HGF estimates about volatility and model-free results (error rates, RT). In what remains, we provide further discussion on the results and limitations reported here in relation to belief updating and our previous chapters, and also present potential future work that may reconcile these observations.

The threat of a psychosocial stressor such as the TSST is a potent means to induce anxiety. Our results show using the TSST to induce anxiety is an intricate and detailed process. There is inherently more variation in administering this experimental manipulation when compared to more replicable manipulations like shock. Certain individuals may be less sensitive ~~more invulnerable~~ to social stress, and factors such as experimenter performance and participant scepticism could dampen anxiety levels. A limitation of this study, therefore, is that the experimenter delivering the anxiety manipulation may be instrumental in inducing the effects. The threat of TSST manipulation worked well in **Chapter 2**, but not in the current experiment. Here, both self-reported anxiety and heart rate variability were equivalent based on BF_{01} among the three groups—which implies the experimenter did not successfully provide the cover story necessary to justify and induce anxiety.

A more consistent and substantial record of results has been achieved in work using the threat of shock (Cornwell et al., 2017; Grillon et al., 2019; Robinson et al., 2013, 2019). However, there is a lack of computational studies using the threat of shock focused solely on rewarded outcomes and volatile environments. In a study conducted by Bublatzky and colleagues (2017) this was partially addressed, but specifically concerning the competition between rewarded and aversive (shock) outcomes where environmental contingencies did not change over time. That study showed that threat interfered with reward seeking behaviours—a striking finding associated with how anxiety may impede safety learning. However, no studies to our knowledge have used the threat of shock in a volatile reward-learning task as used throughout this thesis. We did not observe differences between our experimental and control groups; a shock based approach may help future studies further understand the impact of state anxiety on reward learning in volatile and uncertain conditions. However, the differences between these two lab-based approaches (shock and social stress) and their effectiveness in inducing anxiety is still a matter of ongoing discussion (Grillon et al., 2019). More work also needs to be done investigating the impact of the anticipation of psychosocial stressors on the neural responses of prefrontal and limbic pathways, the enteric nervous system, and other physiological markers such as gastrointestinal functions (Bhatia & Tandon, 2005; Simpson et al., 2021). A further direction for future work could also explore the role of intolerance of uncertainty and safety learning in anxiety using both the threat of a psychosocial stress test and an aversive task, as the loss of reward feedback may be a critical factor in treating anxiety.

A further rationale to use the threat of shock comes from the subtlety of motivation and its potential effect on executive tasks like probabilistic decision making. The effect of motivational incentives on higher-order executive tasks, particularly those that demand attentional resources, are thought weaker than tasks involving physical effort (Schmidt et al., 2012).

Unlike the diffuse and delayed threat of a psychosocial stressor, which may diminish the motivational effect of anxiety, the threat of shock is unpredictable and immediate and is frequently used for within-subject studies (Roxburgh et al., 2020). Moreover, the uncertainty and delay of the future psychosocial stress event may discount the current task goal value, which is the effort taken to perform well during reward learning (Frederick et al., 2002; Green & Myerson, 2004; Pessiglione et al., 2017).

Temporal discounting is a further bias that may cause issues with diffuse threats in the future. The temporal discounting bias describes how we assign a lower value to gains in the future when compared to gains in the present (Berns et al., 2007; Robinson et al., 2015). Thus the potency of motivation from an anxiety-inducing task threatened in the future may be attenuated, and the context biases the effects on reward learning. In fact, others have argued that some higher-level executive tasks (including financial decision-making) may remain entirely unaffected by temporary changes to mood, stress, and anxiety and motivation (Robinson et al., 2015). Our work in **Chapter 2** showing state anxiety shapes reward-learning behaviour has since provided evidence against this argument. However, our modelling of electrophysiological responses and behavioural data suggest it may be informative to drill deeper into the subtle effects of motivation from anxiety by augmenting the experimental manipulation using shock in the lab. One potential design to address temporal discounting using threat of shock is using a within-group design, with an isolated reward learning performance coupled to a subsequent block of unpredictable shock, where rewards lead to reducing the probability of receiving shock. This is particularly true as we could only provide anecdotal evidence from Bayesian analysis that there was no effect of the factor Group on the overall percentage of errors made during reward learning. The nature of these potential motivational effects can then be resolved by future modelling work.

One striking result from this experiment is that unlike in **Chapter 2**, the $HGF_{\mu 3}$ response model—where participants update their beliefs using time-varying estimates of volatility—outperformed the 3-level model where the decision noise parameter is informed by a fixed mathematical description of volatility and its variance (Diaconescu et al., 2014; Mathys et al., 2011). The $HGF_{\mu 3}$ model describes responses where the choice probability depends on dynamic belief estimates of volatility, with more estimated volatility producing noisier and more exploratory responses. By contrast, lower estimates of volatility lead to a less noisy mapping between beliefs and responses, and thus a tighter coupling between belief and response (Cole et al., 2020; Diaconescu et al., 2014; Reed et al., 2020). However, it remains unresolved whether the groups equivalently estimated volatility, as factorial Bayesian analysis revealed only anecdotal evidence in support of the null hypothesis.

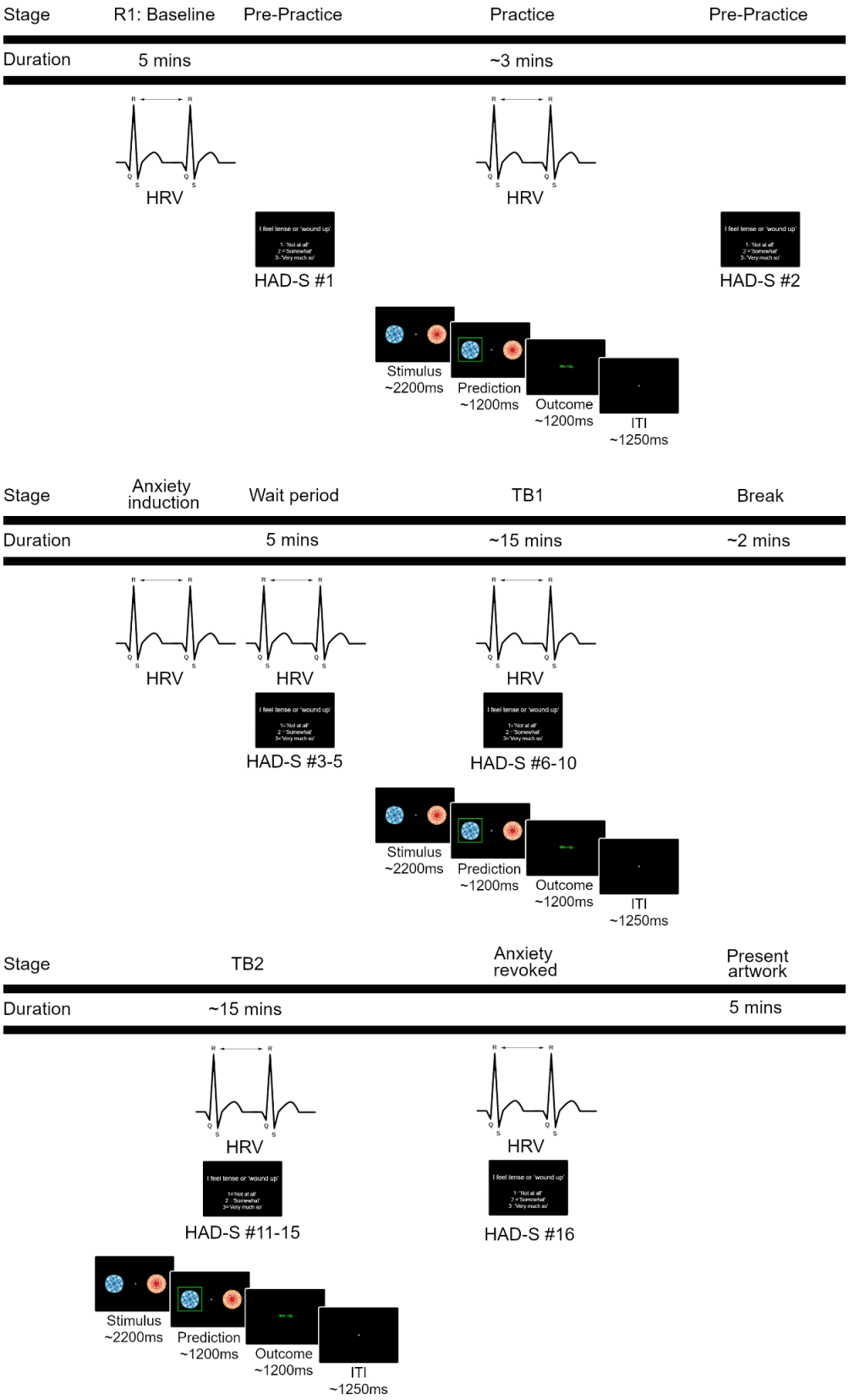
This result suggests that the method learners use to infer the changing statistics of this particular binary reward-learning task is not set. Our previous reasoning in **Chapter 2** (Hein et al., 2021) for the 3-level HGF model best describing behaviour responses was that our task has a set rate of change to probabilistic contingencies. Accordingly, participant decisions are influenced less by the estimated dynamics of volatility (where the ground truth changes are effectively constant). However, here we find that the trial-wise subjectively estimated changes to the contingency blocks inform participants' decisions. The potential inconsistency here between a ground truth constant volatility and the $HGF_{\mu 3}$ model being a better account of the data may be explained by recent work arguing that unexpected feedback from outcome variance (or 'unpredictability') is difficult to distinguish from true changes in the environment (Piray & Daw, 2020b). We speculate that some learners may expect more change in the environment, misattributing violations to predictions about the stimulus outcomes on the second level as changes to the underlying contingency. In these learners, we might expect the dynamics of trial-wise estimates of volatility to inform decisions (even if volatility is relatively stable, like in our task). This may indicate that participants are more tuned in to these changes, predominantly estimating that unpredicted outcomes due to outcome noise are informative of environmental change.

Ultimately, more work defining functional norms in learning from volatile environments is needed to tease apart the effects of outcome noise (unpredictability) and volatility on learning (Piray & Daw, 2020a, 2020b). Intriguingly, alternative code exits in the HGF (*hgf_jget*) to potentially resolve this issue between outcome noise and volatility uncertainty, but it is, as of yet, undocumented and unpublished in a research paper. Using this alternative variational approach, however, may provide compelling new insights into anxiety and the processing of different forms of uncertainty (Frässle et al., 2021; Piray & Daw, 2020b). In **Chapter 5**, we contribute further to our understanding of the role volatility estimates play in informing how anxious participants assimilate environmental statistics by investigating how trait anxiety affects reward-based learning in a volatile task. Beyond that, more models of anxious human behavioural learning responses are needed to expand our basic theoretical accounts and understanding of pathological processes and treatments. Discoveries that reveal how both subclinical trait anxiety and lab-induced state anxiety alter computations of uncertainty can aid in stimulating more hypotheses about cognitive and physiological markers and their subserving neurobiological processes in the pursuit of better treatments.

The HAD-S questionnaire of state anxiety provided a good means to quickly and effectively measure continuous self-reported anxiety levels in participants throughout the task. A

limitation of our task, however, is that we cannot discuss motivation levels. Our aim was to indirectly infer motivation from behavioural performance and model estimates. While self-reported motivation does depend on a participant's internal assessment, which may not be sufficiently granular concerning the potential impact on cognition, there are existing scales (such as Starkstein's apathy scale, see Starkstein et al., 1992) that provide an overall apathy rating based on questions such as '*Do you put much effort into things?*' and '*Do you have motivation?*' (Pessiglione et al., 2017). There are also alternative scales that differ in clinical clarity and the time it takes to fill out the questionnaire (Marin, 1990; Radakovic & Abrahams, 2014; Robert et al., 2002; Sockeel et al., 2006). A limitation of these scales however is that the assessment of motivation is not known to have neural counterparts (Pessiglione et al., 2017). Future analyses of similar datasets might do well to employ this technique of connecting self-reports of anxiety and motivation, behavioural responses to a reward-learning task, and computational modelling of responses to infer the effects of motivation from anxiety and how these interact with reward engagement.

Supplementary Figures



Supplementary Figure 1. Task design and measures. Time starts at the top left panel and finishes in the bottom left panel, proceeding from left to right. The first stage of the experiment is R1: Baseline, where participants sat and relaxed for 5 minutes with eyes open while we recorded ECG (to calculate HRV) continuously. After, the first HAD-S report was taken. Following task instructions (not pictured) 15

practice trials were undertaken lasting approximately 3 minutes. When practice was completed the second HAD-S report was taken. Following this, anxiety was induced following the procedure outlined in **Section 2.3** in the StA_{TD} and $StA_{TD\text{-}Cont}$ groups. In line with previous research (see Birkett, 2011; Kirschbaum et al., 1993; Labuschagne et al., 2019), 5 minutes of waiting time was added for the StA_{TD} and $StA_{TD\text{-}Cont}$ groups only to amplify anxiety. All groups then completed 200 trials of the reward learning task, lasting approximately 15 minutes. Every 40 trials a HAD-S report was filled out (5 HAD-S in TB1). Between task blocks was a self timed break of around 2 minutes. Afterwards, all groups completed the second 200 trial reward learning task block and another 5 HAD-S reports, one every 40 trials. Upon finishing, the anxiety manipulation was revoked in the state anxiety groups (as in **Section 2.3**) and a final HAD-S report was collected. The final part of the experiment was all groups presenting a piece of artwork privately for around 5 minutes.

Supplementary Materials

Model	Prior	Mean	Variance
3-level HGF	κ	1	0
	ω_2	-4	16
	ω_3	-7	16
	$\mu_2^{(0)}$	0	0
	$\sigma_2^{(0)}$	0.1	0
	$\mu_3^{(0)}$	1	0
	$\sigma_3^{(0)}$	1	0
	ζ	48	1
2-level HGF	κ	0	0
	ω_2	-4	16
	ω_3	-7	0
	$\mu_2^{(0)}$	0	0
	$\sigma_2^{(0)}$	0.1	0
	$\mu_3^{(0)}$	1	0
	$\sigma_3^{(0)}$	1	0
	ζ	48	1
HGFμ_3	κ	1	0
	ω_2	-4	16
	ω_3	-7	16
	$\mu_2^{(0)}$	0	0
	$\sigma_2^{(0)}$	0.1	0
	$\mu_3^{(0)}$	1	1
	$\sigma_3^{(0)}$	1	1

Table 1. HGF model parameter values for the estimation of prediction errors and predictions in the HGF for binary inputs without perceptual uncertainty. Listed in the table are the initial values of beliefs $\mu^{(0)}$ and variances $\sigma^{(0)}$ at each level of the 3-Level HGF (HGF_3), 2-level HGF (HGF_2), and 3-level HGF with μ_3 governing decision noise through a negative exponential (HGFμ_3, Diaconescu et al., 2014). Also listed are the perceptual parameters regulating belief updates (the coupling parameter κ, tonic volatility estimates on levels 2 and 3 (ω_2, ω_3) and the decision noise parameter ζ, estimated in log space (HGF_3, HGF_2). Free parameters estimated in unbounded space are log-transformed.

Chapter 5: Subclinical trait anxiety changes the neural signatures of predictions and prediction errors during reward-based learning: a MEG based study

The data reported in this chapter were collected by Zheng Gong and Marina Ivanova, supervised by Maria and Vadim Nikulin, to whom I am very grateful.[11] *Thanks to the Institute for Cognitive Neuroscience, National Research University Higher School of Economics, Moscow, Russian Federation for partially funding this study.*

[11] *This chapter is based on a manuscript that is currently in preparation.*

Author contributions M.H.R., V.N. and T.P.H. designed the experiment, Z.G, and M.I collected the data, M.H.R., T.P.H. analysed the data, M.H.R. and T.P.H. wrote code for data analysis, M.H.R. and T.P.H. wrote the manuscript. Z.G, and M.I applied Maxfilter and calculated questionnaire data.

Abstract

Misestimation of uncertainty has been proposed as a core statistical marker of learning difficulties in anxiety. In Bayesian inference, precision (inverse uncertainty) plays a core role in regulating the equilibrium between prior beliefs (higher cortical levels, descending predictions) and the processing of sensory input (lower cortical levels, ascending prediction errors, PEs). Hierarchical predictive coding (PC) process theories describe higher-level predictions as encoded by lower frequency oscillations (8–30 Hz) and lower-level precision-weighted prediction errors (pwPEs) encoded in higher gamma frequency oscillations (>30 Hz). Recent modelling and electroencephalography work has shown how states of anxiety bias uncertainty estimates in a volatile task, altering the spectral correlates of Bayesian PC and impairing overall reward learning performance. Here, we test whether trait anxiety interferes with reward learning, uncertainty estimates, and the expression of predictions and pwPEs in oscillatory activity using magnetoencephalography (MEG). Participants performed a volatile probabilistic reward learning task. We modelled behaviour using a hierarchical Bayesian learning model, quantifying the parametric effects of trial-wise estimates of pwPEs and predictions on the continuous MEG time-frequency responses using convolution modelling for oscillatory responses. Reward learning performance was poorer in high trait anxiety (HTA) in the first task block relative to low trait anxiety (LTA). Computationally, this was driven by increased volatility estimates and greater levels of environmental uncertainty and informational (estimation) uncertainty about the stimulus outcome tendency. Convolution of the parametric regressors to MEG oscillatory responses showed that HTA attenuated 8–16 Hz activity in central, frontal, and sensorimotor sensor regions relative to LTA. Encoding of predictions was associated with increased beta activity relative to a baseline in the LTA group only. No effects of trait anxiety on gamma modulation were found. Our results demonstrate that high trait anxiety disrupts hierarchical Bayesian inference about the statistical structure of volatile reward environments. Beyond the important implications for understanding the impact of uncertainty on anxiety and cognition, our results are useful for informing learning-based therapy treatments of anxiety disorders.

1. Introduction

Cognitive-affective structures like fear and anxiety serve an adaptive function in response to threats to survival and well-being (Grillon et al., 2019; Grupe & Nitschke, 2013). Fear rallies the fearful to take action. To fight or flee in response to specific and determinable threats. In comparison, anxiety enjoins the anxious to exercise caution. To seek safety in response to indeterminable threats with anticipated adverse outcomes (Grillon, 2008; LeDoux & Pine, 2016; Steimer, 2002). For this reason, anxiety is an uncertainty oriented psychological, physiological, and behavioural state (Carleton, 2016; Tovote et al., 2015)—with responses to uncertainty playing a central role in diagnosing anxiety (Carleton et al., 2012; Quintana et al., 2016). Crucially, anxiety has become one of the most common mental disorders among Western nations, present across most psychiatric disorders, with a heavy price paid by individual sufferers and societies alike (Beddington et al., 2008; Chisholm et al., 2016; Fineberg et al., 2013; Greenberg et al., 2015; Kessler et al., 2005, 2009; Stein & Craske, 2017). Those with anxiety may miss out on invaluable safety signals and rewarding feedback by avoiding potentially unpleasant outcomes (Bublatzky et al., 2017). Finding explanations for difficulties in learning for healthy anxious people represents a significant and understudied area of research that could benefit treatment and learning-based therapy techniques in anxiety disorders (Moutoussis et al., 2018).

Prior work under the predictive coding (PC) process theory and affiliated Bayesian frameworks have provided good evidence that healthy learners continuously refine an internal model of the world using hierarchically related prediction errors (PE) weighted by uncertainty (inverse, precision: precision-weighted PEs [pwPEs], Behrens et al., 2007; Bland & Schaefer, 2012; Doya, 2008; Feldman & Friston, 2010; Friston, 2005; Iglesias et al., 2013; Knill & Pouget, 2004; Mumford, 1992; Payzan-LeNestour et al., 2013; Rao & Ballard, 1999; Summerfield et al., 2011; Yu, 2007; Yu & Dayan, 2005). These computational learning quantities are thought essential to understanding learning difficulties in a constellation of different clinical disorders (de Berker et al., 2016; Parr, Rees, et al., 2018; Pulcu & Browning, 2019; Williams, 2016). This normative hierarchical updating policy outlined by Bayesian PC is thought orchestrated by distinct neural frequencies at particular cortical layers (Bastos et al., 2012; Sedley et al., 2016). Our work in **Chapter 3** investigated the understudied impact of affective states like state anxiety on pwPEs and prediction quantities and how they modulate the neural oscillatory patterns linked with PC. In this experiment, we ask how high levels of subclinical trait anxiety interfere with these learning and neural signals.

Normal levels of anxiety are considered adaptive (Gross & Hen, 2004; Mendl et al., 2010). Pathological anxiety is considered maladaptive—impeding everyday lives and vitiating aspects of cognition (Bishop & Gagne, 2018; Grupe & Nitschke, 2013; Rosen & Schulkin, 1998). Even at subclinical levels, high levels of anxiety can cause anguish and lessen our overall condition of health (Carleton, 2016; Chisholm et al., 2016; Dugas et al., 2005; Grillon et al., 2019). Recent advances in the computational understanding of decision-making have shown that anxiety disrupts how the brain processes information, forms beliefs, and estimates uncertainty—producing both adaptive and maladaptive learning, depending on the context (Aylward et al., 2019, 2020; Browning et al., 2015; Hein et al., 2021; Huang et al., 2017; Kim et al., 2020; Lamba et al., 2020; Mkrtchian, Aylward, et al., 2017; Paulus & Yu, 2012; Pulcu & Browning, 2017, 2019; Zorowitz et al., 2020). Our focus here is on the understudied effects of subclinical "trait" anxiety on reward-based learning. More specifically, using magnetoencephalography (MEG), we test how different levels of trait anxiety alter the neural oscillatory patterns involved with predicting and learning from rewards in an uncertain and changing environment.

Anxiety and uncertainty induce affective and cognitive changes that have behavioural consequences, causing distinct subjectively experienced distress (Birrell et al., 2011; Carleton et al., 2012; Grupe & Nitschke, 2013; Tovote et al., 2015). The highly anxious are susceptible to cognitive biases, such as detecting and processing threatening signals—assigning an overabundant supply of attentional resources (Bar-Haim et al., 2007; Bishop, 2008, 2009; Derryberry & Reed, 2002; MacLeod & Mathews, 1988). Also, anxiety can lead to the biased interpretation of emotional and uncertain stimuli as negative or threatening, driving avoidance behaviours (Arnaudova et al., 2017; Aylward et al., 2019; Blanchette & Richards, 2003; Borkovec et al., 2004; Mathews & MacLeod, 2005; Mkrtchian, Aylward, et al., 2017) and stifling learning (Zorowitz et al., 2020).

When faced with uncertain decisions, anxiety produces risk-averse responses, selecting, for example, more predictable lower payments instead of uncertain higher payments (Charpentier et al., 2017; Maner & Schmidt, 2006) and biasing expectations to pessimistic and adverse outcomes (Borkovec et al., 1999; Butler & Mathews, 1987; Jiang et al., 2018; Kim et al., 2020; Mitte, 2007; Stöber, 1997). Indeterminable threats can give rise to anxiety, leading to a hypervigilant state that can improve performance on perceptual based tasks (Cornwell et al., 2017; Grillon et al., 2019) and threat detection tasks (Bach, 2015; Cisler & Koster, 2010; Richards et al., 2014), also facilitating adaptive inhibitory responding (Aylward et al., 2017; Bach, 2015; Dayan & Huys, 2008; Grillon, 2008)—particularly for Pavlovian responses (Dymond, 2019; Edwards et al., 2010; Grillon et al., 2017; Mkrtchian, Roiser, et al., 2017;

Robinson et al., 2011). In comparison, tasks involving executive functions often expose impairments to performance in anxiety. For example, during distraction by emotion (Cornwell et al., 2011), performing economic decision-making tasks (Miu, Heilman, et al., 2008; Remmers & Zander, 2018), and more recently when engaging with reward and probabilistic learning (de Visser et al., 2010; Hartley & Phelps, 2012; Hein et al., 2021; Jiang et al., 2018; Sporn et al., 2020; Xia et al., 2021).

Suboptimal decision making under anxiety is well studied (Hartley & Phelps, 2012; Miu, Miclea, et al., 2008; Paulus & Yu, 2012). Yet, previous work on anxiety has, for the most part, neglected the reward-based nature of decision making, as depression is more tightly coupled with reward processing impairments while anxiety is linked with amplified aversive and negative stimuli sensitivity (Harlé et al., 2017)—despite the relevance of reward signals in learning-based therapeutic treatment (Moutoussis et al., 2018).

Investigation into the behavioural, neural, and computational underpinnings of reward-based decision making in healthy humans has been considerably successful (Daw et al., 2006; O'Doherty et al., 2001; Rangel et al., 2008; Schultz et al., 1997). Recent work has since extended this understanding into more complex, uncertain, and changing environments (Behrens et al., 2007; Diaconescu, Mathys, et al., 2017; Iglesias et al., 2013; Mathys et al., 2011; Schwartenbeck, FitzGerald, Mathys, Dolan, & Friston, 2015). However, affective states interact with the decisions we make (Lerner et al., 2015)—a dynamic and recurrent procedure that can both facilitate and hinder adaptive learning in uncertain and changing environments (Behrens et al., 2007; Yu, 2007). Thus, a greater understanding of the neurocomputational processes that subserve them needs to be achieved concerning potential deficits in reward functioning and maladaptive learning in anxiety (Meacham & Bergstrom, 2016; Paulus & Yu, 2012; Pulcu & Browning, 2019).

Complex unobservable states that shape our environments make optimal decision making an acutely complicated process. To learn from this complexity, the brain is thought to engage in a strategy almost surprisingly straightforward: the continuous distillation and embodiment of a growingly accurate model of the world, updated using the disparity between what we predict and what we experience (prediction error, PE). Put simply, each level of the cortical hierarchy transmits predictions concerning the level beneath, revised by PEs travelling up the hierarchy. This process of perception and learning by encoding and transmitting error signals resulting from predictions is known as Predictive Coding (PC, de Lange et al., 2018; Friston, 2005; Friston & Kiebel, 2009a; Kok & de Lange, 2015; Rao & Ballard, 1999; Srinivasan et al., 1982). Within this PC framework, the brain is thought to use probabilistic belief estimates to

incorporate uncertainty about the world (Doya et al., 2007). The optimal combination of distinct sources of information, according to probability theory, comes in the form of the sequential updating of beliefs using Bayes' rule (or Bayesian PC, Aitchison & Lengyel, 2017; Dayan et al., 1995; Friston & Kiebel, 2009a). According to Bayesian PC, the brain encodes uncertainty (the variance of probability distributions) and weights PE signals accordingly (using precision: pwPEs, see **Eq. 4–5 Chapter 2** and Feldman & Friston, 2010; Iglesias et al., 2013; Knill & Pouget, 2004). In the context of continuous learning about reward, we can iteratively sample potential choices from our reward environment and select the more probable option given our prior experience and uncertainty.

A detailed and practical taxonomy of the different forms of uncertainty: *expected, estimation, unexpected, and environmental (volatility) uncertainty* has been covered in **Chapters 1–2** and a review by Bland and Schaefer (2012) and later by Soltani and Izquierdo (2019). Unpredictable change can be stimulating and informative. Take again our weather example from **Chapter 1 Section 1.4.5**. In England, we frequently talk about the weather—we have a constant level of volatility. If our environment is regularly changing, it demands that we learn about it faster (if we want to avoid the sudden showers). This type of learning is characterised by adaptive behavioural responses (Soltani & Izquierdo, 2019). Adaptive learning betters our situation, sharpening our understanding of internal signals such as tissue deficits like pain or thermoreceptive feedback like being cold, and pairing these with appropriate actions, such as avoiding harm or the acquisition of a warm and dry environment (Averbeck & Costa, 2017). The statistical framing of inference and learning has driven an influential series of experiments examining how we monitor the uncertainty of our beliefs and how these differently affect the learning process during adaptive decision making (Behrens et al., 2007; Mathys et al., 2014; McGuire et al., 2014; Nassar et al., 2010; Piray & Daw, 2020a; Soltani & Izquierdo, 2019).

Consequently, investigating both individual differences and group differences using computational modelling has become an increasingly popular approach to understanding how both healthy affective states and psychiatric disorders alter this computational process through abnormal learning signals like pwPEs (Brazil et al., 2017; Browning et al., 2015; Cole et al., 2020; Deserno et al., 2020; Diaconescu et al., 2020; Farashahi et al., 2017; Huys et al., 2021; Iglesias et al., 2013; Katthagen et al., 2018; Lawson et al., 2020; Lawson et al., 2014; Paliwal et al., 2019; Piray, Ly, et al., 2019; Powers et al., 2017). Our aim here is to expand upon previous work showing adaptive learning difficulties in subclinical trait anxiety in an aversive learning setting with changing volatility levels (Browning et al., 2015). We investigate the effect of high levels of trait anxiety on reward-based learning signals in an environment with set volatility. We additionally focus on providing neuromagnetic evidence of the oscillatory

correlates of altered computational learning signals in trait anxiety, in line with PC theory (Bastos et al., 2012, 2015, 2020) and our recent work inducing transient states of anxiety in the lab (see **Chapters 2–3**, Hein & Ruiz, 2021).

In clinical work so far, changes to model-based decision making in response to reward have primarily been examined against the backdrop of anhedonia and motivational states in depression (Bishop & Gagne, 2018; Cooper et al., 2018; Harlé et al., 2017; Treadway et al., 2012). As mentioned above, instead of changes to reward processing, anxiety is more commonly linked with aversion sensitivity (Harlé et al., 2017; Olvet & Hajcak, 2008). However, studies concerning stress and trauma have demonstrated that the distinction between threat-related sensitivity in anxiety and decreased responsiveness to reward are not mutually exclusive (Nawijn et al., 2015). Also, recent experimental work inducing anxiety in the lab has been shown to incite reduced reward responsiveness and reward-based maladaptive beliefs in learning behaviour and neural dynamics (see **Chapter 2** and Bogdan & Pizzagalli, 2006; Bublatzky et al., 2017; Hein et al., 2021; Robinson, Overstreet, et al., 2013; Sporn et al., 2020).

On the one hand, we know a lot about anxiety and how it shapes learning from negative and punishing events (Grillon et al., 2019; Robinson et al., 2019). Anxious people are biased toward threat stimuli (Cisler & Koster, 2010), are intolerant to uncertainty (Carleton et al., 2012), and expect negative events and outcomes (Paulus & Yu, 2012). Grupe and Nitschke (2013) offered one explanation for this negative expectation bias: anxiety disrupts the encoding of PE transmission, leading to impaired belief updates about aversive or negative events—although this has since been questioned in light of new evidence (Bishop & Gagne, 2018). A simple hypothesis from Raymond et al. (2017) states that anxiety may increase the likelihood of repeating actions that avoid aversive outcomes.

Evidence from the computational modelling of Bayesian inference during sequential decision making has revealed how anxious participants model the dynamic parts of their environment, and these models can further infer upon their individual learning characteristics. While learning from aversive outcomes, one elegant paper demonstrated that high trait anxiety produces inflexible learning rate adaptation (Browning et al., 2015). The learning style from the high trait anxious group was a reduced capability to adjust outcome expectations between stable and unstable probabilistic outcomes. Moreover, Browning et al. (2015) revealed how volatility correlated with pupil diameter changes in their low trait anxiety control group. This pupil response, however, was relatively dampened in the high trait anxiety group (Browning et al., 2015). These results imply two conclusions: i) that highly trait anxious individuals may experience general impairments in utilising the higher-order statistics of aversive

environments (Raymond et al., 2017), and ii) that this results in interferences with learning rate adaptation as expressed by one physiological marker of volatility tracking, pupil diameter changes—with volatility known to drive learning rate changes (Aston-Jones & Cohen, 2005; Behrens et al., 2007; Vincent et al., 2019; Zénon, 2019). A subsequent study by Pulcu & Browning (2017) replicated these behavioural modelling results, showing impeded learning rate adaptation, but with respect to reward loss. One explanation is that anxious learners estimate an elevated degree of environmental change in all encountered environments (Bishop & Gagne, 2018).

On the other hand, how anxiety interferes with learning from reward-based signals remains largely overlooked, especially from a computational perspective (Paulus & Stein, 2006). While behaviours that actively seek reward can be impoverished in anxiety—a negative corollary of threat avoidance (Bishop & Gagne, 2018; Bublatzky et al., 2017; Raymond et al., 2017)—the ability of anxious participants to efficiently learn from rewarding signals with no loss or punishment outcomes while in an uncertain and volatile environments remains understudied. One paper provided evidence extending that of Browning et al. (2015) and Pulcu & Browning (2017) in high and low trait anxious participants, further confirming the overall misestimation of uncertainty and suboptimal decision making but in a reward-learning context (Huang et al., 2017). In addition to that, our recent work has explored this by inducing states of anxiety in the lab and controlling for trait anxiety levels (Hein et al., 2021; Sporn et al., 2020). By using the threat of an upcoming social stress test, Hein et al. (2021) provided support for overall poorer reward-based learning in a temporarily anxious group. Moreover, behavioural modelling revealed that state anxiety lowered the overall learning rate through reduced tonic volatility estimates (the size of belief updates on the stimulus outcome level), further maladaptively increasing the precision of posterior beliefs about the tendency of reward and producing biased increases to environmental and volatility uncertainty (Hein et al., 2021).

Since the PE signals associated with PC and reinforcement learning have clear neural signatures (Huang & Rao, 2011; O'Doherty et al., 2004; Raymond et al., 2017; Schultz et al., 1997), tracking those signals using electrophysiology and neuroimaging is a powerful means to understand hierarchical learning from uncertainty—especially when applied to detect between-group differences in anxiety (Robinson, Overstreet, et al., 2013). In healthy subjects, the neural correlates of hierarchically organised PE learning have been quite well established. A study by Iglesias et al. (2013) used PE quantities from a probabilistic reversal learning task to explain functional magnetic resonance imaging (fMRI) BOLD responses, revealing a neural hierarchy of lower level PEs about the stimulus outcomes and higher level PEs about stimulus probabilities. This has since motivated considerable research into the neural correlates of

computational learning quantities in the brain (Cole et al., 2020; Deserno et al., 2020; Diaconescu, Litvak, et al., 2017; Diaconescu, Mathys, et al., 2017; Diaconescu et al., 2019; Henco et al., 2020; Iglesias et al., 2021; Kolossa et al., 2015; Liu et al., 2021; Maheu et al., 2019; Mars et al., 2008; Nassar, Bruckner, et al., 2019; Nassar, McGuire, et al., 2019; Powers et al., 2017; Stefanics et al., 2018; Weber et al., 2020).

The neural signatures of feedback-based learning across prefrontal regions broadly overlap with the neural circuitry of anxiety (Robinson et al., 2019). Anxiety is known to dampen responses from the prefrontal cortices interfering with cognitive control (Bishop, 2009; Forster et al., 2015) and to disrupt the balance of amygdala responses, related to emotional appraisal (Bishop, 2007). Several researchers have also linked the amygdala with associative learning, potentially controlling the learning rate and, more broadly, uncertainty over environmental change (Averbeck & Costa, 2017; Holland & Gallagher, 1999; Homan et al., 2019; Li et al., 2011; Phelps et al., 2014; Roesch et al., 2012). As anxiety is an affective state, this may link reported anxiety-induced changes to volatility estimates (Browning et al., 2015; Huang et al., 2017; Pulcu & Browning, 2017) with amygdala modulations. But without further evidence from computational regressors and fMRI data, these links remain speculative.

A study using a two-choice prediction task and fMRI reported higher levels of BOLD activity in the anterior cingulate cortex (ACC) and medial prefrontal cortex (mPFC) in high trait anxiety participants in response to periods of low learning rate relative to lower trait anxious participants (Paulus et al., 2004). The ACC and posterior medial frontal cortex (pMFC) have been shown to encode information about beliefs using EEG (Holroyd et al., 2003; Kennerley et al., 2009; Montague et al., 2004; Yeung et al., 2005), with reward-based PEs in frontal medial electrodes exhibiting a negative ERP response (Cohen et al., 2007; Holroyd & Coles, 2002)—specifically in the ACC, located utilising trial based analysis of both EEG and fMRI data (Debener et al., 2005).

Later, using EEG and a probabilistic response and inhibition (go/no-go) task with both rewards and punishments, Cavanagh et al. (2019) demonstrated improved avoidance learning in high trait anxiety. The authors showed a closer relationship between negative PEs and ERP responses to punishment in anxiety, and further with theta-band (4–8 Hz) oscillatory changes linked with dorsal midline premotor regions (Cavanagh et al., 2019). This evidence aligns with prior work on anxiety and changes to Pavlovian-instrumental responses, and also to increased frontal midline theta responses to signals of control in anxiety when using punishing outcomes (Cavanagh et al., 2017; Cavanagh & Shackman, 2015; Mkrtchian, Roiser, et al., 2017). However, in that study's context using Pavlovian responses to punishments and rewards,

Cavanagh et al. (2019) reported that only depression was associated with reduced reward-related delta (1–4 Hz) oscillatory changes. However, few studies have examined the behavioural and neural responses to isolated reward feedback in anxiety.

Our EEG study in **Chapter 2** (Hein et al., 2021) reported that state anxiety did not significantly modulate single-trial ERP signals explained by pwPEs about rewards—as shown across frontocentral and central and left parietal regions in controls. This lack of significant ERP response in state anxiety is partially consistent with the attenuated neural responses in prefrontal regions reported in anxiety (Bishop, 2009; Forster et al., 2015; Grupe & Nitschke, 2013). Additional fMRI evidence supporting this dampened PFC response to reward in anxiety comes from post-traumatic stress disorder, where reduced activation of reward-processing areas (nucleus accumbens and mesial PFC) has been reported (Sailer et al., 2008). Moreover, a recent study using transcranial direct current stimulation across the dorsolateral prefrontal cortex of healthy participants increased reward-based learning rates, potentially showing that stimulation of the underactive areas that overlap with anxiety may improve engagement with reward, relevant to both depression and anxiety (Overman et al., 2021). However, much more evidence of the effects of anxiety on the neural correlates of predictions and pwPE signals about reward is needed.

Here we expand upon the modelling research of others in subclinical anxiety (Browning et al., 2015; Huang et al., 2017; Pulcu & Browning, 2017) and our electrophysiological work in state anxiety using both trial-wise ERP (see **Chapter 2**, Hein et al., 2021) and continuous time-frequency responses (see **Chapter 3**, Hein & Ruiz, 2021; Sporn et al., 2020) to investigate the effect of subclinical trait anxiety levels on reward-based learning in a volatile environment. In addition to examining how high levels of trait anxiety affect reward learning and model estimates of behavioural responses, we test how high trait anxiety levels shape the associations between reward-based predictions and pwPEs and neural oscillatory activity using high-resolution MEG recordings. Our additional motivation to record neuromagnetic data was to analyse the sources of the oscillatory correlates of pwPEs in a follow-up analysis by combining MEG responses with MRI scans.

Neural oscillations are thought to underlie atypical processes in several clinical disorders (Ronconi et al., 2020; Uhlhaas & Singer, 2010, 2012). At rest, those with high trait anxiety have been shown to exhibit increased alpha power, with separate peaks at both 8–9 Hz and 11 Hz, and decreased delta power (Knyazev et al., 2004). This anticorrelated relationship between alpha and delta is also related to behavioural inhibition and the vigilant detection of potential threat in the environment (Knyazev et al., 2002; Knyazev & Slobodskaya, 2003).

Rumination has been associated with theta (4–8 Hz) and beta band (13–30 Hz) oscillations (Andersen et al., 2009; Pavlenko et al., 2009), while increased beta oscillations have been linked with fixity of thought (Abramowitz et al., 2009; Engel & Fries, 2010; Hamilton et al., 2015; Lang et al., 2015; Williams, 2016). By inducing an anxiogenic state in the lab, we have shown that atypical increases in beta power and burst rate can explain reductions in the magnitude of pwPEs while learning about reward in a motor learning task (Sporn et al., 2020). And more recently, in **Chapter 3** (Hein & Ruiz, 2021), we complemented these findings by linking biased predictions about the tendency of reward in state anxiety to amplified beta oscillations and increased beta oscillations while encoding pwPEs that may represent the inhibition of pwPEs.

In Bayesian PC accounts of learning, there is accumulating evidence from human MEG/EEG and monkey electrocorticography studies suggesting that feedforward PE signals are encoded by faster gamma oscillations (>30 Hz), while backward descending predictions are expressed in lower alpha (8-12 Hz) and beta-band (13-30 Hz) oscillations (Alamia & VanRullen, 2019; Arnal & Giraud, 2012; Auksztulewicz & Friston, 2016; Bastos et al., 2012, 2015, 2020; Bauer et al., 2006; Brodski et al., 2015; Cao et al., 2017; Friston, 2005; Mayer et al., 2016; Pinotsis et al., 2016; Van Kerkoerle et al., 2014; van Pelt et al., 2016; Wang, 2010). Alpha and beta band activity is usually associated with afferent inhibitory effects that gates sensory processing (Arnal & Giraud, 2012; Jensen et al., 2012; Jensen & Mazaheri, 2010; Sedley et al., 2016; van Ede et al., 2011). While gamma-band activity—predominant in superficial layers—mediates the propagation of feedforward PE signals (Arnal & Giraud, 2012; Bastos, Litvak et al., 2015; Bauer et al., 2014; Engel et al., 2001; Michalareas et al., 2016; Sedley et al., 2016; Todorovic et al., 2011; van Pelt et al., 2016; Wang, 2010).

This is most evident in visual cortex studies where asymmetry is shown between alpha- and beta-band synchronisation in infragranular layers and gamma-band in supragranular layers, with alpha-beta, functionally inhibiting the processing of sensory input spiking, suppressing gamma rhythms (Arnal & Giraud, 2012; Bastos et al., 2012, 2015; Buffalo et al., 2011; Gould et al., 2011; Michalareas et al., 2016; Roberts et al., 2013; Xing et al., 2012). More recent work has also shown that precision weights on PEs are regulated in alpha and beta oscillations (Palmer et al., 2019; Sedley et al., 2016). Consequently, as precision values weight the transmission of PEs (Feldman & Friston, 2010), the composite pwPE signal may, as some experiments show, be represented in gamma and alpha/beta changes (Auksztulewicz et al., 2017; Sporn et al., 2020).

Crucially, the reward-related frequency correlates of pwPEs and predictions has been largely overlooked. Whether these rhythmic attributes in the PC framework represent one neural process by which anxiety biases belief updating and compromises learning is an unanswered question. In our work in **Chapter 3** and Sporn et al. (2020), we found beta oscillations were atypically increased in state anxiety during the encoding of pwPEs. In **Chapter 3** we also reported amplified beta activity from predictions about reward in state anxiety. This provides evidence that reward-based pwPE and prediction signals can be tracked in low-frequency oscillations—as outlined by the PC framework. However, similarly to Palmer et al. (2019), we did not observe gamma frequency changes using EEG. Potentially, reward-related pwPE signals did not modulate gamma activity in our EEG recordings. An alternative explanation, however, is that our EEG recordings were not sufficiently sensitive to detect reward-based prediction errors. We address this here by utilising MEG to confirm whether gamma-band activity is involved with reward-based pwPEs, and also for the future aim of investigating the sources of these oscillatory correlates.

To achieve this, we tested a low and high subclinically anxious group on a reward-based learning task. Behavioural responses were modelled using the Hierarchical Gaussian Filter (HGF), a model of perception and learning that includes reinforcement learning and Bayesian models (Mathys et al., 2011, 2014). We then extracted HGF estimates of predictions and pwPEs and used them as input to a convolution model to reveal the oscillatory activity modulated by these computational learning quantities (Litvak et al., 2013). We hypothesised, in line with our previous work in state anxiety, that high trait anxiety would amplify alpha/beta responses associated with predictions and pwPEs relative to low trait anxiety (**Chapter 3**, Sporn et al., 2020). Following Bayesian PC evidence in the sensory domain, we also explored changes to gamma activity during the encoding of reward-based pwPE signals (see Auksztulewicz et al., 2017; Bastos et al., 2018, 2020; Lundqvist et al., 2016, 2020; Miller et al., 2018; Palmer et al., 2019).

Behaviourally, we predict that the state-anxiety-related disregarding of information about the reward outcomes found in **Chapter 2** (Hein et al., 2021) will be replicated in our high trait anxious group, vitiating overall reward learning performance. We predict these effects will be driven by inflated estimates of environmental volatility, as previous work has suggested anxious learners overestimate volatility in all environments (Bishop & Gagne, 2018), corresponding with a recent theory about affective disorders being associated with misestimation of uncertainty (Pulcu & Browning, 2019).

2. Methods

2.1 Participants

Our sample size of 44 participants (19 male) was informed by work using similar tasks and subclinical trait anxiety and recent experiments using both the HGF and MEG recordings (Auksztulewicz et al., 2017; Behrens et al., 2007; Browning et al., 2015; Hein et al., 2021). All participants were healthy and aged between 18 and 36 years (mean 23.1, SEM 0.73) with no reported psychiatric conditions. Ethical approval for this study was granted by the Department of Psychology (National Research University) Higher School of Economics (HSE, Russian Federation) and participants were recruited from the HSE. [12]

2.2 Assessments of anxiety

Participants' trait anxiety level was measured twice using Spielberger's STAI inventory (Spielberger, 1983a): one assessment prior to attending the experiment as a selection procedure, and one before beginning the experiment (to correspond to the pre-screened level). Trait anxiety refers to a relatively stable metric of an individual's anxiety level derived from the self-reported frequency of anxiety from past experiences (Grupe & Nitschke, 2013). Anxiety in subclinical populations is commonly measured using the STAI scale (Spielberger, 1983a), a measure thought to reflect the general risk factor for an anxiety or emotional disorder (Grupe & Nitschke, 2013). This scale taps into the overall exaggerated perspective of the world as threatening, providing a good measure of how frequently a person has experienced anxiety across their life (Raymond et al., 2017).

We used the trait anxiety scores as a selection process to form the two experimental groups: low trait anxiety (LTA, defined as a STAI score below 35) and high trait anxiety (HTA, defined as a STAI score above 40). Trait anxiety scores ranged between 24 and 65. The average anxiety scores for each group were LTA (mean 30.5, SEM 0.8) and HTA (mean 51.7, SEM 1.5). Importantly, the experimental groups were not confounded by age and sex. The high trait anxiety group (HTA, mean age 22.6, SEM = 1.1) consisted of 12 females, while the low trait anxiety group (LTA, mean age 23.7, SEM = 1.0)) consisted of 13 females. In addition to the trait inventory, measures of self-reported state anxiety using the STAI scale (X1, 20 items,

[12] This study was conducted between August 2020 – June 2021 and the study complied with stringent COVID-19 health and safety criteria described in an international protocol (led by NYU).

Spielberger, 1983a) were taken prior to the experiment and after completing the experiment to assess whether state anxiety changed as a function of conducting the experimental task.

In our previous work, we assessed heart rate variability (HRV) as a proxy measure for state anxiety (Hein et al., 2021). We calculated the coefficient of variation (CV = standard deviation/mean) of the difference intervals between consecutive R-peaks (inter-beat interval, IBI) extracted from the continuous electrocardiography (ECG) data as a metric of HRV. In previous empirical studies, states of anxiety have been shown to lower HRV and high frequency HRV (HF-HRV, 0.15 – 0.4 Hz, see Friedman, 2007; Gorman & Sloan, 2000; Hein et al., 2021; Sporn et al., 2020; Thayer et al., 1996). However, experiments on the association between trait anxiety levels and HRV/HF-HRV metrics have reported both reductions (Bleil et al., 2008; Miu et al., 2009; Mujica-Parodi et al., 2009) and small or inverse effects (Dishman et al., 2000; Narita et al., 2007). Thus, the relationship between trait anxiety and HRV remains unclear. Anxiety induced effects are primarily reported in anxiety disorders or responses to transient states of anxiety (Aikins & Craske, 2010; Chalmers et al., 2014; Fuller, 1992; Klein et al., 1995; Pittig et al., 2013; Quintana et al., 2016). In this study, we include a complementary analysis of both HRV and HF-HRV to supplement our self-report measures of trait and state anxiety. See section **2.6 MEG and ECG Recording and Pre-Processing** for further details.

2.3 Experimental design

Our experiment tested the difference between two groups: HTA and LTA. These participants all performed a probabilistic binary reward-based learning task in a volatile learning setting (Behrens et al., 2007; de Berker et al., 2016; Iglesias et al., 2013). The session was split between an initial resting state block (R1: baseline) of five minutes and two experimental reward-learning task blocks consisting of a total 320 trials (TB1, 160 trials – TB2, 160 trials). During the continuous recording of MEG and ECG responses in the baseline block, participants were told to try to relax and fixate on a central point of the screen with their eyes open.

Similarly to **Chapter 2**, participants needed to learn from changing probabilistic reward feedback by predicting whether a blue or orange image in a given trial was the rewarding stimulus (outcome win, 5 points reward – outcome lose, 0 points unrewarded). Participants were informed that the total sum of all their rewarded points would translate into monetary reward at the end of the experiment. The calculation for this remuneration was the sum total of winning points divided by six plus four hundred, given in Russian rubles ₽ (for example, 960

points pays 960/6+400 = 560₽). The probabilistic relationship between the two stimuli was inversely proportional; if one image had a ground truth 0.9 probability of rewarding, the other image necessarily had a 0.1 probability of rewarding (p, 1−p). The mapping between the probability of reward and images changed 10 times across the total 320 trials every 26 to 38 trials, with each contingency occurring twice. The possible contingencies were 0.9/0.1 and 0.1/0.9, 0.7/0.3 and 0.3/0.7, and chance level 0.5/0.5, as in **Chapter 2** and de Berker et al. (2016).

For every trial, a blue and an orange stimulus were shown on the monitor. Their location was either to the right or left of centre, randomly generated in each trial. The maximum time allowed for a response before the trial timed out was 1300 ms ± 125 ms. In contrast to our previous EEG study (**Chapters 2–3**), responses here were given by using a button press with either the left or right thumb (corresponding to selecting either the left or right image). After a prediction was commissioned by the participant, the selected image was outlined in bright green for 1000 ms (± 200 ms) to indicate their response. After, feedback of the trial outcome was provided (win, green; lose or no response, red) in the centre of the screen for 1250 ms (± 250 ms). To conclude a trial, a fixation cross was shown in the centre of the screen (1750 ms [± 250 ms]). Participants were told to select the image they believed would reward them to maximize reward across the 320 trials, and also to modify their selections in response to any inferred changes to their underlying probability. Prior to starting the experimental task blocks (TB1, TB2), each participant performed 16 practice trials and filled out the first state anxiety report. Between the two experimental task blocks, participants rested for a short self timed interval. After completing the second task block, participants filled out the second state anxiety report before finishing the experiment.

2.4 Behaviour and modelling (The Hierarchical Gaussian Filter)

To model behaviour, we used the Hierarchical Gaussian Filter (HGF, Mathys et al., 2011, 2014, version 6.0.0). As with previous **Chapters (2–4)**, we utilised a binary 3-level HGF perceptual model (Mathys et al., 2011) where (hidden) states in our reward-learning task perform Gaussian random walks, generating outcomes across time that are experienced by participants as stimuli (inputs). Following **Chapter 4**, we paired this perceptual model with two alternative softmax response models that map a participant's beliefs to their decisions (Daunizeau, den Ouden, Pessiglione, Kiebel, Stephan, et al., 2010): i) with a fixed decision noise parameter ζ (shaping choice probability, for further detail see **Eq. 21** in **Chapter 2**; Mathys et al. [2011]) ii) where ζ is equal to the exponential of negative log-volatility ($e^{-\mu_3^{(k)}}$),

thus depending on trial-wise changes to volatility estimates—termed hereafter $HGF_{\mu 3}$ (Diaconescu et al., 2014). Here as in **Chapters (2–4),** subject-specific parameters (ω_2, ω_3, ζ) were estimated while κ, the phasic volatility parameter, was fixed to 1 (as the scale of x_3 is arbitrary, see de Berker et al., 2016; Weber et al., 2020). Four free parameters ($\mu_2^{(0)}$, $\sigma_2^{(0)}$, $\mu_3^{(0)}$, $\sigma_3^{(0)}$) determine the starting values of beliefs. All prior settings on hierarchical Bayesian model parameters are summarised in **Supplementary Materials, Table 1**).

Using the prior parameter values (**Supplementary Materials, Table 1**) and series of inputs, maximum-a-posteriori (MAP) estimates of model parameters were then quantified and optimised using the quasi-Newton optimisation algorithm (Cole et al., 2020; Diaconescu et al., 2014; Reed et al., 2020).

2.5 Alternative models

Our expectation given in previous chapters is that learning will be hierarchically coupled. However, as with previous **Chapters (2–4)**, we reasoned that participants' decisions when not undergoing an experimental anxiety manipulation may best be explained by simpler non-hierarchical models. To address this, we compared a Bayesian family of models to a family of reinforcement learning algorithms (Cole et al., 2020; de Berker et al., 2016; Lawson et al., 2020). For the Bayesian family, we included three models. The first was the binary 3-level HGF (HGF_3) that captures volatility estimates and uses a set decision noise parameter for mapping beliefs to decisions (Mathys et al., 2011). The second was the 3-level HGF where decisions depend dynamically on estimated volatility ($HGF_{\mu 3}$, see Diaconescu et al., 2014). In addition, we used a version of the HGF with only two levels (HGF_2) where the third level volatility is fixed. For the family of reinforcement learning models, we used a reward-maximising Rescorla-Wagner (RW) model with a fixed learning rate (see **Chapter 1** and Rescorla & Wagner, [1972] for further details) and a Sutton K1 model (SK1) using a dynamic learning rate (Sutton, 1992). Model comparison was performed using random effects Bayesian model selection (BMS, see Stephan et al., 2009) using code from the MACS toolbox (Soch & Allefeld, 2018).

2.6 MEG and ECG Recording and Pre-Processing

Using a whole-head Elekta Neuromag VectorView MEG scanner (Elekta AB, Stockholm, Sweden) we recorded magnetic fields using 306-sensors (102 magnetometers, 204 planar gradiometers) with a sampling rate of 1000 Hz. We examined only the magnetometers (the portion of the magnetic field perpendicular to the sensor's coil) as we aimed to follow up these

analyses by source localising the MEG signals, and magnetometers are more sensitive to deeper cortical sources than planar gradiometers (Hansen et al., 2010; Parkkonen et al., 2009). Despite most superficial sources being well detected by both kinds of sensors, magnetometers measure fields from a wide distance and are more sensitive to distant sources. (Though see a recent study showing that after signal space separation [SSS] both express the same information, Garcés et al., 2017).

To denoise the MEG data and control for head movements we used a head-position indicator (4 coils affixed to the head, 2 placed on the top of each side of the forehead, and 2 on the mastoid process of each side) and the Elekta MEG workstation's Maxfilter™ (Elektra Neuroscience, 2010). Eye movements were controlled using an electrooculogram (EOG) by applying 4 electrodes recording continuously. Two horizontal EOG electrodes were affixed to each side of the temple, while the 2 vertical EOG electrodes were affixed to above and below one eye. Also, two electrodes were used for electrocardiography (ECG) recording. One electrode was placed on the sternum, while the other was placed on the left side beneath the rib cage. These ECG electrodes were placed in a two-lead configuration (Moody & Mark, 1982) calibrated to fit the Einthoven triangle (Wilson et al., 1931).

In addition, we preprocessed the data using the FieldTrip toolbox (Oostenveld et al., 2011) by high-pass filtering at 0.5Hz for low drifts (Butterworth filter [IIR]) and notch-filtering at 50Hz to remove power line noise. We also downsampled to 250 Hz and removed artefacts using independent components analysis (ICA, fastICA algorithm) and removed on average 3.08 components (SEM 0.08). Subsequently, the continuous MEG signals were epoched around outcome onsets (win, lose, no response) and around stimuli onset (blue, orange image) from −500 to 2000 ms. Due to a low signal-to-noise ratio associated with technical issues during the MEG recording (estimated as in Vidaurre et al., 2020), five participants were excluded from our MEG analysis. In the remaining sample of 39 participants, we excluded noisy epochs exceeding +/− 3 standard deviations of the median femtotesla (fT, 10–15 tesla) signal. In addition, we identified and removed by visual inspection any excessively noisy trials, with the total number of rejected trials in each participant not exceeding 10% of the total number of trials recorded (~32 trials). [13]

We extracted cardiac events (the QRS-complex, R wave peak) from the continuous ECG data using the FieldTrip toolbox. Afterwards, we calculated the latency of each R-peak and

[13] Cleaned MEG and preprocessed behavioural data files are available in the Open Science Framework Data Repository: https://osf.io/b4qkp/

calculated the CV of the inter-beat interval (IBI), our proxy metric for HRV across both experimental task blocks. The CV for each participant was normalised to the resting state block (R1: baseline). To estimate the high frequency content of the HRV (HF-HRV) we used the IBI time series. First, we interpolated at 1Hz with a spline function (order 3), and subsequently estimated spectral power using Welch's periodogram method (Hanning window, following Rebollo et al., 2018). Estimates of power were normalised to the average power in R1 and converted to decibels (dB) for statistical analysis.

2.7 Spectral Analysis

We first tested the average spectral power between our high trait anxiety and low trait anxiety groups in the first resting state block (baseline: R1). Using the raw data, we calculated the standard power spectral density within 1–90 Hz (fast Fourier transform, Welch method, Hanning window of 1 s, 125% overlap) and converted it into decibels (dB: $10*\log_{10}$). Afterwards, we tested how HGF predictors modulated the “phasic” changes in the time-frequency (TF) representations during the experimental task blocks.

To achieve this, similarly to **Chapter 3**, we estimated standard TF representations of the continuous MEG data using Morlet wavelets. TF spectral power was extracted between 8 and 90 Hz. For alpha (8–12 Hz) and beta (13–30 Hz) frequency ranges we used 5–cycle wavelets shifted every sampled point in bins of 2 Hz (Kilner et al., 2005; Litvak et al., 2011; Ruiz et al., 2009). For gamma band activity (31–90 Hz), 7-cycle wavelets sampled in steps of 2 Hz were used.

After transforming the MEG continuous signal to TF representations, we used linear convolution modelling for oscillatory responses to model the continuous TF data (Litvak et al., 2013). Our approach was identical to that taken in **Chapter 3**, to use trial-wise regressors from the HGF (model estimates of predictions and pwPEs) to examine their effect on the oscillatory responses while controlling for the modulation induced by the stimuli presentation, outcome feedback (win, lose), and response regressors. Potentially, in line with **Chapter 3**, a prediction is formed and represented in the frequency response of the brain somewhere between the stimuli presentation and the outcome resolution, co-occurring with response commission. As the convolution approach can isolate and distinguish between overlapping events, it is ideal for analysing the expression of predictions in MEG data.

For a more in-depth treatment of our convolution general linear model (GLM) approach see **Section 2.6** of **Chapter 3**. It is sufficient here to mention that the convolution GLM approach

seeks to explain our continuous TF transformed MEG signals across time as a linear combination of explanatory variables (regressors like HGF model estimates of predictions and participant responses) and residual noise (see Litvak et al. [2013] for further details). In this way, the convolution GLM method estimates coefficients (β) for each regressor and frequency using ordinary or weighted least squares. The resulting TF images can be interpreted as deconvolved TF responses to particular event types and parametric regressors.

The procedure for performing convolution modelling is informed by code developed by Spitzer et al. (2016). [14] To adhere to the GLM error assumptions (see Kiebel et al., 2005; Litvak et al., 2013) we first needed to convert the spectral power to amplitude by executing a square-root transformation. After, we coded a matrix of discrete regressors denoting the stimuli presentation, the participant's prediction (response: left, right, no response), the feedback from the prediction (outcome: win, lose), and the two parametric regressors extracted from the sequence of estimates from the HGF model: unsigned predictions on level 2 ($|\hat{\mu}_2|$) about the tendency towards a stimulus-reward contingency (henceforth: 'predictions') and precision-weighted prediction errors (pwPEs) about stimulus outcomes encoding the magnitude of the belief update about the reward contingency ($|\varepsilon_2|$, henceforth: 'pwPEs').

Similarly to **Chapter 2**, we found extremely high linear correlations between the second level HGF estimates of pwPE about stimulus outcomes and the third level pwPEs about environmental change (ε_3), and high linear correlations between predictions on the second level about the reward outcomes and predictions about volatility on the third level ($|\hat{\mu}_3|$). For pwPEs, the Pearson correlation coefficients ranged from 0.67 to 0.95 among all 44 participants (mean 0.86, SEM 0.01). And for predictions the range was −0.95 to 0.37 (mean −0.66, SEM 0.05) for the Pearson correlation coefficients in 44 participants. Because of this significant multicollinearity, both pwPE and predictions on level 3 have been excluded from subsequent analysis (for further information on the effect of multicollinearity among regressors on the GLM, we refer the reader to **Chapter 2** and Mumford et al. [2015] and Vanhove [2020]). The remaining second level HGF pwPE and prediction regressors were uncorrelated (<0.25, mean −0.05, SEM 0.01, similar to **Chapter 3** and Auksztulewicz et al., 2017; Iglesias et al., 2013; Vossel et al., 2015).

Informed by our previous work in state anxiety, the pwPE convolution model here was estimated using a window from −500 to 1800 ms relative to the outcome event, but this window was subsequently refined for the statistical analysis based on our previous findings (**Chapter**

[14] Available at https://github.com/bernspitz/convolution–models–MEEG

3). For predictions, we similarly used the results from **Chapter 3** to motivate focusing only on a stimulus-locked analysis, running convolution modelling from −500 to 1800 ms. In all convolution analyses (alpha, beta, gamma), each discrete and parametric regressor was convolved with a 20th-order Fourier basis set (40 basis functions, 20 sines and 20 cosines), following our **Chapter 3** gamma approach and Litvak et al. (2013). Each convolution model, (outcome-locked and stimulus-locked) was run between −500 to 1800 ms affording the GLM to resolve modulations up to ~9.7 Hz (20 cycles / 2.3 seconds; or ~103 ms).

2.8 Statistical Analysis

For model-free and model based analysis, we followed a similar process to **Chapter 2** to examine differences in behavioural and computational model variables between our group factor (low trait anxiety, LTA; high trait anxiety, HTA) and block (TB1, TB2). Our dependent variables (DVs) were i) averaged model-free behavioural data over task blocks (error rates, reaction time: RT); ii) coefficient of variation (CV) and spectral measures (expressed in dB) of HRV (normalised to R1 baseline, averaged over task blocks); iii) averaged HGF model estimates over task blocks of computational trajectories: a) informational uncertainty about the stimulus outcomes x_2 (σ_2); b) belief estimates about volatility (mean, μ_3, and variance, σ_3); and c) environmental uncertainty: $\exp(\kappa\mu_3 + \omega_2)$ i.e., the volatility of the environment; iv) HGF perceptual model parameter quantities ω_2 and ω_3. We used non-parametric factorial synchronised permutations tests (Basso et al., 2007) with 5000 permutations to test the main effects and interactions in our DVs i-iii, and used planned pair-wise comparisons using permutation tests (5000 permutations) for DV iv (ω_2 and ω_3). Following **Chapter 4**, in the event that code for our preferred N x 2 non-parametric factorial tests based on synchronised permutations are not available (where N > 2), we instead use 2 x 2 analysis of variance (ANOVA).

To address the multiple comparisons problem, where it arises, we control the false discovery rate (FDR) using an adaptive linear step-up procedure set to a level of q = 0.05 providing an adapted threshold p-value (P_{FDR}, Benjamini et al., 2006). In the results section below, we provide the mean and standard error of the mean (SEM) for our dependent variables with estimates of the non-parametric effect sizes for pair-wise comparisons and associated bootstrapped confidence intervals (Grissom & Kim, 2012; Ruscio & Mullen, 2012). As in **Chapters 2–4**, the within-group effect sizes used the probability of superiority for dependent samples (Δ_{dep}) while the between-group effect sizes used the probability of superiority (Δ, see Grissom and Kim [2012]).

For MEG responses, we used the FieldTrip Toolbox for statistical analysis (Oostenveld et al., 2011) by converting the SPM TF images (in arbitrary units, a.u.) to a Fieldtrip structure. We first examined the difference in raw spectral power at rest (R1: Baseline) between HTA and LTA by testing the average in each frequency band (delta 1–3 Hz, theta 4–7 Hz, alpha 8–12 Hz, beta 13–30 Hz, low gamma 31–49Hz, high gamma 51–90 Hz) using a cluster-based permutation approach (Maris & Oostenveld, 2007; Oostenveld et al., 2011). After, we performed 3D statistical tests in 102 magnetometers in alpha and beta (8–30 Hz) and gamma (31–90 Hz) frequency bands across time using a cluster-based permutations (Maris & Oostenveld, 2007; Oostenveld et al., 2011).

Informed by our results of the modulation of time-frequency responses in state anxiety from **Chapter 3**, we tested within and between-group effects independently. Here, dependent-samples two-sided tests were used to assess the within-group changes in oscillatory responses by HGF regressors (t-distribution critical value threshold: 2.5^{th} and 97.5^{th} quantiles). As highlighted in **Chapter 3**, time-frequency images in convolution modelling for oscillatory responses are not baseline corrected by subtraction or division by the average baseline level but are rather estimated comparably with post-event activity. This is achieved by incorporating the latency variation of events from the continuous recording (Litvak et al., 2013). Between-group modulation of the HGF regressors on neural oscillations was tested using independent samples two-sided tests. We tested the hypothesis that high levels of trait anxiety would be associated with changes in alpha/beta activity and additionally explored whether there would be reduced gamma activity when compared to low trait anxious participants.

For the dependent samples outcome-locked analysis of pwPEs, TF images were compared in a time window between 500 to 1000 ms relative to a baseline period averaged between −200 to 0 ms. This window of analysis was motivated by the within-group effects in our experimental state anxiety group from **Chapter 3** between 500–1000 ms. Similarly, in **Chapter 3**, the between-group effect of pwPE on neural oscillatory responses occurred approximately between 1000–1500 ms. Accordingly, here we used independent samples tests between 1000–1500 ms post-outcome normalised by subtracting the average response in a.u. between −200 to 0 ms pre-outcome.

The prediction regressor was examined by locking to the time of the stimulus. Informed by the results of **Chapter 3**, within-group changes were tested between 100–1000 ms relative to a baseline averaged between −200 to 0 ms pre-stimulus. Again, we controlled for the effects of the participant's responses on the expression of activity modulated by predictions by including a response regressor in the convolution GLM. The independent samples equivalent for

stimulus-locked analysis on the prediction regressor was carried out between 100–1000 ms baselined by subtracting the average response in a.u. between −200 to 0 ms pre-stimulus.
To handle multiple comparisons across our time, sensor space and frequency, all analyses controlled the family-wise error rate (FWER) at level 0.05 (5000 iterations, Maris & Oostenveld, 2007; Oostenveld et al., 2011).

3. Results

3.1 Measures of anxiety

We found no correlation between trait anxiety scores and the HRV index (Spearman rank correlation, ρ = 0.19, *P* = 0.26) and no correlation between trait anxiety and high-frequency HRV (HF-HRV, 0.15 – 0.4 Hz, ρ = 0.19, *P* = 0.25). Factorial testing on the normalised HRV index was conducted using non-parametric 2 x 2 permutation tests with synchronised rearrangements. The results revealed there was no significant main effect of Group (*P* = 0.57), Block (*P* = 0.21) and interaction effect (*P* = 0.91, see **Supplementary Figure 1A**). Analysis of the spectral characteristics of the IBI time series using a non-parametric 2 x 2 factorial test also demonstrated no significant difference in HF-HRV for the main effect of Group (*P* = 0.78), Block (*P* = 0.68) and interaction effect (*P* = 0.08, see **Supplementary Figure 1B**).

Subjective self-reported measures of state anxiety showed a significant main effect of the Group factor (HTA mean 35.4, SEM 1.9; LTA mean 27.9, SEM 1.1, *P* = 0). However there was no effect of the factor Block (*P* = 0.23) or interaction effect (0.22). This analysis demonstrates that the HTA were subjectively more anxious than LTA, both prior to and following the reward-learning task.

3.2 Model-free analysis

3.2.1 Error Rate

To assess overall reward learning performance we tested the error rate (number of ‘lose’ trials divided by the total number of trials) in each task block (TB1, 120 trials; TB2, 120 trials) using non-parametric factorial tests with synchronised rearrangements. This approach yielded a significant main effect of the factor Block (*P* = 0.002) and a significant interaction effect (*P* = 0.004), but a trend level Group effect (*P* = 0.06, see **Figure 1A**).

To explore the significant interaction effect, post-hoc analysis with pair-wise permutation tests revealed that the main effect of Group was driven by a significantly higher error rate in HTA during TB1 (mean 38.72, SEM 1.11) relative to LTA (mean 34.71, SEM 0.81, $P_{FDR} < 0.05$, $\Delta = 0.74$, CI = [0.64, 0.89]). There was no significant difference between the two groups in TB2 (LTA, mean 34.48, SEM 0.71; HTA, mean 33.44, SEM 0.95, $P_{FDR} > 0.05$, see **Figure 1A**).

Because LTA did not perform significantly better in TB2 relative to TB1 (despite the result of a significant main effect of the Group factor), this may suggest that this group reached an optimal task performance already in TB1, which could confound the results. To examine this further, we calculated the 'optimal' solution to our task by estimating the average error rate in TB1 given absolute knowledge of the ground truth changes in probabilistic contingencies. This simulated agent would have, on average across all series of observed inputs from all participants, an error rate of 26%, which is significantly lower than the LTA group rate (35%). This outcome demonstrates two points: i) the LTA group did not reach the 'optimal' level of performance during TB1 ii) both groups could improve performance from TB1 to TB2. We thus interpret these model-free results as poorer reward-learning performance in HTA relative to LTA in block 1, with faster learning about reward in HTA leading to a similar reward-learning performance by block 2.

3.2.1 Reaction time

Averaged reaction times (RT) were assessed in milliseconds (ms), revealing no significant main effects in our factor Group or interaction effect ($P = 0.27$, $P = 0.44$, respectively), consistent with our previous work in **Chapters 2–4** and prior studies in anxiety (Bishop, 2009). In contrast to our previous **Chapters 2–4**, there was no significant effect of the factor Block ($P = 0.23$). Following **Chapter 2**, as a separate analysis in line with our lack of between-group differences in RTs, we tested both groups (HTA + LTA) average RT taken from the current trial for lose outcome and win outcome trials. In contrast to our previous results in **Chapters 2–4**, the difference between the RT on lose trials (mean 466.04, SEM 22.81) and win trials (mean 459.33, SEM 22.52) was not significant ($P = 0.74$).

As in **Chapters 2–4**, to show that classical attentional effects do not account for the between-group differences in the percentage of errors in overall reward-learning performance reported above, we examined RT among the predictable (0.9–0.1) and unpredictable contingency phases (0.5–0.5). Anxiety-related attentional deficits during the task would be expressed by increases in RT in the 0.9–0.1 phase but not the 0.5–0.5 phase. To test this, we first averaged the RT independently in 0.5–0.5 and 0.9–0.1 probabilistic mappings from both blocks. After,

we executed planned pair-wise permutation tests. These assessed the mean RT between HTA and LTA groups independently for each probabilistic mapping. For the 0.9–0.1 probabilistic mapping there was no significant difference between HTA (mean 496.12, SEM 16.21) and LTA (mean 512.56, SEM 17.62, $P = 0.41$). Similarly, no difference existed between the 0.5–0.5 probabilistic mapping in HTA (mean 497.35, SEM 17.49) and LTA (mean 523.14, SEM 20.78, $P = 0.38$). The results above are both consistent with our previous **Chapters 2–4** and with the conclusion that no deficit in classical attentional resources was driving poorer task performance in our HTA group (Prinzmetal et al., 2009).

3.3 Bayesian model selection

We extracted the log-model evidence (LME) for all participants (N = 44) in all models (HGF: 3-Levels [HGF_3], 2-Levels [HGF_2], $HGF_{\mu 3}$, Rescorla Wagner [RW], and the Sutton K1 [SK1]) and used the LME to conduct Bayesian model selection (BMS). The family of Bayesian models yielded stronger evidence (exceedance probability = 1; expected frequency = 0.94, see **Figure 2**) for explaining the task behaviour when compared with the reinforcement-learning models. Afterwards, we ran BMS comparing each Bayesian model (HGF_3, HGF_2, and $HGF_{\mu 3}$). This procedure demonstrated that the $HGF_{\mu 3}$ model was more likely to explain the data (exceedance probability = 1; expected frequency = 0.95, see **Figure 2**). Importantly, we confirmed that the $HGF_{\mu 3}$ model was equivalently the best model at describing responses independently in each group (LTA exceedance probability = 1; expected frequency = 0.91; HTA exceedance probability = 1; expected frequency = 0.91). As with **Chapter 3**, this result revealed that each group's task behaviour was best explained by the $HGF_{\mu 3}$, permitting comparison of parameter estimates between groups.

3.4 Model-based analysis

3.4.1 High trait anxiety increases volatility belief estimates

The mean of the belief distribution on level 3, μ_3, represents mean estimates of volatility x_3. The step size of x_3 depends on the exponential of a positive constant parameter ω_3 (the lower ω_3 the slower a participant updates their beliefs about volatility). We observed differences between the groups in model estimates of μ_3, with a significant main effect of the factor Group ($P = 0.003$), but no significant Block ($P = 0.84$) or interaction effect ($P = 0.92$). Pair-wise permutation tests demonstrated higher estimates of volatility in the HTA group (mean −0.34, SEM 0.11) relative to the LTA group when averaging over experimental blocks (mean −0.67,

SEM 0.11, P = 0.001, Δ = 0.71, CI = [0.52, 0.85], see **Figure 2**). No difference was found between groups in the associated parameter ω_3.

The higher posterior belief estimate of volatility in the high trait anxiety group represents a subjectively greater level of task environmental change in comparison to the low trait anxiety group. From our model comparison, the $HGF_{\mu 3}$ model, which dynamically assimilates belief estimates about volatility across trials, provided a better account of the data. In the $HGF_{\mu 3}$ response model, if more volatility is estimated by a participant the inverse decision temperature parameter ($\beta = e^{-\mu_3}$) becomes lower, leading to a noisier mapping between beliefs and responses and thus more exploratory behaviour. When a participant estimates lower levels of volatility in the probabilistic contingencies leading to reward, they will exhibit a more deterministic coupling between belief and response.

Our results suggest that choice probability in the HTA group is more stochastic and exploratory as a higher level of volatility is estimated in that group relative to the LTA group. Also, this result reveals the importance of individual differences in belief estimates about volatility varying across trials and its impact on learning (Diaconescu et al., 2014). In addition, for trait anxiety, this result emphasises the use of a hierarchical model with a core function of volatility estimates for inferring the underlying environmental statistics and deciding upon responses.

To further test this model-based result that describes more exploratory behaviour in HTA, we tested an additional model-free metric corresponding to exploratory response choice, the empirical switch rate (trial-to-trial response switches, see Aylward et al. [2019]). Several authors have also reported impaired switching strategies in anxiety (Ansari et al., 2008; Huang et al., 2017; Paulus et al., 2004; Xia et al., 2021).

3.4.1.1 Empirical switch rate

We first calculated the average probability of a response switch (e.g. choosing the blue (orange) image after having chosen orange (blue) in the previous trial) and the average estimate of volatility in all trials in each participant. Using a non-parametric Spearman rank correlation we revealed a highly significant relationship in these variables across participants (N = 44, ρ = 0.91, P = 3.3270e−18). We observed a similar outcome when separately testing correlations within the LTA group (ρ = 0.90, P = 2.6781e−06) and HTA group (ρ = 0.91, P = 2.5989e−14). Permutation-based factorial testing on the switch rate revealed a significant main effect of the factor Group (P = 0.01) but not for Block (P = 0.23) or the interaction effect (P = 0.29, see **Figure 1B**).

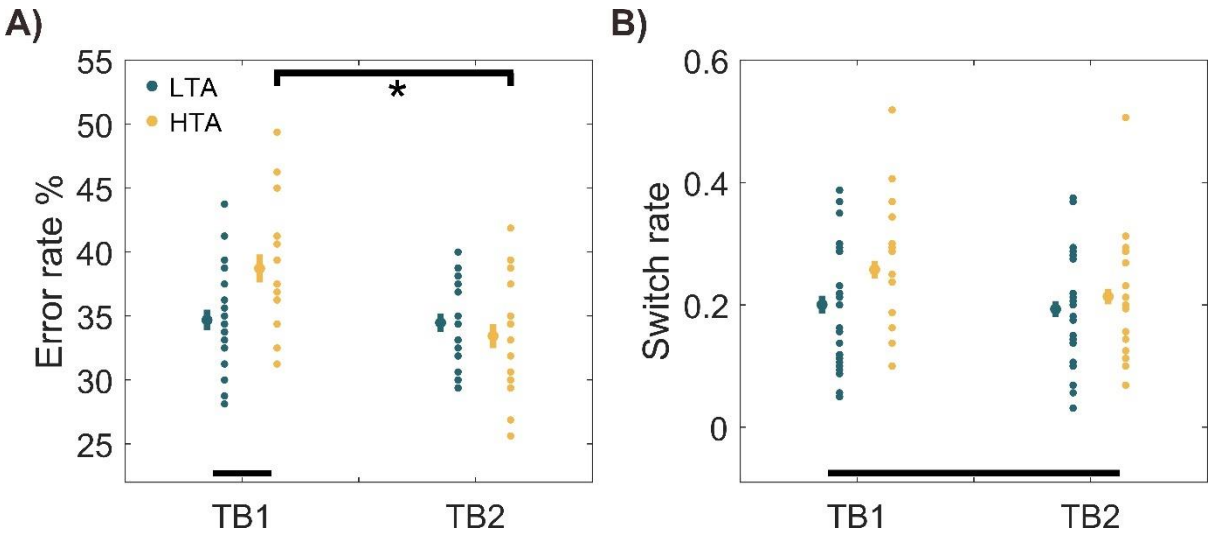


Figure 1. Trait anxiety modulates the percentage of reward learning errors and the empirical switch rate. **A)** High trait anxiety (HTA, yellow) increases the percentage of errors during reward learning relative to low trait anxiety (LTA, dark blue, $P = 0.001$). Each group is presented using the average with SEM bars and to the right are individual data points to display dispersion. LTA significantly increased the error rate in TB1 but not TB2 ($P_{FDR} < 0.05$) and significantly dropped from TB1 to TB2 in HTA ($P_{FDR} < 0.05$). Significant between-group differences are represented by black bars on the x-axis and significant within-group differences are identified by black bars and a '*' symbol at the top. **B)** The empirical switch rate was calculated as trial-to-trial response switches from one image (blue, orange) to the other. HTA anxiety significantly raised the empirical switch rate relative to LTA ($P = 0.01$). Observed across both task blocks on the x-axis, the empirical switch rate appears larger in the first task block in HTA, which aids in the interpretation of the results that HTA increased volatility estimates and that the model-free error rate was significantly higher in HTA in TB1 but not TB2.

Despite a significantly higher overall switch rate in HTA than LTA across both blocks (main effect of Group), in **Figure 1B**, we observe a higher empirical switch rate in TB1 in the HTA group compared to LTA and TB2. This observation can help in our interpretation of the model-free results about the higher percentage of errors in HTA in TB1 and the overestimation of μ_3 in HTA relative to LTA (main effect of Group). We interpret this as HTA estimating more changes in the reward environment, which in the $HGF_{\mu 3}$ response model leads to more exploratory behavioural responses, observed initially in TB1 in increased switching responses. Higher volatility estimates in HTA were also associated with a higher percentage of errors in TB1. After, in TB2, HTA appears less 'explorative' and more 'exploitative' (fewer switches trial to trial, see **Figure 1B**), which is also reflected in their percentage of errors decreasing to an equivalent level with LTA. As an additional check, in **Supplementary Figure 2** we present the empirical switch rate from each contingency block (0.9/0.1, 0.7/0.3 and 0.5) between groups.

3.4.1.2 Empirical stay/switch rate by outcome

Increased volatility and high learning rates may lead to participant responses exhibiting an interaction between response (stay, switch) and outcome (win, lose, see Huang et al., 2017; Jiang et al., 2018; Xia et al., 2021). To test this, we calculated the stay/switch rate following a trial's outcome (win, lose). The interaction effect using a three-way repeated-measures ANOVA between the factors Group (LTA, HTA), Response (Stay, Switch), and Outcome (Win, Lose) was significant ($F(1,41) = 5.9$, $P = 0.02$, see **Supplementary Figure 3**). Post-hoc testing using pair-wise permutation tests revealed that HTA are significantly more likely to switch after a win trial (mean 0.06, SEM 0.09) compared with LTA (mean 0.21, SEM 0.01, $P_{FDR} < 0.05$, $\Delta = 0.65$, CI = [0.51, 0.85]). Also, HTA had a significantly lower stay rate following wins (mean 0.56, SEM 0.01) relative to LTA (mean 0.6, SEM 0.01, $P_{FDR} < 0.05$, $\Delta = 0.68$, CI = [0.52, 0.87]). In addition, the HTA group had a significantly higher switch rate after a lose trial (mean 0.21, SEM 0.01) compared with LTA (mean 0.12, SEM 0.01, $P_{FDR} < 0.05$, $\Delta = 0.72$, CI = [0.56, 0.86]). HTA also stayed significantly less after a lose trial (mean 0.18, SEM 0.01) relative to LTA (mean 0.21, SEM 0.01, $P_{FDR} < 0.05$, $\Delta = 0.68$, CI = [0.53, 0.84]).

Analysis of the empirical switch rate as a function of the preceding trial's outcome reveals that the HTA group switches more from both win and lose trials while also staying less in response to win and lose trials. We interpret this according to the results that HTA exhibits higher estimates of volatility and the $HGF_{\mu 3}$ response model leading to more 'exploratory' switching behavioural responses, even in cases in which switching is unwarranted (e.g. following a win outcome).

3.4.2 Amplified informational uncertainty about the stimulus outcomes in high trait anxiety

Informational (belief) uncertainty about the stimulus outcomes, σ_2, depends on the volatility estimate μ_3 from the preceding trial (among other quantities, see Eq. 11 and 13 in Mathys et al., 2014). Factorial analysis with synchronised rearrangements showed a significant main effect of the factor Group ($P = 0.02$), but no Block ($P = 0.39$) or interaction effect ($P = 0.81$). When comparing the blocks-average σ_2 between groups, pair-wise permutation tests revealed the HTA group (mean 1.53, SEM 0.14) significantly increased σ_2 relative to LTA (mean 1.21, SEM 0.13, $P = 0.01$, $\Delta = 0.71$, CI = [0.55, 0.83], see **Figure 2**). The higher belief uncertainty about the stimulus outcomes in the HTA group translates to a greater influence on the update equations for beliefs about x_2 and consequently faster learning on that level—potentially driven by higher volatility estimates.

3.4.3 Environmental uncertainty is higher in high trait anxious individuals

Uncertainty about the task environment relies on the quantities ω_2 (tonic volatility) and the trial by trial estimates of volatility $\mu_3^{(k-1)}$ (see Eq. 13–14 in Mathys et al., 2014), here as in **Chapter 2** κ was set to one). In that way, environmental uncertainty is driven by changes to the task environment, with subjectively estimated higher levels of volatility generating more environmental uncertainty. Permutation-based factorial testing revealed a significant main effect of the factor Group (P = 0.03), but not for the Block (P = 0.65) or interaction effect (P = 0.85). Follow-up pair-wise permutation tests of the Group factor averaging over task blocks demonstrated that the HTA group had greater environmental uncertainty (mean 0.49, SEM 0.07) when compared with LTA participants (mean 0.34, SEM 0.06, P = 0.03, Δ = 0.67, CI = [0.54, 0.83], **Figure 2**). There was, however, no significant difference between groups in the related parameter ω_2, the expected tonic uncertainty on level 2.

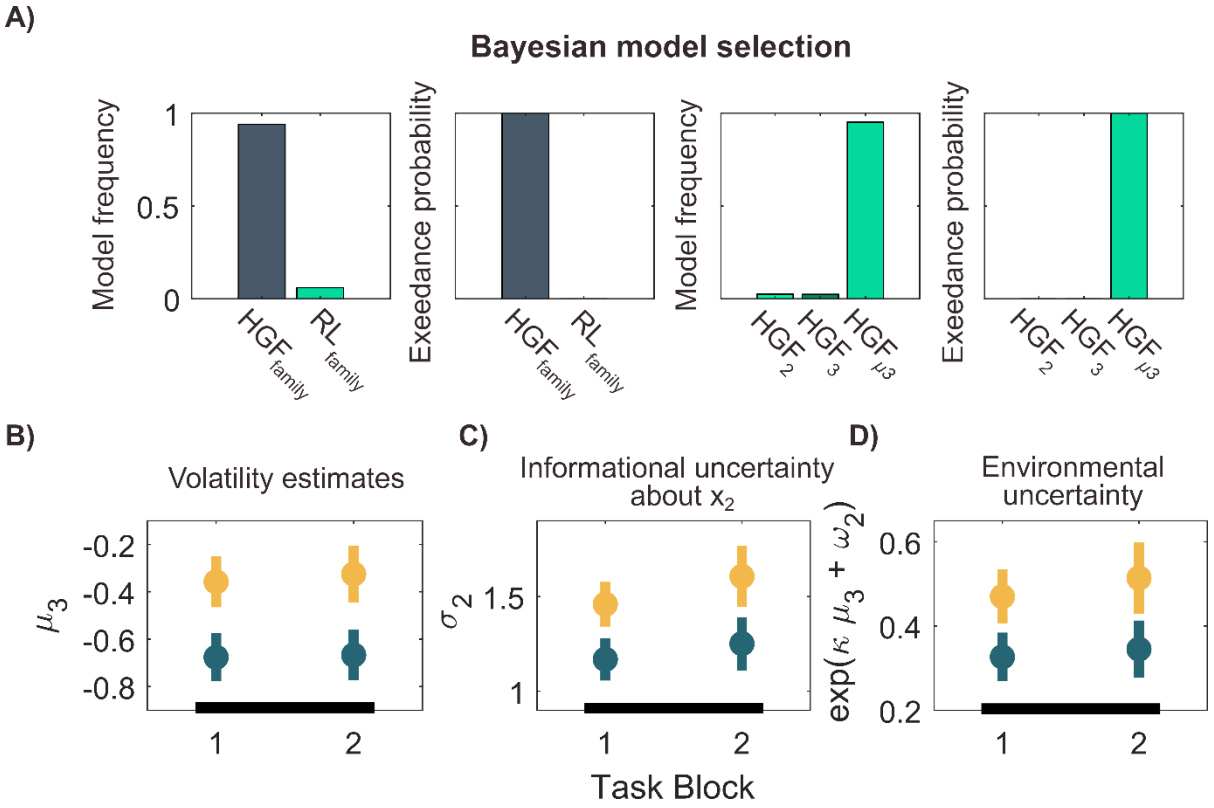


Figure 2. Bayesian Model Selection and Hierarchical Gaussian Filter Results. **A)** Bayesian model selection (BMS). The panels to the left show the model frequency and exceedance probability for the family of models 'HGF Fam' consisting of the 2-level HGF (HGF_2), the 3-level HGF (HGF_3), and the 3-level HGF informed by trial-wise estimates of volatility ($HGF_{\mu 3}$) given in dark blue. The family of reinforcement learning models 'RL FAM' (Rescorla-Wagner, RW, Sutton K1, SK1, presented in green). The family of HGF models better describes task learning behaviour. In the right-hand panels are the model frequency and exceedance probability comparing the three HGF models (HGF_2: light green,

HGF_3: dark green and HGFμ_3: green). The HGFμ_3 model best explained the data. **B)** Volatility estimates (μ_3) in HTA (yellow) were significantly higher relative to LTA (dark blue, $P = 0.003$). However, there was no significant Block ($P = 0.84$) or interaction effect ($P = 0.92$, given by black bars). **C)** Informational (estimation) belief uncertainty about the stimulus outcome tendency was raised in HTA compared with LTA (significant factor Group, P = 0.02; shown using black bars on the x-axis). **D)** The HTA group were also significantly more uncertain about the environment ($P = 0.03$, as given by black bars). However no significant differences were found in uncertainty about volatility (σ_3) or the tonic learning rates at levels 2 (ω_2) and 3 (ω_3).

3.5 Time-frequency responses

3.5.1 General modulation of spectral power

At rest, we found no significant differences in the average raw spectral power between HTA and LTA groups ($P > 0.05$, cluster-based permutation test, **Supplementary Figure 4)**. In contrast to previous work in high trait anxiety, this result suggests that our HTA group's general spectral profile at rest was not different from the low trait anxious group (Knyazev et al., 2004).

3.5.2 Precision-weighted prediction errors about stimulus outcomes

Within-subject analysis showed there was a significant negative modulation of the oscillatory responses by the pwPE regressor in 8–30 Hz in the LTA group relative to baseline (one negative cluster, $P = 0.001$, two-sided test, FWER-controlled, see **Figure 3**). The significant cluster extended within 500–900 ms in 10–20 Hz and had a centroparietal distribution. In the HTA group, a significant decrease relative to baseline was found in 8–30 Hz (one negative cluster, $P = 0.001$, two-sided test, FWER-controlled, see **Figure 3**). The negative modulation of the significant cluster was within 500–760 ms in 10–20 Hz and was distributed across centroparietal sensors. No significant effect of pwPEs on the gamma-band oscillatory responses relative to baseline were found (see **Supplementary Figure 5**).

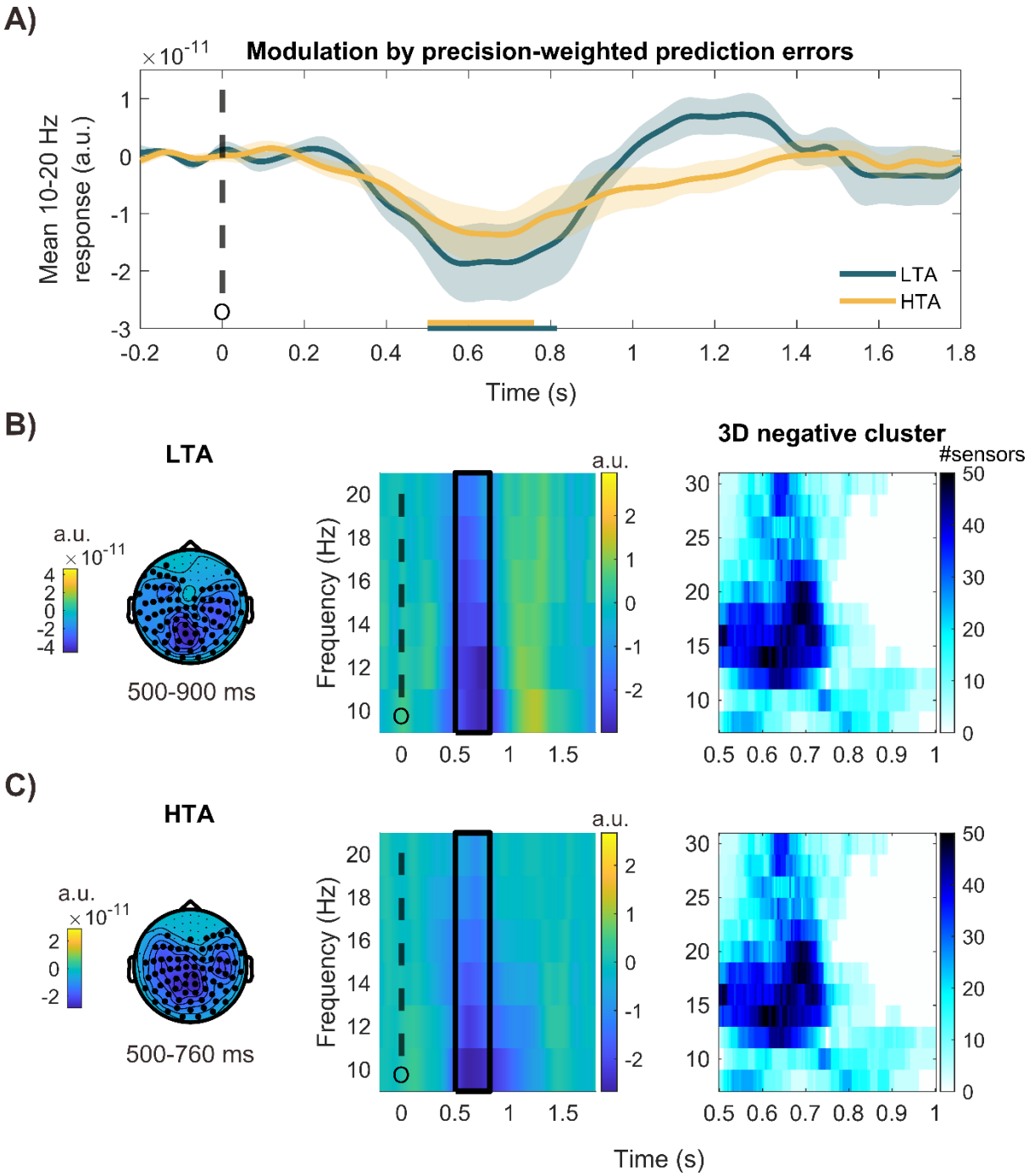


Figure 3. Oscillatory activity between 10–20 Hz is modulated by precision-weighted prediction errors about stimulus outcomes: within-group effects. **A)** Time course of the average oscillatory response (10–20 Hz) to pwPEs in each group (LTA, dark blue; HTA, yellow), given in arbitrary units (a.u.). Dependent-samples significant clusters in each group are given by horizontal bars on the x-axis in their respective colours. **B)** LTA sensor correlates of oscillatory activity modulated by pwPEs. One negative cluster was found in 10–20 Hz spanning 500–900 ms relative to baseline (P = 0.001). On the left is the topographic distribution of this effect, primarily over posterior centroparietal sensors, but also in temporal and fronto-central regions. On the right are time-frequency images for the cluster averaged across the sensors (the outcome is denoted by the black dashed line; black squares show the significant cluster) and the 3D negative cluster showing the total number of sensors (#sensors) forming the significant cluster. **C)** Follows the same format as (B) but showing the HTA group. Similarly we observed a significant negative difference from baseline in 10–20 Hz between 500–760 ms relative to baseline (P = 0.001) in posterior and centroparietal sensors.

At the between-subject level, an independent sample test between 8–30 Hz revealed a significant decrease in oscillatory responses in the HTA group relative to LTA (one negative cluster, *P* = 0.01, two-sided test, FWER-controlled). The latency of the significant cluster was within 1100–1330 ms in 8–16 Hz and was widely distributed across central and sensorimotor and right frontal sensors (see **Figure 4**). There was no significant difference between groups in gamma-band modulation by pwPEs.

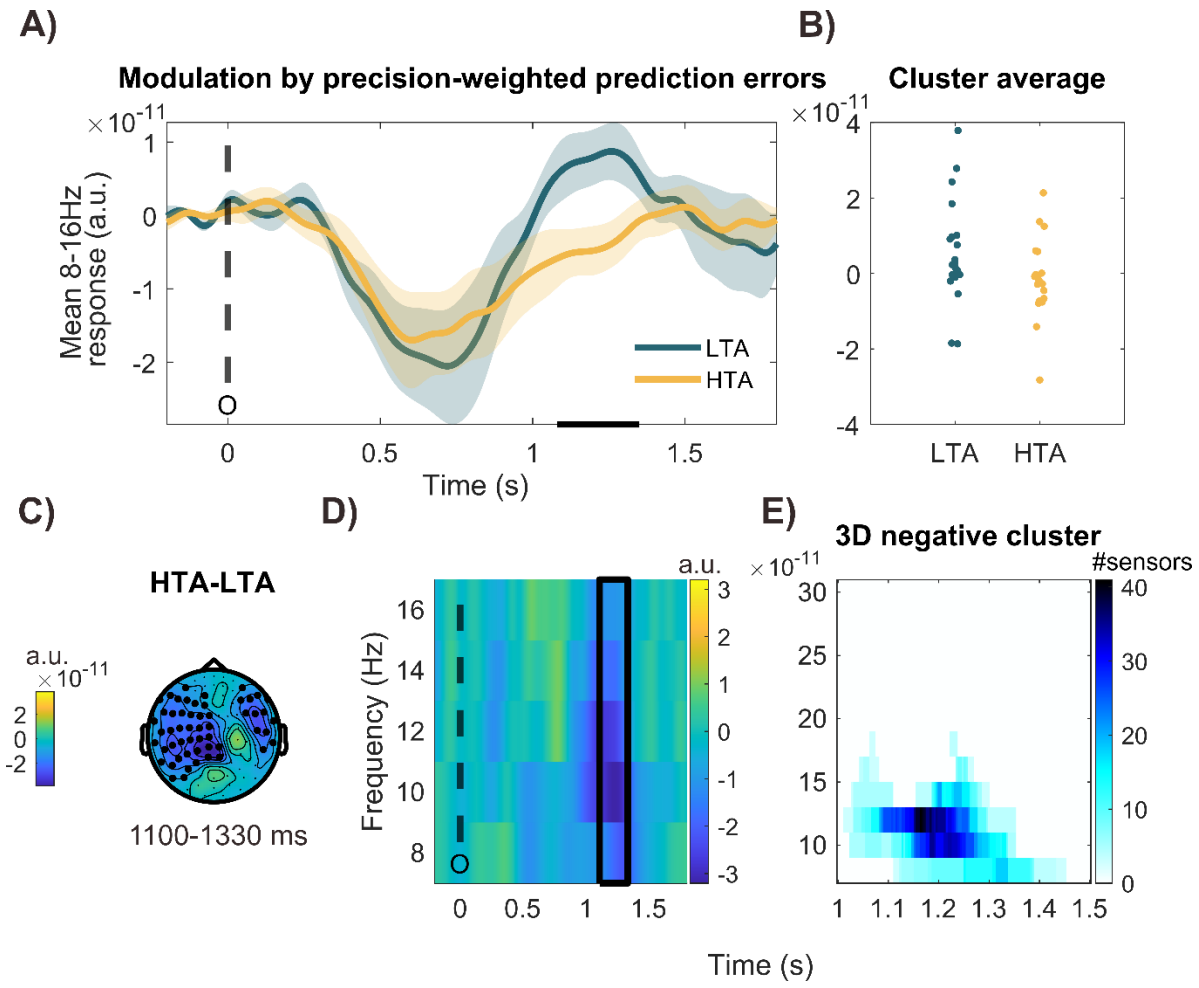


Figure 4. Between-group differences in alpha-beta oscillatory activity modulated by precision weighted prediction errors about stimulus outcomes. **A)** Time course of the average 8–16 Hz oscillatory response to pwPEs in each group (LTA, dark blue; HTA, yellow), given in arbitrary units (a.u.). The time of the significant cluster is shown using a horizontal black bar on the x-axis. **B)** The average 8–16 Hz oscillatory response (a. u.) in the cluster to pwPE for individuals in the LTA (dark blue) and HTA (yellow) group. **C-D)** Between-group differences in the oscillatory activity (8–16 Hz) modulated by pwPEs about stimulus outcomes (one significant negative cluster between 1100–1330 ms; *P* = 0.01, FWER-controlled). On the left **C)** is the topographical distribution and on the right **D)** is the time-frequency image for the cluster averaged across the sensors (the outcome is denoted by the black dashed line; black squares show the significant cluster) **E)** The total number of sensors (#sensors) contributing to the significant 3D cluster at each frequency and time point.

3.5.3 Stimulus-locked predictions about reward tendency

A within-subject level test between 8–30 Hz revealed a significant positive modulation of oscillatory activity by the prediction regressor in the LTA group relative to baseline (one positive cluster, P = 0.024, two-sided test, FWER-controlled, see **Figure 5**). The significant cluster was within 552–845 ms in 14–20 Hz and had a posterior centroparietal distribution. There was no difference from baseline in oscillatory responses modulated by the prediction regressor in the HTA group. Independent samples tests also showed there was no between-group difference in oscillatory activity modulated by the prediction regressor. Of note, the statistical tests of predictions in alpha-beta activity were not confounded by modulation from the response regressor (motor responses) as we independently accounted for activity associated with the response regressor. In fact, the response regressor in the stimulus-locked representation led to different effects as those reported for the prediction regressor. This was assessed in a control analysis using independent-samples cluster based statistical tests where we found a significant increase in oscillatory responses from the response regressor in HTA relative to LTA (one positive cluster, P = 0.01, two-sided test, FWER-controlled, see **Supplementary Figure 6**). The significant cluster was within 350–640 ms in 14–26 Hz across sensorimotor and central sensors. We also observed a prominent within-group classic sensorimotor beta rebound effect in both groups locked to the stimulus (**Supplementary Figure 7**).

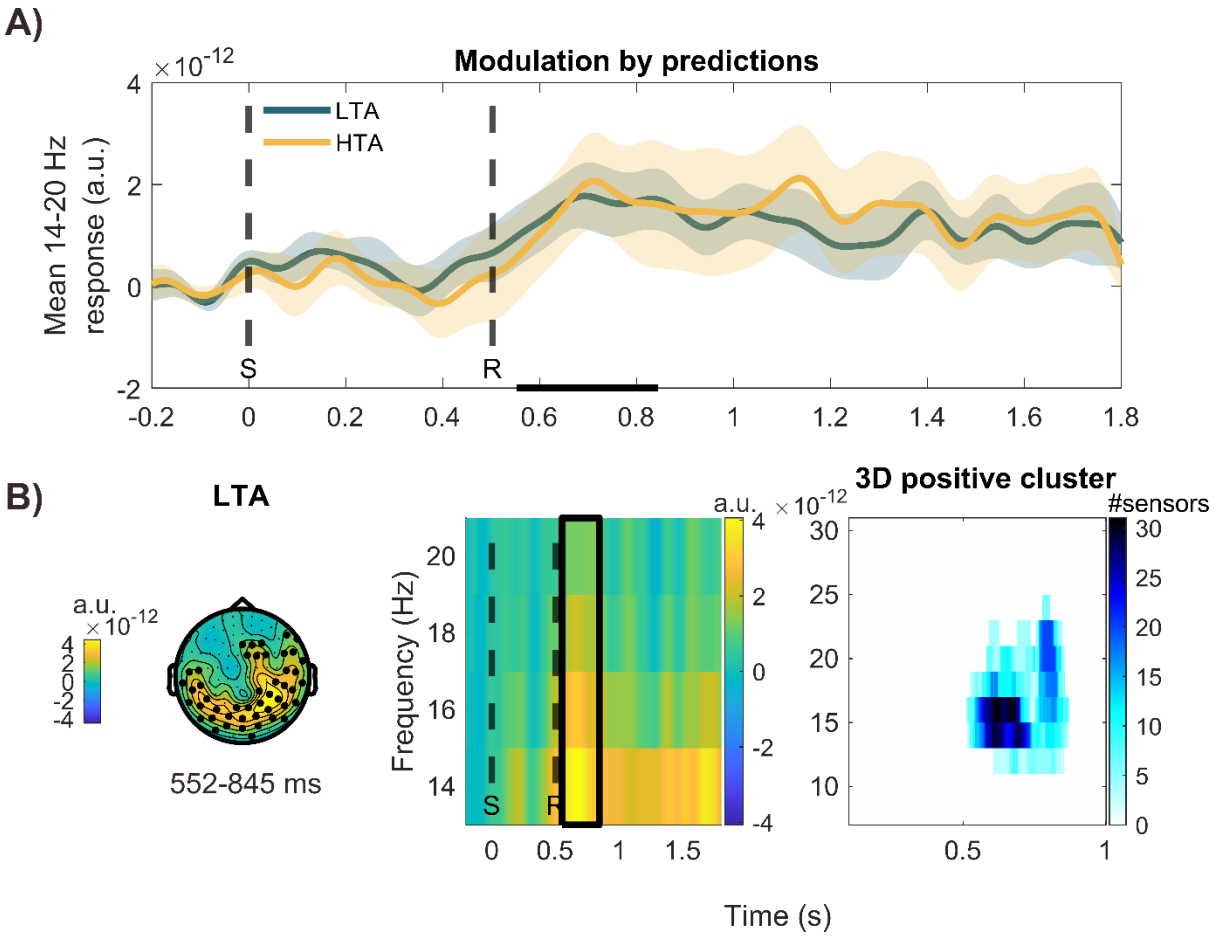


Figure 5. Within-group modulation by predictions to oscillatory activity between 14–20 Hz. **A)** Time course in arbitrary units (a.u.) of the average oscillatory response between 14–20 Hz by the prediction regressor in each group (LTA, dark blue; HTA, yellow). The dependent-samples significant cluster for the LTA group is shown using a horizontal bar on the x-axis. In dashed black lines are given the time of the stimulus 'S' and the average response time 'R'. **B)** Sensor-level correlates of oscillatory responses modulated by predictions in LTA. One negative cluster was discovered in 14–20 Hz within 552–845 ms post-stimulus, relative to baseline (P = 0.024). Left: The topographic distribution of this effect is in posterior centroparietal sensors. Right: Time-frequency images for predictions about the stimulus outcomes averaged across the cluster sensors (dashed black lines give the time of the stimulus and average response; black squares show the significant cluster) and the 3D negative cluster showing the total number of sensors (#sensors) comprising the significant cluster.

4. Discussion

Current theories of affective disorders conceptualise some of the psychiatric symptoms as divergent hierarchical Bayesian inference, described by difficulties estimating uncertainty and maintaining an equilibrium between prior beliefs (or predictions) and sensory input (Paulus & Yu, 2012; Pulcu & Browning, 2019). At the neural level, the computational quantities of hierarchical Bayesian inference are thought to be encoded by distinct neural oscillations, with predictions expressed in low frequency changes and PEs expressed by high frequency changes (Bastos et al., 2012).

We provide evidence supporting these predictions in a subclinical group with high trait anxiety using MEG data and a probabilistic reward learning task. We first tested for the effect of high trait anxiety (HTA) on behavioural responses relative to low trait anxiety (LTA). We found that the overall percentage of errors during reward learning in HTA was higher, driven by performance in the first experimental task block. Also, HTA raised the average probability of switching responses from one choice to the alternative, observed primarily in the first task block.

Following model-free analysis, we looked for between group differences in model estimates of beliefs and belief uncertainty considered vital to hierarchical Bayesian inference. The reward learning data were best explained by a Bayesian model that depends on trial-wise belief estimates of volatility. HTA raised volatility estimates across both task blocks, which was associated in the first block with an increase in the probability of switching responses, a model-free metric of stochastic response behaviour. Combined, these results suggest that choice probability in the HTA group is more exploratory. Unlike in our state anxiety group in **Chapter 2** (Hein et al., 2021), both environmental uncertainty and informational (estimation) uncertainty about the stimulus outcomes were higher in HTA, which is related to amplified volatility estimates and consequently larger model updates about stimulus outcomes.

We then looked at the oscillatory correlates of precision-weighted prediction errors (pwPEs) and predictions about the stimulus outcomes. Trial-wise model estimates of these computational learning quantities were used as parametric regressors in a convolution model to explain modulations in the amplitude of oscillatory MEG activity (Litvak et al., 2013). In contrast to our previous work in state anxiety (**Chapter 3**), we found that HTA dampened beta oscillations (8–16 Hz) during the encoding of pwPEs between 1100–1330 ms. Interestingly, this is consistent with our modelling results. HTA were more uncertain about the environment

and overestimated volatility, which raises belief uncertainty and learning rates on the second level. This could indicate more stimulus-driven processing, enhancing the role of PEs in updating predictions (Bauer et al., 2014; Sedley et al., 2016). However, we did not find that pwPEs modulated gamma oscillatory activity. Also, the encoding of predictions was associated with increased beta activity (14–20 Hz, 552–845 ms post-stimulus) from baseline, but only in the LTA group.

As with our previous work (**Chapter 3**), these results provide evidence supporting generalised predictive coding (PC, Brown & Friston, 2013; Feldman & Friston, 2010). In PC, attention regulates the precision of PEs by modulating synaptic gain. Functionally, this either inhibits new sensory information that elicits PEs (encoded in alpha/beta oscillations) or heighten their impact on belief updates (encoded in gamma oscillations, anticorrelated with alpha/beta, see Hoogenboom et al., 2006; Lundqvist et al., 2016, 2018, 2020; Miller et al., 2018; Potes et al., 2014). Taken as a whole, our results expand the computational work on maladaptive learning in anxiety but suggest that the effects of anxiety on higher-order cognition are not homogeneous, depending heavily on context. Even within the learning domain, the context (e.g. rewarded or punishing outcomes) and the type of anxiety (state, trait, or clinical) play a core role in the behavioural and neural responses elicited (**Chapter 3** and Browning et al., 2015; Hein et al., 2021; Piray, Ly, et al., 2019; Pulcu & Browning, 2019; Sporn et al., 2020). To understand how anxiety research can inform diagnosis and treatment, further empirical work is necessary to reconcile or connect the seemingly incompatible findings in the different realms of learning and variety of anxious experiences (Grillon et al., 2019; Grupe & Nitschke, 2013; Hein et al., 2021; Pulcu & Browning, 2019; Robinson et al., 2019; Sporn et al., 2020).

Trait anxiety alters computations of uncertainty during reward-based learning

Here, we provide supporting evidence for the prediction that trait anxiety impedes reward learning. Prior work showed a link between high levels of trait anxiety and difficulties in higher-order executive tasks (de Visser et al., 2010; Hartley & Phelps, 2012; Miu, Heilman, et al., 2008; Remmers & Zander, 2018), and also that high levels of trait anxiety present distinct challenges to learning in changing task environments (Browning et al., 2015; Huang et al., 2017). An essential contribution of this paper is that we bridge these findings. We show that high levels of trait anxiety impairs performance in an executive task and is associated with biased uncertainty estimates, especially concerning changes to the environment. Our results indicate that misestimation of uncertainty takes centre stage in anxiety, and in our task, gave rise to initial suboptimal switching behaviours that vitiated reward learning performance. The

striking positive side for our HTA group is that despite initially performing worse than LTA, HTA learned quickly to meet similar performance levels by the second task block.

Biased estimates of environmental volatility influence lower-level prediction errors through precision weights (that is to say, raised volatility estimates increase lower-level belief uncertainty and lead to greater model updates). As such, larger updates about the stimulus outcomes in our HTA group were driven by both increased environmental uncertainty and informational uncertainty relative to LTA. However, this leaves the unresolved question of why increased learning about stimulus outcomes in HTA does not equate to an overall boost to reward learning performance.

The LTA anxiety group learned well in the first task block, but not at an optimal level. HTA were more uncertain about the changing environmental statistics, and thus used larger model updates—for the first task block at least. Biased estimates of uncertainty in HTA may have driven suboptimal switching behaviour (with higher learning rates in more stable blocks), making it an acute challenge to establish if the outcome of a choice was by chance (expected uncertainty associated with outcome noise) or by some meaningful statistical instability (unexpected environmental uncertainty). This relates to the work of Browning et al. (2015), who showed that highly trait anxious individuals exhibit difficulties fitting a learning rate to changes in the task environment. In the aversive learning context of Browning et al. (2015) this was overall a maladaptive response, with the authors concluding that the HTA group were not sensitive to environmental change (lower pupil response and deficit in using volatility to inform task choices relative to controls). However, Browning et al. (2015) did not model estimates of volatility or expected outcome unpredictability. More recent work has shown how anxiety is linked with the inflated valuation of information—driving exploratory uncertainty reduction behaviours (Aberg et al., 2021; Bennett et al., 2021).

In our task, as the subjectively estimated degree of volatility dictates the learning rate, the learners are required to infer the degree of environmental change dynamically (in spite of our relatively set level of ground truth changes). This learning is expressed by monotonic changes in log-volatility estimates. In HTA, the increased volatility estimates and switching between probabilistic outcomes in the first task block eventually delivered success by the second task block—with HTA performing equivalently to LTA. We speculate that the previously hypothesised widespread elevated prior expectations for environmental change in anxiety (Bishop & Gagne, 2018) may relate to our HTA group needing to learn the stable volatility changes in our task through initial maladaptive exploration. Their LTA counterparts may either not expect a particular level of change or respond less rigidly to prior expectations and were

thus better able to estimate volatility and outcome uncertainty accurately, and adapt their learning rate accordingly. This account goes some way to explaining why our HTA group experienced an initial deficit in reward learning performance that later reached level terms with LTA.

One key insight from our results is that inflated volatility estimates in HTA led to nosier switching responses. The misestimation of uncertainty that is thought central to the expression of anxiety (Pulcu & Browning, 2019) implies a malfunction at higher levels of the belief hierarchy that represent the structure of the environment. In particular, this disposition for excessively inferring the probability that a true change to the statistics of the world has occurred could instigate a policy of heightened prediction error transmission (Mathys et al., 2011; Weber et al., 2020), as observed in other learning domains (Robinson et al., 2019). This may also require particularly precise beliefs at higher levels to offset this tendency.

To complement our finding of higher volatility estimates in the HTA group, we analysed the empirical switch rate as a model-free proxy indication of exploratory choice behaviour. We would expect a noisier mapping of beliefs to response in HTA, given that our winning model describes choice probability as dependent on dynamic volatility estimates and that higher levels of volatility raise the learning rate at lower levels. We found a higher empirical switch rate in HTA which we observed was primarily in the first task block. HTA also increased switching responses following rewarded outcomes and increased staying after unrewarded outcomes compared with LTA. This may explain the initially poorer reward learning performance in HTA, as higher levels of switching response combined with a maladaptively high learning rate would make it difficult to learn the underlying probabilistic contingencies and potentially may make distinguishing between meaningful environmental changes and outcome noise challenging. More generally, this finding is consistent with evidence from task/set switching and reversal learning tasks where high trait anxiety is associated with inflexible responses poorly adapted to changes in the task (Ansari et al., 2008; Derakshan et al., 2009; Edwards et al., 2015; Gustavson et al., 2017; Wilson et al., 2018).

Commonly, this cognitive inflexibility in highly anxious people is explained by issues with attentional control, especially when distracted (Berggren & Derakshan, 2013). However, in our probabilistic reward learning task, there are no distractors, and we have consistently shown using reaction time data that attentional deficits do not explain impaired reward learning performance (see **Chapters 2–4** and Hein et al., 2021; Sporn et al., 2020). Browning et al. (2015) also provided evidence for learning difficulties in HTA during an aversive learning task by establishing there were no differences in their learning rate across the changing

contingency phases. This describes a suboptimal adaptation to environmental uncertainty. Similarly, recent empirical work by Huang et al. (2017) using a volatile reward-based visual search task reported both an inflated lose-shift rate and an inflexible adjustment of learning rates in HTA. In that study, the HTA group suboptimally used a higher learning rate. Both studies, however, did not model volatility estimates. Speculatively, the higher learning rate and lose-shift behaviour in Huang et al. (2017) resembles our finding of amplified volatility estimates in HTA, leading to more exploratory and shifting responses.

High trait anxiety has also been associated with fewer lose-shift responses in an earlier study using a two-choice prediction task (Paulus et al., 2004). A more recent experiment by Jiang et al. (2018) demonstrated that high trait anxious individuals initially make suboptimal choices about rewards in a probabilistic learning task, providing evidence that negative feedback reduced the feedback-related negativity (FRN) amplitude. Strikingly, the authors note, similar to our findings, that initially, their HTA group learned the environmental statistics slower than controls, but that HTA task performance reached an equivalent level later in their task (Jiang et al., 2018). In addition, Xia et al. (2021) reported that HTA impeded learning in a reversal-learning task with both reward and loss outcomes. Notably, in that study, HTA individuals switched less from trials following a loss relative to controls. The authors also explain this rigid choice behaviour as diminished sensitivity to negative feedback, supported by evidence of a dampened FRN amplitude. Our data add to these findings by showing that switching responses and exploratory behaviour during suboptimal learning relate to maladaptive volatility estimation biases in high trait anxiety.

Lower frequency encoding of precision-weighted prediction errors and predictions

In **Chapter 3**, we found that state anxiety increased beta activity during the encoding of pwPE and predictions relative to controls. The results followed from the findings reported in **Chapter 2**, where anxious individuals had more precise beliefs about the reward outcome tendency, reducing the learning rate and diminishing learning (Hein et al., 2021). This motivated our hypothesis here that HTA would show a similar behavioural and oscillatory profile in response to our reward-based learning task. In line with the Bayesian PC account, we expected higher levels of alpha and beta encoding pwPEs and predictions. More specifically, as we expected reduced estimation uncertainty about stimulus outcomes to lead to smaller pwPEs (lowering learning rates on that level), we predicted increased alpha and beta during the processing of pwPEs in our HTA group. We also explored related gamma oscillatory activity. Heightened gamma and diminished alpha/beta oscillations have been attributed to encoding pwPE (Auksztulewicz et al., 2017; Bastos et al., 2020), with gamma activity anticorrelated to

alpha/beta activity across the cortex (Hoogenboom et al., 2006; Lundqvist et al., 2016, 2018, 2020; Miller et al., 2018; Potes et al., 2014).

We found that the encoding of pwPEs in oscillatory activity decreased alpha/beta oscillations (10–20 Hz) from baseline during 500–900 ms post-outcome in both groups' central and parietal sensors. This finding closely aligns with the temporal and spatial results from our previous work in state anxiety in **Chapter 3**, providing more evidence that a reduction to lower frequency neural oscillations encodes pwPEs independently of anxiety. High trait anxiety later attenuated the alpha/beta oscillatory response (8–16 Hz) relative to LTA between 1100–1330 ms over central and sensorimotor sensors. This between-group difference goes in the opposite direction to those found in beta activity for state anxiety reported in **Chapter 3** and in our recent work in reward-based motor learning—despite sharing similar spatial and temporal dimensions (beta power and burst events, see Sporn et al., 2020). We were unsuccessful in finding complementary increases to gamma oscillations during the processing of pwPEs, as has been reported in previous perceptual/sensory processing work (Auksztulewicz et al., 2017).

Additionally, in the prediction regressor, we observed a significant within-group increase in alpha/beta (14–20 Hz) oscillatory activity exclusively in the LTA group between 552–845 ms post-stimulus in posterior centroparietal sensors. In contrast to our previous results (**Chapter 3**), we found no difference in oscillatory activity processing predictions relative to baseline in HTA, and no significant between-group changes. As MEG provides better spatial resolution of source localisation (2–3 mm, Singh, 2014), our future work will focus on using individual MRI scans to resolve three-dimensional magnetic source imaging (MSI)—determining the active site of origin of the changes in spectral correlates due to pwPEs.

We built here on recent progress in the modelling of computational learning quantities and their representation in neural responses, providing evidence for the altered encoding of pwPEs in high trait anxiety (Diaconescu, Mathys, et al., 2017; Iglesias et al., 2013; Kolossa et al., 2015; Mars et al., 2008; Nassar, Bruckner, et al., 2019; Nassar, McGuire, et al., 2019; Powers et al., 2017; Stefanics et al., 2018; Weber et al., 2020). Although our between-group result appears incompatible with our previous finding in **Chapter 3**, the finding is consistent with the hierarchical Bayesian model analysis reported here. Unlike our state anxiety group in Hein et al. (2021), HTA were characterised by higher estimates of task volatility (μ_3) and excessive environmental uncertainty ($\exp(\kappa\mu_3^{(k-1)} + \omega_2)$). Combined, this misestimation of uncertainty about the environment amplified uncertainty about the probabilistic reward outcomes (σ_2), driving larger updates on that level. Given this opposite direction of effect on uncertainty

estimates in HTA compared with state anxiety, one would expect a lower expression of alpha/beta activity in a similar time window post-outcome. This is precisely the result we observed, with pwPEs in HTA represented by a significant reduction in 8–16 Hz oscillatory activity relative to LTA. We speculate that this lowered precision in the posterior belief about the reward tendency is represented by attenuated 8–16 Hz activity in HTA during the encoding of pwPEs, reflecting larger updates to predictions.

In visual cortex studies, robust changes in gamma oscillatory responses have been observed (Arnal & Giraud, 2012; Bastos, Litvak et al., 2015; Buffalo et al., 2011; Gould et al., 2011; Michalareas et al., 2016; Roberts et al., 2013; Xing et al., 2012). Modulation to gamma oscillations by PEs has also been shown during learning in perceptual/sensory based tasks using MEG (Auksztulewicz et al., 2017) and electrocorticography (Bastos, Vezoli et al., 2015). Yet, for reward-based pwPEs, we did not observe changes to gamma oscillations by pwPE—however, see Palmer et al. (2019) for a lack of gamma-band effects during sensory and sensorimotor learning. Our lack of gamma frequency results here and in **Chapter 3** suggests a more complex relationship between reward-based pwPEs and observable cortical oscillatory changes, rendering it difficult to interpret negative results.

As discussed in **Chapter 3**, neural evidence of cortical gamma responses modulated by reward prediction errors have been shown in areas of research outside of the sensory/perceptual studies cited above (Berke, 2009; Ellwood et al., 2017; Domenech et al., 2020; Lohani et al., 2019). Potentially, as our reward-based pwPEs do not encode reward PEs (the unsigned difference between received and predicted reward) the phasic responses are not robust enough to be detected. Moreover, our MEG based approach may not be sufficiently sensitive to pick up on these gamma-band oscillatory signals. Future work may profit from using alternative analysis approaches (independent component analysis, see Dickinson et al., 2015) or design choices (using reward PEs separately) to increase the signal of reward-based pwPEs. Use of local field potential recordings in humans, if possible, could also represent an excellent means to test the oscillatory proposals of Bayesian PC in reward learning.

As previously mentioned, we hope to capitalise on source localisation techniques using individual MRI scans to reveal where the processing of pwPEs evokes activity in deep brain sources. Still, future work that follows up on our current findings could also benefit from combining models able to infer pwPE and prediction signalling in cortical hierarchies from combined EEG and high-resolution laminar fMRI data (Stephan et al., 2019). Inserting the trajectories of HGF estimates into laminar dynamic causal models could supply the missing spectral evidence linking misestimation of uncertainty in anxiety to hierarchical Bayesian PC

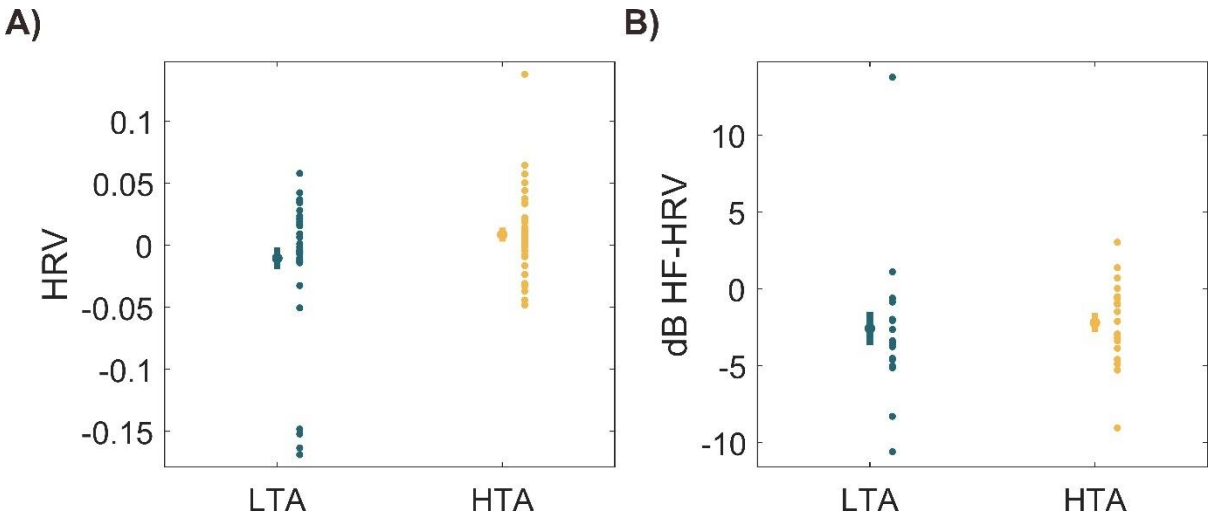
processing hypotheses about altered ascending connections (Heinzle et al., 2016; Scheeringa & Fries, 2019).

Recent work has provided evidence that high levels of trait anxiety attenuate the pupil response during the highly volatile blocks during probabilistic aversive learning (Browning et al., 2015)—connected with suboptimal adaption of learning rates to changes in the environment. Pupil changes have been associated with phasic responses of the noradrenaline system, representing surprise about changing environmental conditions in both healthy and clinical populations (Jepma & Nieuwenhuis, 2011; Lawson et al., 2017; Muller et al., 2019; Pulcu & Browning, 2017; Vincent et al., 2019). Others have reported dynamic changes to the learning rate in the dorsal anterior cingulate cortex (dACC, Behrens et al., 2007, 2008). How anxiety modulates the learning rate and which computational and neural processes are involved remains an understudied area of research (Piray, Ly, et al., 2019). By combining hierarchical Bayesian modelling of behaviour with fMRI and EEG recordings, future reward-based studies designed to include marked changes of volatility across time could determine the spectral correlates of hierarchically-related pwPE signalling and associate them with previous neuroimaging evidence of learning rate changes. Accordingly, one could test the prediction that trait anxiety attenuates responses from the cingulate and prefrontal cortex while learning about reward, as these regions are in a core network known to subserve disruptions in anxiety and undergird the processing of both volatility and pwPE about stimulus outcomes (Bishop et al., 2004; Bishop, 2009; Grillon et al., 2019; Grupe & Nitschke, 2013; Iglesias et al., 2013). Probing these questions could bridge the misestimation of uncertainty reported here in trait anxiety and the computational difficulties reported in Browning et al. (2015) and Huang et al. (2017).

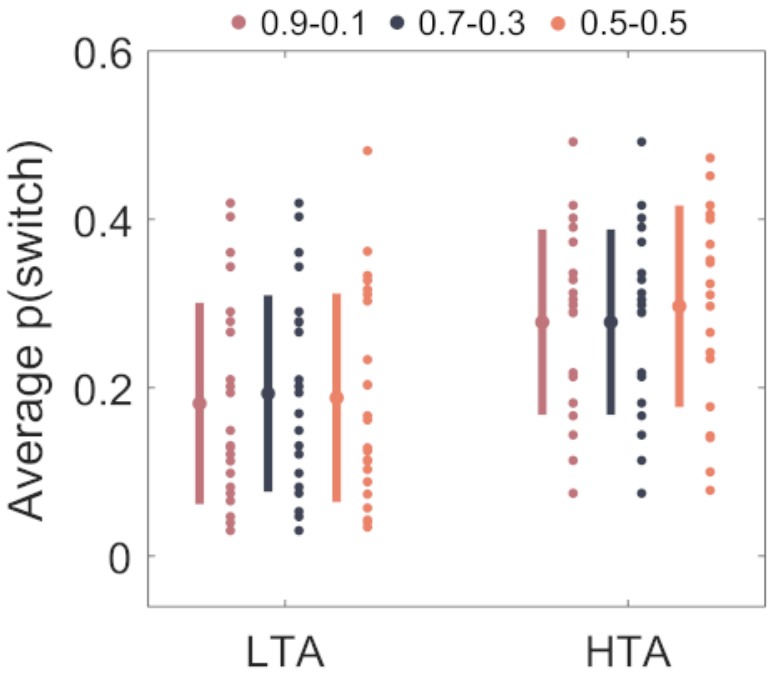
In summary, our findings provide a new understanding of the behavioural, computational, and neuromagnetic spectral correlates of high levels of trait anxiety on reward learning. Our results draw particular attention to the impact of higher-order environmental statistics on reward-learning behaviour in anxiety. Nevertheless, the connection between unique forms of anxiety and different learning domains remains uncertain. Our result that high trait anxiety impairs reward learning is consistent with our previous studies on transient states of anxiety (Hein et al., 2021; Sporn et al., 2020). However, our finding that high trait anxiety elicits larger model updates about the stimulus outcomes is similar to prior work in alternative learning fields where anxiety drives better learning in perceptual or sensory processing tasks, including threat detection and safety seeking (Aylward et al., 2019; Cornwell et al., 2017; Grillon, 2008; Robinson et al., 2011; Wise et al., 2019). Anxiety is a widespread phenomenon of related negative affect and a many-sided field of research; we hope our work can eventually contribute

to the development of learning-based therapy techniques to help reduce excessively high levels of anxiety.

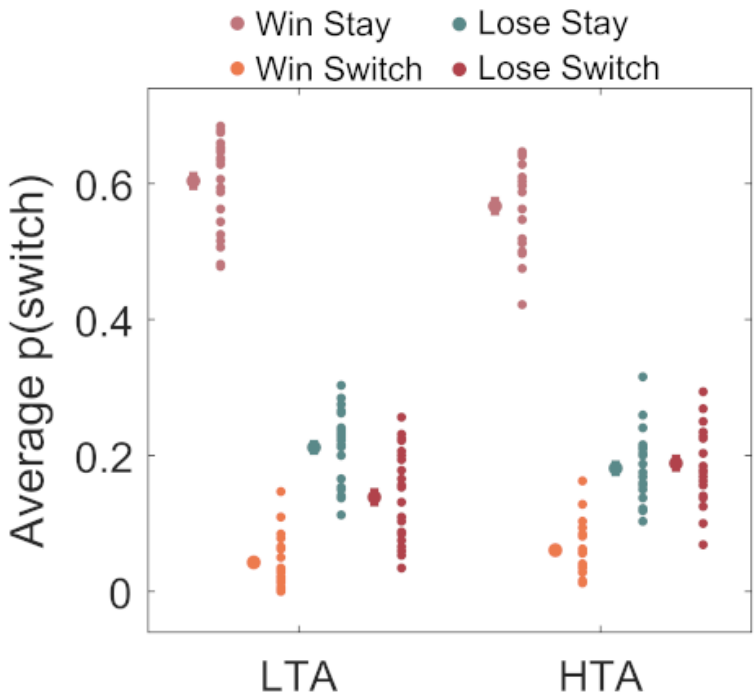
Supplementary Figures



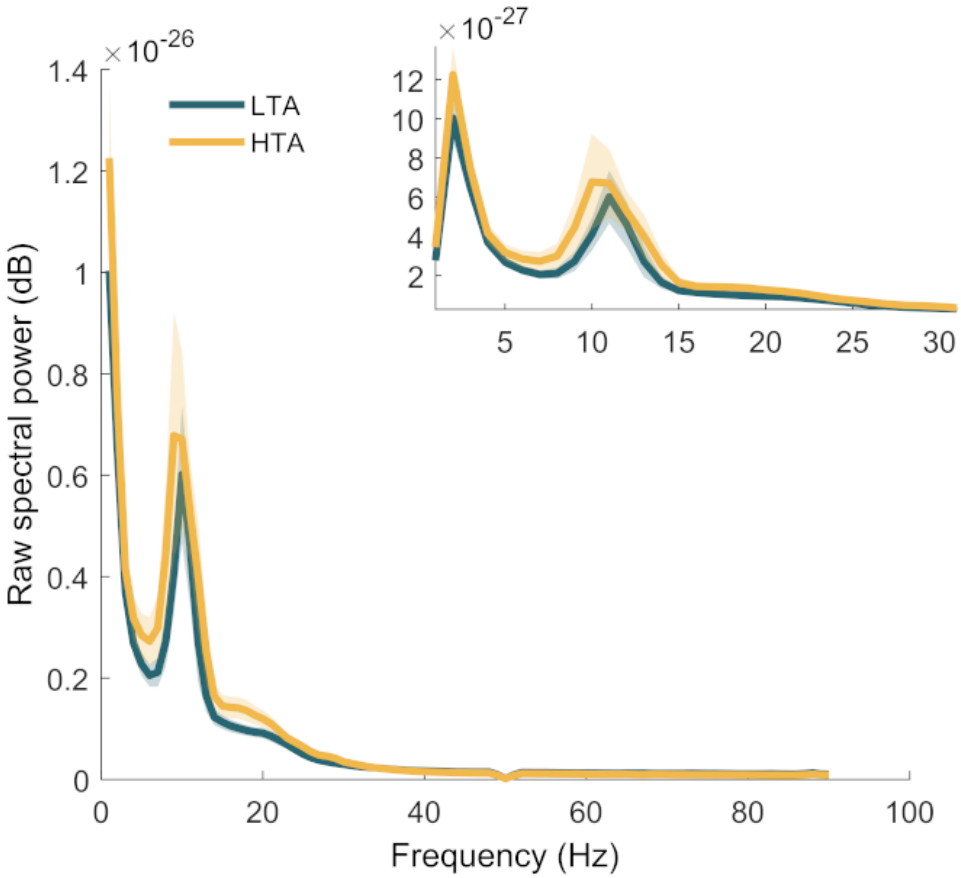
Supplementary Figure 1. Heart-rate variability in trait anxiety. **A)** Normalised heart-rate variability (HRV). Average HRV (coefficient of variation of the inter-beat-interval of the ECG signal) in high trait anxiety (HTA, yellow, mean 0.009, SEM, 0.006) and low trait anxiety (LTA, dark blue, mean −0.011, SEM 0.009). The HRV from both experimental task blocks has been normalised by subtracting the average HRV in the resting state baseline (R1). No significant differences were found using a non-parametric 2 x 2 permutation test with synchronised rearrangements. **B)** Normalised high-frequency HRV. Analysis of the high frequency (0.15 – 0.40 Hz) spectral content of the IBI time series data revealed there was no significant difference between HTA (displayed in yellow, mean −2.7, SEM, 0.4) relative to LTA (dark blue, mean −1.6, SEM 0.8).



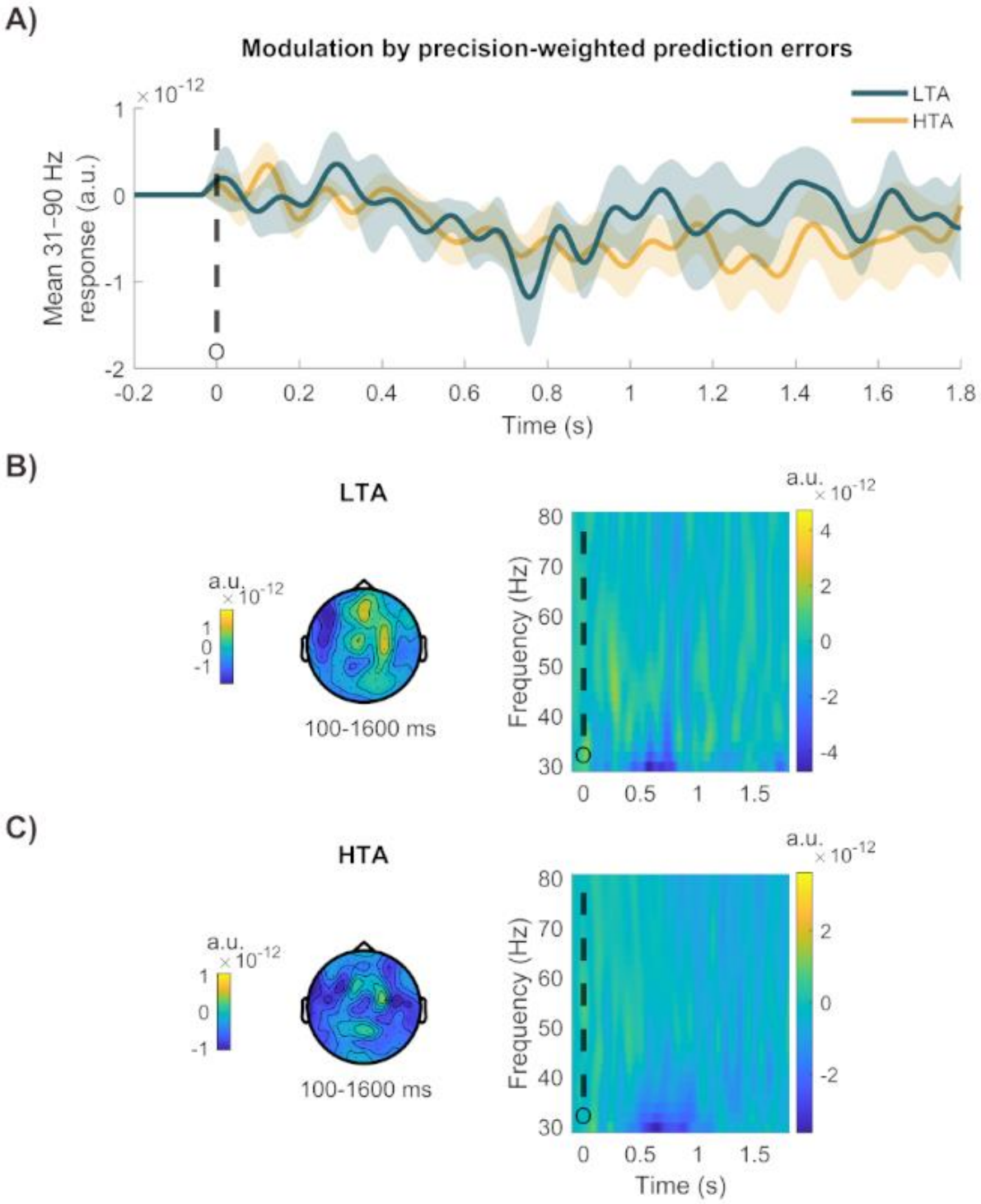
Supplementary Figure 2. Average empirical switch rate by contingency block. The switch rate was calculated for each contingency mapping to show if switching behaviour was higher in an individual contingency phase. Probabilistic contingencies ranged from strongly biased (0.9/0.1; for example the probability of reward for blue p = 0.9), to moderately biased (0.7/0.3), to unbiased (0.5/0.5). The probability of reward assigned to each stimulus was reciprocal: p, 1−p. We observed here that both groups exhibit equivalent switch rates among each contingency phase.



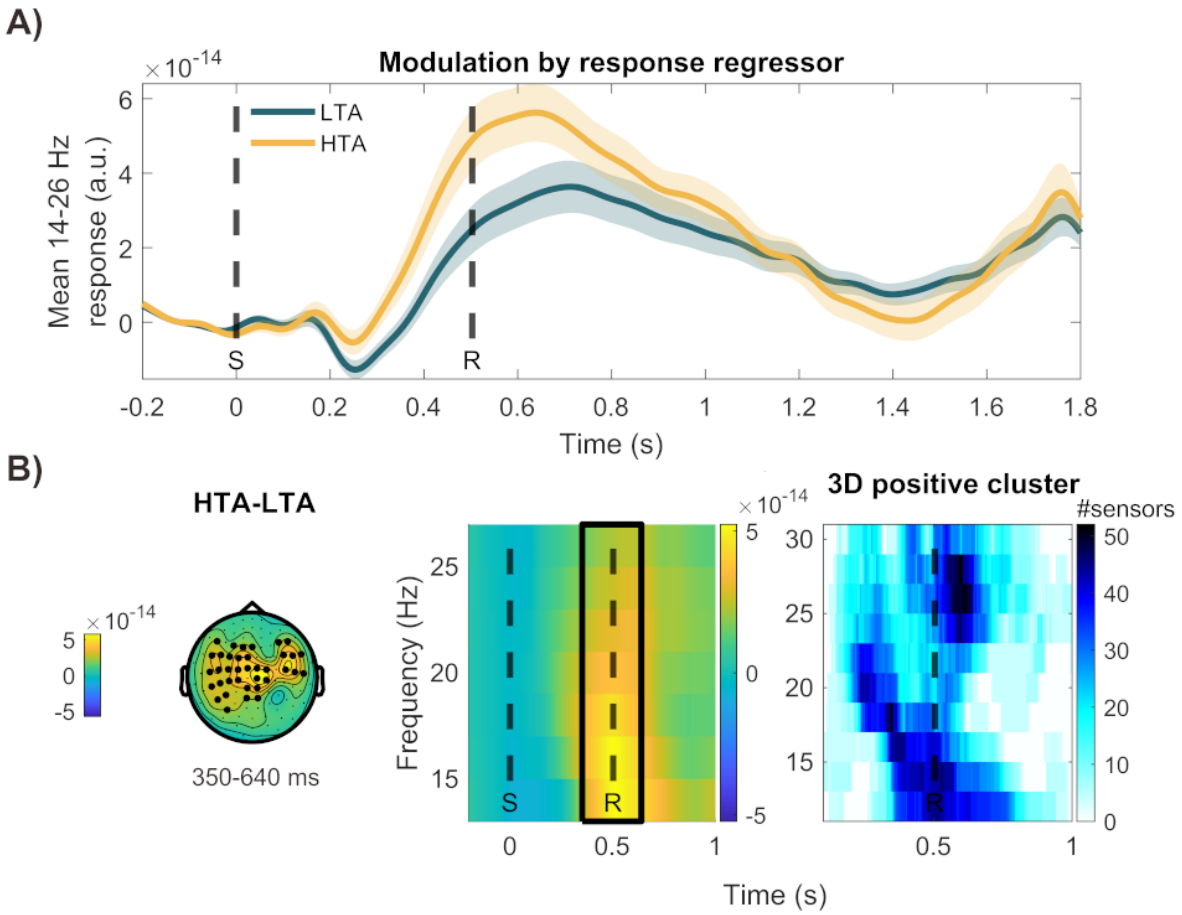
Supplementary Figure 3. Average stay/switch rates by trial outcome. We calculated the stay/switch rate following a trial's outcome (win, lose). The three-way interaction effect between Group (LTA, HTA), Response (Stay, Switch), and Outcome (Win, Lose) was significant (P = 0.02). Post-hoc tests demonstrated that HTA are significantly more likely to switch after a win trial, also showing lower stay rates after a win ($P_{FDR} < 0.05$). Moreover, the HTA group switched significantly more than LTA following a lose trial and stayed less after a lose trial ($P_{FDR} < 0.05$).



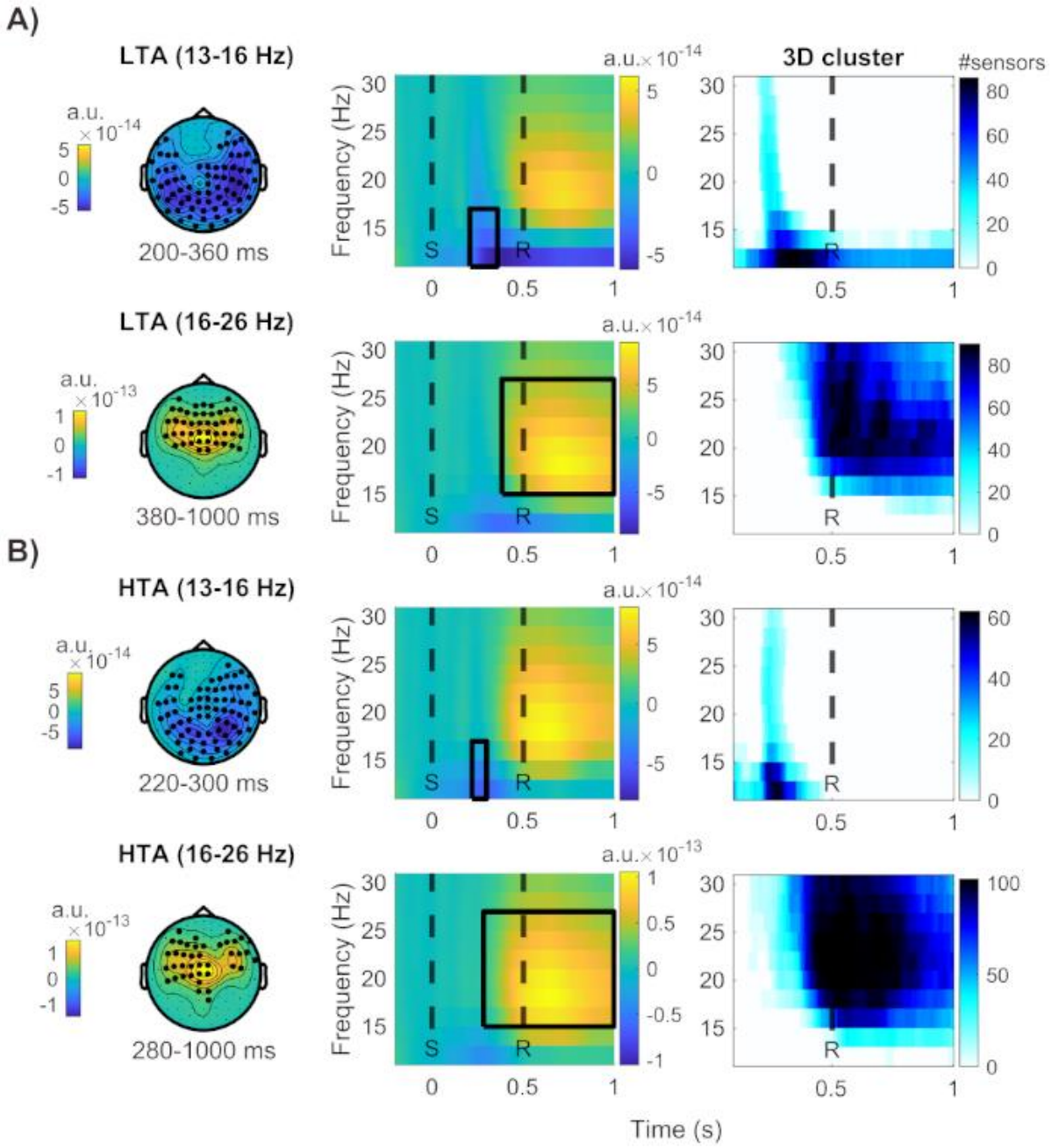
Supplementary Figure 4. Resting-state grand-average of the raw spectral power. The raw power spectral density was converted into decibels (dB: 10*log10), and averaged in HTA (yellow) and LTA (dark blue) groups. Shaded areas denote the standard error of the mean (SEM). There was no significant difference in any frequency band tests (1–3, 4–7, 8–12, 13–30, 13–90 Hz). However, in the inset (top right panel) we observe marginally higher alpha/beta in HTA relative to LTA, in line with Knyazev et al. (2004).



Supplementary Figure 5. Gamma-band oscillatory activity modulated by precision-weighted prediction errors updating belief estimates about the stimulus outcome tendency. **A)** The average gamma-band response (31–90 Hz) in arbitrary units (a.u.) modulated by pwPEs on level 2 ($|\varepsilon_2|$). LTA is given in dark blue, HTA in yellow, with time in seconds (s) on the x-axis. **B)** LTA outcome-locked gamma oscillatory activity modulated by $|\varepsilon_2|$. The left topography shows activity between 100–1600 ms. The right time frequency image is averaged over all sensors (black dashed line, 'O' representing the time of the outcome). **C)** HTA outcome-locked topographic representation of $|\varepsilon_2|$ in gamma activity (left). On the right is the time frequency image averaged over all sensors (outcome given by black dashed line, 'O').



Supplementary Figure 6. Between-group difference on the stimulus-locked oscillatory activity in beta 12–30 Hz modulated by the response regressor. **A)** The average 12–30 Hz response in arbitrary units (a.u.) to responses in each group (LTA, dark blue; HTA, yellow), with time in seconds (s) on the x-axis. **B)** An independent-samples cluster-based permutation statistical analysis on stimulus-locked beta activity by the response regressor—in arbitrary units (a.u.)—showed a significant increase in 14–26 Hz in HTA when compared with LTA. The time of the effect was between 350–640 ms across sensorimotor and central sensors (one positive cluster, P = 0.01, two-sided test, FWER-controlled). The right columns show the time-frequency image response in 14–26 Hz activity for the between-group data averaged over the sensors associated with the cluster and the total number of sensors comprising the 3D cluster. Dashed black lines represent the average time of the stimuli presentation 'S' and the response 'R'.



Supplementary Figure 7. Stimulus-locked oscillatory activity in beta 12–30 Hz is modulated by the response regressor: within-group effects relative to baseline. Responses were given using both hands. **A)** Top row: Dependent samples statistical analysis on stimulus-locked beta activity by the response regressor in the LTA group revealed a significant pre-movement decrease in low beta (13–16 Hz) between 200 to 360 ms post-stimulus (P = 0.02). On the left is the topographical distribution of the significant negative cluster primarily over sensorimotor and central and parietal sensors. On the right are the time-frequency images showing the time and frequency of the significant cluster (in black bars) and the number of sensors contributing to the cluster. Bottom row: Following the initial decrease in 13–16 Hz, dependent samples tests showed a significant increase in post-movement activity in 16–26 Hz (P = 0.002) in the LTA group. The topographical effect is seen over sensorimotor sensors between 380 to 1000 ms post-stimulus. On the right are the respective time-frequency and 3D cluster images. **B)** Similar differences from baseline were observed in the LTA group. Top row: Dependent sample test revealed a trend level pre-movement decrease in low beta (13–16 Hz) between 220 to 300 ms post-stimulus (P = 0.06). On the left is the topographical distribution with the peak of the effect over central

and parietal sensors. On the right are the time-frequency images indicating the time and frequency of the cluster and total sensors comprising the cluster. Bottom row: A significant increase in post-movement activity in 16–26 Hz ($P = 0.001$) was seen in the HTA group. The topography of this difference is across sensorimotor sensors between 280 to 1000 ms post-stimulus. On the right are the time-frequency images showing the time and frequency of the significant cluster and the 3D cluster by frequency across time consisting of a number of sensors given on the z-axis. All tests are two-sided tests FWER-controlled. Plots are given in arbitrary units (a.u.). Dashed black line show events ('S', stimulus; 'R', response). Solid black boxes outline the time and frequency of the cluster.

Supplementary Materials

Model	Prior	Mean	Variance
3-level HGF	κ	1	0
	ω_2	-4	16
	ω_3	-7	16
	$\mu_2^{(0)}$	0	0
	$\sigma_2^{(0)}$	0.1	0
	$\mu_3^{(0)}$	1	0
	$\sigma_3^{(0)}$	1	0
	ζ	48	1
2-level HGF	κ	0	0
	ω_2	-4	16
	ω_3	-7	0
	$\mu_2^{(0)}$	0	0
	$\sigma_2^{(0)}$	0.1	0
	$\mu_3^{(0)}$	1	0
	$\sigma_3^{(0)}$	1	0
	ζ	48	1
HGFμ_3	κ	1	0
	ω_2	-4	16
	ω_3	-7	16
	$\mu_2^{(0)}$	0	0
	$\sigma_2^{(0)}$	0.1	0
	$\mu_3^{(0)}$	1	1
	$\sigma_3^{(0)}$	1	1

Table 1. Table of priors for HGF model parameters. An HGF for binary inputs with no perceptual uncertainty uses the prior values of beliefs $\mu^{(0)}$ and variances $\sigma^{(0)}$ at each level of the 3-Level HGF (HGF_3), 2-level HGF (HGF_2), and 3-level HGF with μ_3 informing the mapping between decision and response ($HGF_{\mu 3}$, Diaconescu et al., 2014). The table also gives the perceptual parameters κ (coupling parameter) and ω_2, ω_3 (tonic volatility estimates on levels 2 and 3) and for the HGF_3 and HGF_2, the decision noise parameter ζ, estimated in log space. Free parameters estimated in unbounded space are log-transformed.

Chapter 6: General Discussion

The work in this thesis examined whether subclinical anxiety in healthy individuals disrupts optimal hierarchical Bayesian learning about rewards in a volatile environment. We used a combination of EEG, MEG, ECG, and computational modelling to ask how anxiety influences reward-based learning responses. In our empirical studies (**Chapters 2–5**), we extensively detailed how anxiety shapes ongoing decision making and the electrophysiological and neuromagnetic encoding of learning signals in the brain.

In **Chapter 2**, healthy volunteers experiencing a state of anxiety showed impairments to reward learning and biases to estimates of uncertainty about the stimulus outcomes and task volatility. Afterwards, we re-analysed the results from **Chapter 2** to demonstrate how state anxiety changes the neural oscillations encoding predictions and precision-weighted prediction errors (**Chapter 3**). In **Chapter 4** we changed tack to focus on the potential adaptive motivational component of anxiety. There we reported null results, discussed the differences in the effectiveness of the threat of social stress as an experimental manipulation, and provided evidence that participants used an alternative response model informed by trial estimates of task volatility. In the final study (**Chapter 5**), we aimed to expand upon the effects of state anxiety reported in **Chapters 2–3** by testing the effects of trait anxiety levels on reward learning and neuromagnetic responses using MEG recordings. We provided evidence that high levels of trait anxiety biased belief estimates about environmental volatility and increased uncertainty about the environment and the reward tendency of the stimulus outcomes. This led to poorer reward learning performance and altered spectral encoding of pwPEs relative to LTA.

The discussion sections of each chapter have already addressed a significant portion of the particular limitations and implications relevant to each study. As such, first we briefly summarise the results of each chapter, and after we concentrate on pulling the related discoveries and their conceptual relation to extant work on anxiety together and discuss dissimilarities and limitations. We specifically also focus on the potential connection of our work to the related field of computational psychiatry, providing selected questions for subsequent researchers aiming to use computational models for capturing the effects of everyday affective states on decision making.

Uncertainty estimation in anxiety

In **Chapter 2**, we discovered that inducing anxiety in the lab using the threat of a psychosocial stressor altered belief uncertainty. In our volatile reward learning context, our state anxiety group's more precise belief estimates about the stimulus outcomes led to the disregarding of information (a lower learning rate)—vitiating learning. However, the influence of the lower-level uncertainty estimates on weighting PEs about stimulus outcomes was expressed in the trial-wise amplitude changes of EEG activity in control participants only. This lack of ERP response in state anxiety was partially consistent with the result that HGF model estimates of the tonic learning rate were significantly reduced under state anxiety. For the controls, modulation of ERP responses by the lower-level pwPEs was consistent with the temporal and spatial profile of the P300 ERP component. However, the lack of a between-group difference hindered our interpretation of neural processing in state anxiety.

In **Chapter 3**, we followed up on this result of group differences in model-based reward learning between state anxiety and controls by reanalysing the EEG data to test how both pwPEs about stimulus outcomes and predictions about the reward tendency are represented in the oscillatory responses of the continuous EEG signal. We tested the hypothesis from hierarchical predictive coding (PC, Bastos et al., 2012) that distinct neural oscillations encode predictions and PEs. We predicted that the biased computations of uncertainty in state anxiety from **Chapter 2** would alter the spectral correlates of these computational learning signals (predictions, pwPEs). Interestingly, using the GLM convolution modelling approach on estimated time-frequency images, we found increased beta oscillations encoding pwPEs in state anxiety in frontal and sensorimotor electrodes, consistent with our recent results in state anxiety during reward-based motor learning (Sporn et al., 2020). This result potentially represented the attenuation of pwPEs. Also, we showed that state anxiety increased beta oscillations during the processing of predictions about the reward tendency, potentially representing a firmer reliance on prior beliefs (Bauer et al., 2014; Sedley et al., 2016). Our results supported generalised Bayesian PC, where one would expect more resistant posterior beliefs (predictions) to be encoded by lower alpha/beta oscillatory activity. Proponents of PC also assert that gamma frequencies encode PEs, and others have shown that gamma frequencies are anticorrelated with alpha/beta frequencies. Our findings suggest that overly precise beliefs about the reward tendency in state anxiety drove lower (tonic) learning rates in our task, potentially inhibiting the encoding of PEs through increased beta oscillations. However, we did not observe the predicted reduction to gamma-band activity encoding pwPEs.

In **Chapter 4,** we attempted to address the potential motivational component of anxiety by coupling learning about reward to reducing anxiety. We aimed to discover if the motivation to reduce the threatened social stress test manipulation generates belief estimates consistent with previously reported studies where anxiety produces adaptive benefits to task performance—characterised by expected unpredictability but with low precision. We hypothesised that given the opportunity to reduce the total time spent performing the social stress test, the state anxiety group might have engaged more in behaviour designed to resolve that uncertainty. This hypothesised result could have explained the differences between adaptive and maladaptive anxious responses reported among different learning domains in anxiety. However, these interpretations and discussion points remain speculative as we failed to induce anxiety in **Chapter 4** successfully, with equivalent self-reported anxiety levels and physiological proxy measures for anxiety, heart rate variability among the groups. As discussed in **Chapter 4**, this may have been due to the experimenter not successfully expressing the cover story required to induce anxiety.

In **Chapter 5**, we examined how trait anxiety shapes reward learning, computations of uncertainty, and pwPE and prediction spectral correlates using MEG data. We showed that the high trait anxiety group (HTA) overestimated volatility relative to low trait anxious controls (LTA). The model that best explained response data was a hierarchical Bayesian filter model (HGF) informed by trial-wise volatility estimates—where overestimating volatility leads to more exploratory choices. We confirmed this in a separate model-free analysis of choice switching responses. Further model-based results indicated that both environmental uncertainty and uncertainty about the stimulus outcome tendencies were higher in HTA. This led to impaired reward learning in HTA, which we observed primarily in the first experimental task block. Analysis of the time-frequency responses of the continuous MEG data revealed that pwPEs were encoded by 10–20 Hz oscillatory activity and that HTA decreased 8–16 Hz oscillatory activity relative to LTA in an equivalent time window as reported in state anxiety in **Chapter 3** over central and sensorimotor sensors.

Our work in these empirical chapters broadly indicates that affective states influence the tracking of reward in a changing environment. As such, the work presented in this thesis more broadly adds to the literature on uncertainty processing in affective states, providing further evidence that the picture concerning maladaptive and adaptive biases in anxiety is still unresolved. There seems to be somewhat of a ravine between different domains of learning in anxiety. Within sensory processing, we observe some potential adaptive functions. While in most executive tasks, we observe impairments and maladaptive biases. In this thesis exclusively testing reward learning, healthy participants under a state of anxiety and high

levels of trait anxiety both expressed difficulties in learning. However, there were critical differences between state and high trait anxiety groups exposed by using a computational modelling approach, which we discuss in more detail below. Much further investigation is required to make connections between these independent areas of learning research in anxiety. However, as a function of the transdiagnostic impact uncertainty has on mental health and the importance of learning for treatment gains, we believe that the work in this thesis has a broad potential impact for informing future work and potential learning-based therapy techniques in anxiety disorders.

Observed differences between state and trait anxiety

We observed similarities and dissimilarities between our experiments manipulating state anxiety and our experiment focusing on trait levels of anxiety. The similarities are simpler to tie together. Both state and trait anxiety disrupt reward learning performance. Several researchers have provided evidence for anxious physiological arousal diminishing cognitive resources and control (Bishop, 2007, 2009; Grupe & Nitschke, 2013; Robbins & Arnsten, 2009), while others have reported an overreliance on habitual or prepotent actions (Dias-Ferreira et al., 2009; Schwabe & Wolf, 2009).

Our modelling approach yielded further insights into how this disruption to learning was different between state and trait anxiety. Bayesian hierarchical filtering models like the HGF can be fit to and describe trial-based empirical learning. This also makes the HGF a simple tool for making precise predictions about group-level learning characteristics as the model provides individual estimates of belief updates.

The HTA group from **Chapter 5** estimated higher levels of volatility and exhibited increased environmental uncertainty and uncertainty about stimulus outcomes relative to LTA, leading to larger updates and noisier (more switching) response choices. By contrast, the state anxious group from **Chapter 2** exhibited a biased decrease in uncertainty about stimulus outcomes and environmental uncertainty while estimating higher levels of uncertainty about volatility estimates (but no overall differences in volatility estimates compared with controls). This leaves the question of why we observed an almost opposite pattern of model results in state compared with high trait anxiety? Moreover, why were the behavioural responses in trait anxiety best explained by a model informed by dynamic volatility estimates when in state anxiety volatility was useful in explaining response data but did not inform online decisions?

Trait anxiety is a self-reported measure of the frequency of anxious experiences, tapping into an individual's estimates of how these symptoms describe them generally (Grupe & Nitschke, 2013; Raymond et al., 2017). The impact of high levels of subclinical trait anxiety may consequently be more representative of general negative affect than state anxiety, bearing a closer resemblance to the characteristics of generalised anxiety disorder (GAD). Increased worry tendency (Hirsch & Mathews, 2012) and negative emotions (Spinhoven et al., 2017) are important factors elevated in high trait anxiety that predispose individuals to the development of GAD (Kanuri et al., 2015). Despite high trait anxiety being a risk factor for developing GAD, mechanistic links remain understudied (Chambers et al., 2004; Kertz et al., 2014; Olatunji et al., 2010). However, more recent neuroimaging work suggests increased resting-state functional prefrontal-amygdala connectivity differentiates GAD from high trait anxiety (Porta-Casteràs et al., 2020). Using our hierarchical Bayesian learning model in **Chapter 5**, we provide additional insights on the computational processes affected by high trait anxiety, which we can also compare to state anxiety.

HTA amplified volatility estimates and drove task choices defined by increased switching responses. Increased uncertainty about the environment and stimulus outcome tendencies ramped up learning about the reward tendency on the second level. Importantly, HTA overestimated volatility in response to our task where probabilistic contingencies changed regularly (there were not distinct stable and volatile blocks, see Browning et al., [2015])—potentially at the cost of learning about the reward contingencies and the probability of an aberrant event (outcome noise). This result is consistent with previous research, which has shown specific difficulties inferring the underlying statistical regularities of the environment and fitting the rate of learning accordingly (Browning et al., 2015; Huang et al., 2017).

Speculatively, in overestimating volatility, highly trait anxious individuals may be underestimating the unpredictability of reward outcomes (the trial by trial noisiness of outcomes rather than systematic change as indicated by volatility, represented by the width of the likelihood distribution). In our work, we did not model both unpredictability and volatility. But the two compete to explain away observed noise. Hence, if high trait anxious participants overestimate volatility, unpredictability will be underestimated (Piray & Daw, 2020b; Pulcu & Browning, 2019). Extant research has typically manipulated just one type of uncertainty or has made visible the expected changes in the experimental uncertainty, or explicitly exaggerated the differences between, for example, expected and unexpected uncertainty (Behrens et al., 2007; Browning et al., 2015; Diederen et al., 2016; Diederen & Schultz, 2015; Nassar et al., 2012, 2016; O'Reilly et al., 2013). Consequently, we know very little about how humans can distinguish and locate the origin of the unique types of uncertainty when not isolated by

experimental manipulation. Studies looking to extend this line of work in anxiety, whether using reward or aversive outcomes, may do well to examine how anxious participants identify, learn from, and use unique sources of uncertainty—and explicitly model them (Piray & Daw, 2019).

One key difference to behavioural responses between our results in **Chapter 5** on trait anxiety and **Chapter 2** on state anxiety is switching responses. Behavioural inhibition is normative in anxiety where threat/reward associations are present (Bach, 2015). As a result, experimentally induced anxiety states may tap into the behavioural expression of hard-wired inhibition mechanisms adapted for predator/prey scenarios (Grillon et al., 2019; Grupe & Nitschke, 2013). By contrast, our HTA group and the experimental procedure in **Chapter 5** was absent of threat—which could have instead revealed the tendency for increased prior expectations (predictions) for the unexpected uncertainty of adverse events (Bach, 2015; Browning et al., 2010; Indovina et al., 2011; LeDoux & Pine, 2016; Pulcu & Browning, 2017).

In **Chapter 2**, we experimentally induced a state of anxiety by threatening a psychosocial stress task. As just mentioned, this may activate the evolutionarily baked in behavioural inhibition induced by threat in anxiety (Bach, 2015). We showed this to some extent in our earlier work using an identical state anxiety manipulation to **Chapter 2** in Sporn et al. (2020). In that reward-based motor learning study, state anxiety was associated with reduced variability in performance in an initial exploration phase, subsequently leading to poorer reward learning (Sporn et al., 2020). In **Chapter 2**, we indirectly showed reduced learning through inflated estimates of the reward outcome tendency (more resistant beliefs and smaller updates about the stimulus outcomes) coupled with less environmental uncertainty. We speculate that this is related to the autonomic flexibility-neuro visceral integration model of anxiety and further to the disconfirmation of maladaptive beliefs.

The autonomic flexibility-neuro visceral integration model asserts that anxiety consists of an intrinsic cognitive rigidity (Friedman & Thayer, 1998). This explains the weaker inhibition of worrying thoughts, rumination, and biased processing of negative outcomes as inflexible cognitive patterns. Especially when under threat, anxiety is linked with more stable physiological responses from the heart (Friedman & Thayer, 1998), rigid responses to changing task demands (Ansari et al., 2008; Derakshan et al., 2009; Edwards et al., 2015), and is also a key factor in OCD (Meiran et al., 2011) and social anxiety disorders (Arlt et al., 2016). In **Chapter 2** we showed reduced physiological variation in responses from the heart (heart rate variability) and more rigid beliefs about probabilistic rewards.

Slower learning about the stimulus outcome tendencies from our HGF modelling results in **Chapter 2** means that those in the state anxiety group were slower to update their beliefs in the face of new information. A potential explanation for our different findings in state anxiety and trait anxiety is avoidance behaviours. In **Chapter 2**, we induced state anxiety using the threat of a social stress test, which may relate more closely to social anxiety, where, for example, avoiding social events can prevent the anxious individual from disconfirming the often erroneous belief that they will be socially embarrassed (Moscovitch, 2009). As such, avoidance of threat can distort inference impairing the disconfirmation of maladaptive beliefs (Moutoussis et al., 2018). Avoidance behaviours can also be viewed as deficient safety learning (Grupe & Nitschke, 2013). While we did not show avoidance behaviours in state anxiety in **Chapter 2**, we did report biased decreases to uncertainty about the stimulus outcomes despite poorer performance. We speculate that this more refractory estimation of uncertainty may represent an attempt to preserve, even if artificially, a safe environment (a bias widely reported in anxiety: see Grupe & Nitschke, 2013). To illustrate, when we precisely estimate that we are on average correct, there would be little need to update our beliefs, and consequently, a subjectively threatening world may become more certain and manageable. Findings reported in high and low trait socially anxious participants in Piray et al. (2019) are also indirectly consistent with this more rigid response style, as disruption of optimal learning was shown through the inflexible application of their learning rate to a dynamic learning environment. Also, a recent study in socially anxious participants showed slower updates to beliefs about reward after experiencing social punishment (Beltzer et al., 2019). Interestingly, inflexibility of beliefs in anxiety may also be true of the influential Browning et al. (2015) study on high trait anxiety, as, in that paper, they used threatening shock outcomes and reported inflexible learning rate adaptation to task changes.

Bringing this back to **Chapter 2**, the above implies that the presence of a future social threat for our state anxiety group may have produced a more rigid inhibited learning style. One further intriguing observation from our state anxiety group is that model estimates of uncertainty were similar to those shown in depression (precise priors that are resistant to change and difficulties engaging and learning from reward: see Parr, Rees, et al., 2018). Depression and social anxiety disorder are highly comorbid (correlation calculated 12 months after diagnosis is 0.43; see Kessler et al., 2005), and thus similar processes like attenuated reward sensitivity may subserve both conditions. As a result, the threat of social stress may result in a self-maintaining negative emotional state that gives rise to beliefs resistant to revision.

A prediction that falls out from the understanding that different forms of anxiety modulate perception and learning in unique ways is that reward-based learning may be differently

affected by the two standard methods for inducing anxiety in the lab: using aversive shocks (Robinson et al., 2013) and psychosocial stress tests (Cavanagh et al., 2011). If this is an accurate interpretation, follow-up studies could examine this claim by testing the differential effect of the threat of shock and threat of social stress in high and low trait anxious individuals on learning in volatile task environments. Additionally, future work using the HGF could test the effects of state anxiety in an environment where volatility is not fixed, to examine how learning is influenced by dynamic changes to phasic volatility.

In summary, a limited number of studies have investigated the potential mechanism of uncertainty estimation in anxiety (Pulcu & Browning, 2019). The results so far indicate that environmental uncertainty plays a central role in learning about threatening or aversive stimuli in anxiety (Grillon et al., 2019), and now we have expanded this to a purely reward-based learning environment. However, to the author's knowledge, few studies have examined the effect of unexpected and expected uncertainty on learning simultaneously while modelling these quantities independently. Furthermore, more experiments need to be conducted testing the relationship between anxiety, misestimations of uncertainty, and prior experiences to reveal the processes that may contribute toward fitting a profile of emotional/mood disorders or clinical anxiety. Evaluating if the computational models that usefully describe anxiety in the laboratory will prove to be a clinically relevant tool in recognising and treating anxiety disorders represents a vital and exhilarating challenge.

Neural responses in time and frequency space in anxiety

The work in this thesis expanded upon prior computational work on learning in anxiety by providing electrophysiological and neuromagnetic results (Browning et al., 2015; Huang et al., 2017; Pulcu & Browning, 2017). In previous studies, pupillometry had been used as an index of the trial-wise encoding of surprise (and estimates of volatility) in phasic noradrenergic responses, dependent on activity from the central noradrenaline system (Browning et al., 2015; Pulcu & Browning, 2017; Yu & Dayan, 2005). These phasic pupil responses indicate the observation and enhanced processing of more informative events (higher volatility relating to greater pupillary dilation) by increases in the gain of sensory input (equivalent to raising the learning rate, see Aston-Jones & Cohen, 2005; Browning et al., 2015; Nassar et al., 2012; Pulcu & Browning, 2017). Highly trait anxious individuals have been shown to exhibit dampened pupil dilation in response to changes in the environment, a result paired with minor adjustment of the learning rate to task changes (Browning et al., 2015).

Our work focused on computations of uncertainty and prediction error signalling in anxiety during reward-based learning. We provided evidence of single-trial ERP responses correlated with model-based pwPEs in controls in **Chapter 2**, and of the oscillatory correlates of pwPEs and predictions about stimulus outcomes in anxiety and controls using EEG and MEG in **Chapters 3 and 5**. Our results broadly indicate that our brains are using predictive models of the reward statistics generated by our volatile task environment, assimilating top-down predictions about reward encoded in alpha/beta oscillations with bottom-up sensory input about stimulus outcomes. The findings give novel empirical evidence for altered oscillatory activity in anxiety and evidence of PC in higher-order cognitive functioning.

Computational simulation studies have shown an asymmetry in the spectral profile of feed-forward and feedback connections (Bastos et al., 2012; Lee et al., 2013). The direction of our effects in alpha/beta-band oscillatory activity encoding predictions follows the cortical processing hierarchy outlined in PC: that alpha/beta oscillations are strongest in the associated descending (top-down) feedback (Bastos et al., 2012; van Pelt et al., 2016). In a Bayesian PC context, this means that stronger priors (predictions) are modulating lower-level error transmission by providing inhibitory input to superficial pyramidal neurons encoding PEs in gamma. This attenuation on PE signals is also thought to happen by postsynaptic gain regulation (the precision ratio modulating PEs, see Bauer et al., 2014; Brown & Friston, 2013; Larkum et al., 2004).

The parametric amplification of beta oscillations encoding predictions in anxiety reported in this thesis thus implies that the precision of descending predictions is processed by lower frequency oscillations (van Pelt et al., 2016). The novelty of our results are that biased estimates of precision in anxiety inhibit encoding of PEs used for updating beliefs about the tendency of different stimuli to be rewarding, which impacts reward learning behaviour. Our findings align with a recent study showing more attentional alpha modulation with more predictable targets (Bauer et al., 2014), increased alpha/beta oscillations encoding predictions about sensory input (Arnal & Giraud, 2012; Bastos, Litvak et al., 2015; van Pelt et al., 2016; Wang, 2010), and more generally with active inference (Friston et al., 2011).

In **Chapter 2**, we reported that state anxiety increased precision about the stimulus outcome tendency and impacted the learning rate (associated with higher levels of beta-band activity in **Chapter 3**). But how do we reconcile this discovery with the fact that the state anxiety group was still learning from each trial? Higher precision of predictions represents only that the predictions in state anxiety are more precise, not that priors are updated less by pwPEs. We did not find evidence of gamma activity encoding pwPEs in controls or in state anxiety.

Observing increased beta oscillations during the encoding pwPEs in state anxiety still allows for updates about the stimulus outcomes to occur, we just did not observe gamma-band changes in our study, a limitation discussed below. We stress that the resultant smaller (not fewer) updates in state anxiety represent a suboptimal strategy/maladaptive learning, as shown by overall poorer reward learning.

In **Chapter 5**, we revealed that HTA was associated with an opposite spectral profile encoding pwPEs compared with state anxiety in **Chapter 2**, which fits with the opposite pattern of modelling results reported in **Chapter 5** in HTA. Overestimates of volatility in HTA led to higher uncertainty about the environment and reward outcome tendencies, with lower beta-band activity encoding pwPEs relative to LTA in an equivalent time window and spatial topography to those reported in state anxiety in **Chapter 2**. This outcome may represent a weaker reliance on prior beliefs (Sedley et al., 2016) that up-weights PEs for revising predictions, associated in the work of others with increased gamma responses (Bauer et al., 2014). Yet, pwPEs in our study did not significantly modulate gamma activity as a function of high or low trait anxiety. This leaves the question of how HTA participants had larger updates about the stimulus outcomes (but showed poorer learning, at least initially) and did not exhibit increased gamma oscillations?

We did not observe gamma-band activity modulated by pwPEs using both EEG and MEG. Proponents of PC assert that gamma represents the unexplained portion of the propagated sensory input (Bastos et al., 2012; Friston, 2010). As a result, unpredicted and lower probability outcomes elicit larger PE responses encoded by increased higher frequency gamma oscillations. In **Chapter 3**, we reasoned that state anxiety decreased PEs and associated gamma activity through lower estimation uncertainty and inhibitory input to superficial pyramidal neurons. However, this did not explain why reward-related pwPE signals did not modulate gamma activity in our control group. Alternatively, we reasoned that our EEG recordings were not sensitive enough to detect reward-based PEs in gamma activity. We also reported in **Chapter 5** that reward-based PE signals may be harder to detect in observable cortical changes in oscillatory activity after we did not observe changes to gamma oscillations during the processing of pwPEs using MEG.

Our findings suggests that, in our task, encoding pwPEs is not associated with gamma modulation. This outcome was unforeseen as there is now increasing data suggesting cortical gamma activity is modulated by reward in multiple cognitive domains outside of perception (Berke, 2009; Ellwood et al., 2017; Lohani et al., 2019). As one example, a recent experiment using invasive local field potential (LFP) recordings showed dmPFC gamma oscillations

encoding unsigned reward PEs during a exploration-exploitation task (Domenech et al., 2020). Perhaps changes in gamma oscillations encoding reward-based pwPEs are simply not observable using the reduced sensitivity of EEG and MEG relative to LFP.

A potential explanation beyond the limitations of M/EEG is that in our task the pwPE regressor is not directly coding reward PEs, but both i) a positive reward PE (rewarded when weak association of stimulus and reward) and ii) a negative reward PE (unrewarded despite selecting the stimulus strongly associated with reward). Alternatively put, our pwPE estimate codes for how much the participant needs to update the contingency (tendency) belief, not how much more or less (than expected) reward they received.

Further studies into reward-based learning are needed to resolve this limitation in our work. Using invasive LFP recordings in humans may be especially important in future work to expand the oscillatory hypotheses of Bayesian PC to more general learning contexts. It would also be interesting to reintroduce manipulations to task volatility, as with earlier work (Browning et al., 2015), and record M/EEG responses to track the hypothesised larger gamma-band activity elicited by environmental change and unpredictable outcomes. Combining pupillometry and MEG data with hierarchical Bayesian modelling would be an excellent way to collate evidence on computations of uncertainty, learning rates, and altered oscillatory responses by pwPEs and predictions in anxiety.

One limitation with the studies in this thesis is our inability to use the third level log-volatility estimates in GLM analyses due to multicollinearity, the high linear correlation between pwPEs/predictions about stimulus outcomes and pwPEs/predictions about volatility. Volatility in the HGF model is estimated as a continuous value evolving as a Gaussian random walk. We have discussed alternatives that model changes in the environment as sudden changes (Nassar et al., 2010), highlighting how, in practice, both generative models (HGF and change-point models) can successfully handle both sudden and diffuse environments (Marković & Kiebel, 2016). That analysis showed that Bayesian inference and model comparison methods could correctly distinguish data generated by the HGF versus a reformulation of a change-detection model. However, perhaps they still might lead to different interpretations. An exciting avenue of research to explore would be if change point models (Nassar et al., 2010; Wilson et al., 2013) can make distinct predictions concerning neural processing and circuity and whether these models can better explain the associated behavioural data; this is associated with the still unresolved question of whether the brain utilises a generic strategy for solving tasks, possibly at the expense of specialised tasks (Weber, 2020). Interestingly, recent accounts of learning have asserted that simple but general form algorithms may perform the

learning process in one part of the brain, like the dopamine system, and that these parts may be used to train other areas like the prefrontal cortex to use more specialised solutions to particular task demands (Wang et al., 2018).

Key challenges for work to come will be in the examination of model update principles. An interesting outstanding question is how PEs afford updates to the generative model to provide refined posterior beliefs about upcoming outcomes based on alpha/beta and gamma oscillations. Moreover, we still know very little about how the descending oscillatory signals of predictions interact with ascending signals at a neurophysiological level to alter gamma-band activity. Suggestions so far have included cross-frequency phase–power, power–power or phase–frequency coupling (Jensen & Colgin, 2007). Future studies could specifically target these cross-frequency interactions in anxiety in a learning task where the experimental manipulation methodically modulates model updating.

The role of the prefrontal cortex in reward learning and anxiety

Despite increasing data strengthening the striatal reward PE model of learning over the past two decades, continued investigation has revealed somewhat of a conundrum in the involvement of the prefrontal cortex (PFC, Wang et al., 2018). Evidence suggests that the PFC realises reward-based learning processes astonishingly similar to the mechanisms detailed for dopaminergic reinforcement learning. Although the neural expression of expected value in the PFC has been strongly supported for some time (Padoa-Schioppa & Assad, 2006; Rushworth & Behrens, 2008; Seo & Lee, 2008), a more contemporary line of investigation has revealed that the PFC encodes a recent record of actions and rewards (Barraclough et al., 2004; Kim & Shadlen, 1999; Seo & Lee, 2008; Seo et al., 2012; Tsutsui et al., 2016; Wang et al., 2018). Put simply, some activity in the PFC seems to encode an independent and self-sufficient reinforcement learning algorithm (Wang et al., 2018).

The contribution of the PFC to reward-based reinforcement learning and its overlap in anxiety states provokes the question of how anxiety might interfere with reward-based learning in this self-contained PFC network. Some have reported that dopamine drives model-free learning, whereas the PFC drives model-based learning. This is because the PFC is known to represent task structure (Bromberg-Martin, Matsumoto, Hong, et al., 2010; Daw et al., 2005). Recent work attempts to reconcile the two systems, in light of evidence that dopaminergic PEs also represent task structure (Bromberg-Martin, Matsumoto, Hong, et al., 2010; Daw et al., 2011; Nakahara & Hikosaka, 2012; Sadacca et al., 2016), by suggesting that the dopaminergic model-free system may instruct and calibrate the model-based PFC system to seize upon

particular answers to particular changing task environments (Wang et al., 2018). Wang and colleagues (2018, p. 861) call this 'meta-reinforcement learning' from the recurrent neural network of the PFC.

The relevance for our discussion on anxiety is both the diminished activity in the PFC reported in anxiety (denuding capabilities for control of attention, working memory, and predicting safety Bishop, 2007, 2009; Grupe & Nitschke, 2013; Robinson et al., 2013), and the reported deficit in adapting to changing environments (Browning et al., 2015; Huang et al., 2017). The expectation of threat in anxiety may thus use up the resources necessary for processes unrelated to threat, such as working memory. That said, this account is likely oversimplified, as different circuits between several brain regions are implicated in anxiety (Grupe & Nitschke, 2013). Accordingly, anxiety may interfere with the optimal and efficient learning of the statistical task structure and reward environment, linking to the neuroimaging results of Behrens et al. (2007), who showed dorsal anterior cingulate cortex (dACC) activity in the PFC associated with adapting learning rates to environmental change, and more broadly to work reporting the dACC is involved in reward-based decision making (Bush et al., 2002).

It is difficult to interpret our electrophysiological findings (**Chapters 2** and **3**) with respect to the above neuroimaging work. However, disruptions to PFC activity could be related to our lack of trial based ERP responses under state anxiety in **Chapter 2**, where controls showed an ERP response in frontocentral channels. In **Chapter 3,** we provide one potential neurophysiological process that may drive these reward-based learning disruptions in anxious states in the increased beta oscillations encoding predictions and pwPEs in frontocentral electrodes. Later in **Chapter 5,** we provide neuromagnetic evidence of decreased beta oscillations encoding pwPEs in HTA compared with LTA. However, this effect was seen across central, sensorimotor, and right frontal sensors. Our future source space study on this MEG dataset should clarify the origins of these oscillatory differences in trait anxiety. And further studies in this field would also benefit from utilising dynamic causal modelling to reveal in greater detail the role of the interconnected regions in anxious reward learning (such as prefrontal-amygdala connectivity, see Soltani and Izquierdo [2019] for further details) and the encoding of pwPEs and predictions.

Outstanding questions for continued work into how the brain computes uncertainty is how the different forms of uncertainty are represented in the brain and whether these are distinguishable for different outcomes, such as reward and punishment. Even though studies in animals and humans show an array of separable neural processing regions associated with uncertainty estimates, little is known about how affective outcomes (e.g. reward, punishment)

and states (e.g. mood, state anxiety, trait anxiety) modulate uncertainty processing in the brain.

Limitations

In **Chapters 2–5**, we showed that the hierarchical Bayesian filter (HGF, Mathys et al., 2011) model explains reward learning behaviour in an unpredictable (in)stable environment better than alternative reinforcement learning models and a HGF with a fixed volatility term. But how confident should we be in the conclusion that the brain is making use of the HGF model quantities? Alternatively put, it might be difficult to create a model that explains behavioural data that exhibits no relationship to neural responses. Predictive coding (PC) describes a algorithmic process of perception and inference (Friston, 2005; Rao & Ballard, 1999) based on the hierarchical organisation of the cortex, on which similar hierarchical principles the Bayesian HGF is constructed. Uncertainty is in the foreground of both PC and HGF models, with the strength of belief updates depending on the precision (inverse uncertainty) of prior beliefs (predictions) scaling prediction errors (pwPEs). However, in contrast to the HGF, PC is set in continuous time, with the posterior beliefs and pwPEs that develop from sensory input seized upon by differential equations, with higher levels generally predicting the state of lower levels (i.e. the mean [versus the variance] of the probability distribution, see Friston, 2005, 2010; Rao & Ballard, 1999). [15] By contrast, the HGF functions in discrete time using one-step update equations, where higher levels predict the volatility of lower levels (Mathys et al., 2011, 2014). As such, higher levels of estimated volatility modulate the uncertainty at lower levels giving rise to faster belief updates. Also, the HGF calculates belief updates across all levels of the hierarchy and uses a response model that provides it the ability to explain trial based behavioural and neural responses. Some outstanding questions that fall out from these differences between PC and the HGF are whether the HGF can be developed to handle both volatility-based and mean value-based updates, using differential equations for the within-trial dynamics of belief updates, and how this might differently operate in the cortical microcircuits outlined by Bastos et al. (2012). Future studies could then start to show the equivalency between these two hierarchical Bayesian models of perception and learning, and demonstrate through simulations the predicted neural responses.

A limitation of our work is that we used only the HGF as a Bayesian model of learning about environmental stability. There are, however, alternatives such as change-point models and an

[15] Although for an exception to predicting the mean of probability distributions in PC see Kanai et al. (2015).

extension of the Kalman filter that includes volatility (volatility Kalman filter [VKF]) estimates that could be tested in future work (Moens & Zénon, 2019; Nassar et al., 2010; Piray & Daw, 2020a, 2020b; Wilson et al., 2010). While we discussed previously that in practice both change point detection models and the HGF perform equivalently in a comparative analysis (Marković & Kiebel, 2016) and can be approximated with one another (Mathys, 2012), these models may produce unique interpretations from identical data. Potentially change point models or the VKF might be better at explaining behavioural responses in anxiety. A further intriguing avenue for future work might be whether change point detection models and the VKF form independently distinguishable predictions about underlying cortical microcircuitry and the distinct types of uncertainty, and which can best describe behaviour.

A second limitation of working with the HGF is the inability to model outcome noise. This can purportedly be addressed by the VKF through the simultaneous modelling of both volatility estimates and 'unpredictability' or outcome noise (Piray & Daw, 2020a, 2020b). Distinguishing between these two origins of uncertainty is an important and understudied feature of adaptive learning behaviour (Piray & Daw, 2020a, 2020b), particularly concerning neuropsychiatric populations and affective disorders related to misestimation of uncertainty (Pulcu & Browning, 2019). Future work could explore anxious learning during changes between stable and volatile blocks using the VKF to determine the relative influence of both types of uncertainty on belief updates. Perhaps anxiety might best be described by difficulties distinguishing between outcome noise and meaningful changes to the environmental statistics. We would anticipate based on our findings in this thesis that high trait anxiety would be associated with the increased misattribution of outcome noise as ground truth changes to the environment.

A final limitation of our work is in the suboptimal design choices of **Chapter 4** that aimed to test the motivational component of anxiety in driving uncertainty reduction and reward-seeking behaviour in state anxiety. We estimated that the motivation to perform well during our reward-learning task and reduce the total time spent performing a secondary anxiety-inducing social stress test would be a sufficiently powerful motivational effect to alter learning. However, two design issues limiting **Chapter 4** included the subtle effect of this kind of external motivation, and the issue discussed in **Chapter 4** concerning the performance of the experimenter delivering a cover story to successfully induce anxiety. It would be interesting to examine this motivational component of anxiety in follow-up reward-based studies where we addressed these design limitations. Future work could make use of the shock methodology in inducing anxiety. A useful design could be where participants experience blocks under threat of shock that are coupled to independent blocks of safety where reward based learning is isolated. In

reward-based blocks participants can reduce the unpredictability or probability of later shock-based blocks.

Challenges and future directions

As a final discussion point, we attempt to at least partially connect the results of this research thesis to computational psychiatry and clinicians' efforts to treat aspects of anxiety (Moutoussis et al., 2017, 2018; Nair et al., 2020). One intriguing connection is between our findings of biased uncertainty estimates and resistant belief updates in state anxiety and the 'counterfactual gating' hypothesis from Moutoussis et al. (2018). In that paper, the authors outline an account of maladaptive belief updates where absent feedback (counterfactual outcomes, ones that did not materialise) of predicted aversive events, when aligning with maladaptive beliefs, could prevent the updating of more credible alternatives (Moutoussis et al., 2018). Thus, Moutoussis et al. (2018, p. 60) conclude, "counterfactual thinking can then strengthen maladaptive beliefs and drive an increase of avoidance behaviours in the absence of actual feared outcomes". Counterfactual gating is then the process whereby harmless events are mis-estimated as behaviours resulting in the reaffirmation of avoidance or safety behaviours. This type of maladaptive attribution of harmful outcomes to benign causes can maintain dysfunctional beliefs by inhibiting the impact of new evidence (Moutoussis et al., 2018). One speculation from the results in this thesis is that states of anxiety may be related to counterfactual gating through the biasing of uncertainty estimates. This line of research into dysfunctional biased beliefs, misestimates of uncertainty, and maladaptive evidence accumulation in anxiety is an exhilarating and challenging direction for future research.

One important issue with probing these biases further using current learning and decision-making tasks in cognitive neuroscience and psychology (as was used in this thesis) is the lack of ecologically valid tasks. The bespoke task environments we use for investigating and fitting models to human behaviour are oftentimes excessively artificial, with an immoderate number of assumptions concerning the information agents might use to guide behaviour to achieve some task goal. Ecologically valid tasks are methodologically and ethically more difficult, demanding a whole different level of experimenter and participant trust (Moutoussis et al., 2018). Computational psychiatry (and the related work using models in subclinical samples as reported here) would benefit from concentrating on reproducibility, generalisability, and the test-retest reliability of parameter estimation and tasks (Nair et al., 2020; Wilson & Collins, 2019). The changes observed in the artificial contexts in cognitive neuroscience need to generalise between real life and the experimental setting, of which there is little evidence so

far (Eisenberg et al., 2019). This is particularly important when we seek to inform treatment based psychological therapy (Scholl & Klein-Flügge, 2018).

Outside of the artificial tasks used in human cognitive neuroscience, alternative research areas have benefited from utilising more naturalistic designs (computational ethology, see Mobbs et al., 2021). While some steps have been taken to achieve this in neuroscience, with immersive tasks (Nord et al., 2017) and by using a host of digital graphics and virtual reality (Bouchard et al., 2017; Gega, 2017), future studies would gain tremendously from taking on this approach to achieve an understanding of behaviours not easily accessible using traditional tasks; we could thus considerably advance our insights into behaviours observed outside of the laboratory.

Future work also needs to establish how stable computational parameters are across time in control participants to be helpful in a therapeutic context. As mentioned above, the test-retest reliability for parameters of specific tasks is at present relatively moderate (Enkavi et al., 2019). Increasing parameter reliability can boost the power and sensitivity of use for tracing brain changes and behaviour changes in clinical states and assessing the efficacy of treatments (Nair et al., 2020). As touched on before, specifically tailored or artificial tasks also can reduce how much we can interpret and assimilate new evidence. Bringing this back to anxiety, this is particularly vital when aiming to understand the impact of an intervention on an outcome, for example, the impact of cognitive behavioural therapy in anxiety on the learning rate in volatile environments (Nair et al., 2020). In contrast to neuroimaging/magnetic and electrophysiological methods, computational modelling techniques are uniquely positioned to afford scalable and pragmatic biomarkers (e.g. using online and smartphone apps for collecting large datasets) extracted from behaviour and validated by neural physiology. Using this method also affords better post-study/treatment longitudinal examinations for tracking the expression of symptoms and treatment efficacy. Good examples of this have been in mood and decision making research (Eldar et al., 2018; Rutledge et al., 2014, 2017) and more recently in anxiety (Wise & Dolan, 2020).

In summary, throughout the work in this thesis, we have explored and revealed how anxiety is not a uniform affective state homogeneously interacting with cognition. Rather, we have found that the different types of anxiety generate distinct effects on behaviour and neural responses in our reward learning task. We observe in the work of others that the different forms of anxiety also have different effects on behaviour depending on the learning context. Our focus has been on the understudied aspect of learning from rewards in subclinical anxiety, examining the divergence from Bayes optimal learning in volatile environments. We utilised a

modelling approach to understand further the computational dimensions of individual variability pertinent to these learning and inference states. This approach is particularly important for clarifying whether sub-optimality in Bayesian learning and cognitive biases are normative, as with anxiety, depending on context. Future research can then take the critical next steps in identifying risk factors specific to individuals from biases that may help to create and sustain anxiety disorders; as despite identifying biased reward processing in several neuropsychiatric disorders, we are not much closer to understanding their symptoms and trajectories (Moutoussis et al., 2015; Whelan et al., 2014).

Bibliography

Abend, R., Gold, A. L., Britton, J. C., Michalska, K. J., Shechner, T., Sachs, J. F., Winkler, A. M., Leibenluft, E., Averbeck, B. B., & Pine, D. S. (2020). Anticipatory Threat Responding: Associations With Anxiety, Development, and Brain Structure. *Biological Psychiatry*, *87*(10), 916–925.

Aberg, K. C., Toren, I., & Paz, R. (2021). A neural and behavioral tradeoff underlies exploratory decisions in normative anxiety. *bioRxiv*, 2020-09.

Abramowitz, J. S., Taylor, S., & McKay, D. (2009). Obsessive-compulsive disorder. *The Lancet*, *374*(9688), 491–499.

Adamantidis, A. R., Tsai, H.-C., Boutrel, B., Zhang, F., Stuber, G. D., Budygin, E. A., Touriño, C., Bonci, A., Deisseroth, K., & de Lecea, L. (2011). Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(30), 10829–10835.

Adams, C. D., & Dickinson, A. (1981). Instrumental Responding following Reinforcer Devaluation. *The Quarterly Journal of Experimental Psychology Section B*, *33*(2b), 109–121.

Adams, R. A., Huys, Q. J. M., & Roiser, J. P. (2016). Computational Psychiatry: towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery, and Psychiatry*, *87*(1), 53–63.

Adams, R. A., Perrinet, L. U., & Friston, K. (2012). Smooth pursuit and visual occlusion: active inference and oculomotor control in schizophrenia. *PloS One*, *7*(10), e47502.

Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: active inference in the motor system. *Brain Structure & Function*, *218*(3), 611–643.

Aftanas, L. I., Pavlov, S. V., Reva, N. V., & Varlamov, A. A. (2003). Trait anxiety impact on the EEG theta band power changes during appraisal of threatening and pleasant visual

stimuli. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *50*(3), 205–212.

Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., & Breiter, H. C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron*, *32*(3), 537–551.

Ahmed, S. H. (2010). Validation crisis in animal models of drug addiction: beyond non-disordered drug use toward drug addiction. *Neuroscience and Biobehavioral Reviews*, *35*(2), 172–184.

Aikins, D., & Craske, M. G. (2010). Autonomic expressions of anxiety: heart rate and heart period variability in the anxiety disorders. *Handbook of Cognitive and Affective Neuroscience of Psychopathology. Oxford University Press, USA, New York*.

Ainley, V., Apps, M. A. J., Fotopoulou, A., & Tsakiris, M. (2016). "Bodily precision": a predictive coding account of individual differences in interoceptive accuracy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1708), 20160003.

Aitchison, L., & Lengyel, M. (2017). With or without you: predictive coding and Bayesian inference in the brain. *Current Opinion in Neurobiology*, *46*, 219–227.

Alamia, A., & VanRullen, R. (2019). Alpha oscillations and traveling waves: Signatures of predictive coding? *PLoS Biology*, *17*(10), e3000487.

Albin, R. L., Young, A. B., & Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends in Neurosciences*, *12*(10), 366–375.

Al-Ezzi, A., Kamel, N., Faye, I., & Gunaseli, E. (2020). Review of EEG, ERP, and Brain Connectivity Estimators as Predictive Biomarkers of Social Anxiety Disorder. *Frontiers in Psychology*, *11*, 730.

Allen, A. P., Kennedy, P. J., Cryan, J. F., Dinan, T. G., & Clarke, G. (2014). Biological and psychological markers of stress in humans: focus on the Trier Social Stress Test. *Neuroscience and Biobehavioral Reviews*, *38*, 94–124.

Allen, J. G., Fonagy, P., & Bateman, A. W. (2008). *Mentalizing in Clinical Practice*. American Psychiatric Pub.

Amaral, D. G., & Price, J. L. (1984). Amygdalo-cortical projections in the monkey (Macaca fascicularis). *The Journal of Comparative Neurology*, *230*(4), 465–496.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.

Andersen, S. B., Moore, R. A., Venables, L., & Corr, P. J. (2009). Electrophysiological correlates of anxious rumination. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *71*(2), 156–169.

Anderson, M., & Braak, C. T. (2003). Permutation tests for multi-factorial analysis of variance. *Journal of Statistical Computation and Simulation*, *73*(2), 85–113.

Ang, Y. S., Lockwood, P., Apps, M. A. J., Muhammed, K., & Husain, M. (2017). Distinct Subtypes of Apathy Revealed by the Apathy Motivation Index. *PloS One*, *12*(1), e0169938.

Ansari, T. L., Derakshan, N., & Richards, A. (2008). Effects of anxiety on task switching: evidence from the mixed antisaccade task. *Cognitive, Affective & Behavioral Neuroscience*, *8*(3), 229–238.

Apps, M. A. J., Grima, L. L., Manohar, S., & Husain, M. (2015). The role of cognitive effort in subjective reward devaluation and risky decision-making. *Scientific Reports*, *5*, 16880.

Ariely, D., & Jones, S. (2008). *Predictably irrational*. New York, NY: Harper Audio.

Arlt, J., Yiu, A., Eneva, K., Taylor Dryman, M., Heimberg, R. G., & Chen, E. Y. (2016). Contributions of cognitive inflexibility to eating disorder and social anxiety symptoms. *Eating Behaviors*, *21*, 30–32.

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

Arnaudova, I., Kindt, M., Fanselow, M., & Beckers, T. (2017). Pathways towards the proliferation of avoidance in anxiety and implications for treatment. *Behaviour Research and Therapy*, *96*, 3–13.

Arvanitogiannis, A., & Shizgal, P. (2008). The reinforcement mountain: allocation of behavior as a function of the rate and intensity of rewarding brain stimulation. *Behavioral Neuroscience*, *122*(5), 1126–1138.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450.

Atance, C. M., & O'Neill, D. K. (2001). Episodic future thinking. *Trends in Cognitive Sciences*, *5*(12), 533–539.

Auksztulewicz, R., & Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*. https://www.sciencedirect.com/science/article/pii/S0010945216000101

Auksztulewicz, R., Friston, K. J., & Nobre, A. C. (2017). Task relevance modulates the behavioural and neural effects of sensory predictions. *PLoS Biology*, *15*(12), e2003143.

Auksztulewicz, R., Friston, K., & Nobre, A. C. (2016). *Prediction, precision, and context: dynamic causal modelling of MEG and ECoG*.

Averbeck, B. B., & Costa, V. D. (2017). Motivational neural circuits underlying reinforcement learning. *Nature Neuroscience*, *20*(4), 505–512.

Awesome Sports. (2018). *Squash Coaching Blog: Tactics – When to play a volley drop*. http://awsomesports.co.uk/squash-coaching-blog-tactics-play-volley-drop/

Aylward, J., Hales, C., Robinson, E., & Robinson, O. J. (2020). Translating a rodent measure of negative bias into humans: the impact of induced anxiety and unmedicated mood and anxiety disorders. *Psychological Medicine*, *50*(2), 237–246.

Aylward, J., Valton, V., Ahn, W.-Y., Bond, R. L., Dayan, P., Roiser, J. P., & Robinson, O. J. (2019). Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nature Human Behaviour*, *3*(10), 1116–1123.

Aylward, J., Valton, V., Goer, F., Mkrtchian, A., Lally, N., Peters, S., Limbachya, T., & Robinson, O. J. (2017). The impact of induced anxiety on affective response inhibition. *Royal Society Open Science*, *4*(6), 170084.

Azim, E., Mobbs, D., Jo, B., Menon, V., & Reiss, A. L. (2005). Sex differences in brain activation elicited by humor. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(45), 16496–16501.

Baars, B. J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences*, *7*(4), 166–172.

Bach, D. R. (2015). Anxiety-Like Behavioural Inhibition Is Normative under Environmental Threat-Reward Correlations. *PLoS Computational Biology*, *11*(12), e1004646.

Bach, D. R., & Dayan, P. (2017). Algorithms for survival: a comparative perspective on emotions. *Nature Reviews. Neuroscience*, *18*(5), 311–319.

Bach, D. R., & Dolan, R. J. (2012). Knowing how much you don't know: a neural organization of uncertainty estimates. *Nature Reviews. Neuroscience*, *13*(8), 572–586.

Bach, D. R., Hulme, O., Penny, W. D., & Dolan, R. J. (2011). The known unknowns: neural representation of second-order uncertainty, and ambiguity. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(13), 4811–4820.

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(31), 8161–8165.

Bang, D., Kishida, K. T., Lohrenz, T., White, J. P., Laxton, A. W., Tatter, S. B., Fleming, S. M., & Montague, P. R. (2020). Sub-second Dopamine and Serotonin Signaling in Human Striatum during Perceptual Decision-Making. *Neuron*, *108*(5), 999–1010.e6.

Bannert, M. M., & Bartels, A. (2013). Decoding the yellow of a gray banana. *Current Biology: CB*, *23*(22), 2268–2272.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological Bulletin*, *133*(1), 1–24.

Barnes, L. L. B., Harp, D., & Jung, W. S. (2002). Reliability Generalization of Scores on the Spielberger State-Trait Anxiety Inventory. *Educational and Psychological Measurement*, *62*(4), 603–618.

Barone, P., Batardiere, A., Knoblauch, K., & Kennedy, H. (2000). Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. *Journal of Neuroscience*, *20*(9), 3263–3281.

Barraclough, D. J., Conroy, M. L., & Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neuroscience*, *7*(4), 404–410.

Barrett, L. F. (2006). Are Emotions Natural Kinds? *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *1*(1), 28–58.

Barrett, L. F., Quigley, K. S., & Hamilton, P. (2016). An active inference theory of allostasis and interoception in depression. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1708). https://doi.org/10.1098/rstb.2016.0011

Barto, A. G., Sutton, R. S., & Watkins, C. (1989). Learning and sequential decision making. *Learning and Computational*. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.42.7260

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, *76*, 412–427.

Bassett, D. S., Zurn, P., & Gold, J. I. (2018). On the nature and use of models in network neuroscience. *Nature Reviews. Neuroscience*, *19*(9), 566–578.

Basso, D., Chiarandini, M., & Salmaso, L. (2007). Synchronized permutation tests in replicated I×J designs. *Journal of Statistical Planning and Inference*, *137*(8), 2564–2578.

Bast, N., Boxhoorn, S., Super, H., Helfer, B., Polzer, L., Klein, C., Cholemkery, H., & Freitag, C. M. (2021). Atypical Arousal Regulation in children with Autism but not with ADHD as

indicated by pupillometric measures of Locus Coeruleus Activity. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*. https://doi.org/10.1016/j.bpsc.2021.04.010

Bastos, A. M., Litvak, V., Moran, R., Bosman, C. A., Fries, P., & Friston, K. J. (2015). A DCM study of spectral asymmetries in feedforward and feedback connections between visual areas V1 and V4 in the monkey. *NeuroImage*, *108*, 460–475.

Bastos, A. M., Loonis, R., Kornblith, S., Lundqvist, M., & Miller, E. K. (2018). Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(5), 1117–1122.

Bastos, A. M., Lundqvist, M., Waite, A. S., Kopell, N., & Miller, E. K. (2020). Layer and rhythm specificity for predictive routing. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(49), 31459–31469.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*(4), 695–711.

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., De Weerd, P., Kennedy, H., & Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, *85*(2), 390–401.

Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *20*(7), 1391–1397.

Bauer, M., Oostenveld, R., Peeters, M., & Fries, P. (2006). Tactile spatial attention enhances gamma-band activity in somatosensory cortex and reduces low-frequency activity in parieto-occipital areas. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*(2), 490–501.

Bauer, M., Stenner, M.-P., Friston, K. J., & Dolan, R. J. (2014). Attentional modulation of alpha/beta and gamma oscillations reflect functionally distinct processes. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(48), 16117–16125.

Baumeister, R. F., Vohs, K. D., DeWall, C. N., & Zhang, L. (2007). How emotion shapes behavior: feedback, anticipation, and reflection, rather than direct causation. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*, *11*(2), 167–203.

Baum, M. (1970). Extinction of avoidance responding through response prevention (flooding). *Psychological Bulletin*, *74*(4), 276–284.

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1), 129–141.

Bayer, H. M., Lau, B., & Glimcher, P. W. (2007). Statistics of midbrain dopamine neuron spike trains in the awake primate. *Journal of Neurophysiology*, *98*(3), 1428–1439.

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, *52*(2), 336–372.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*(3), 295–307.

Bechara, A., Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain: A Journal of Neurology*, *123 (Pt 11)*, 2189–2202.

Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., Shadlen, M. N., Latham, P. E., & Pouget, A. (2008). Probabilistic population codes for Bayesian decision making. *Neuron*, *60*(6), 1142–1152.

Beddington, J., Cooper, C. L., Field, J., Goswami, U., Huppert, F. A., Jenkins, R., Jones, H. S., Kirkwood, T. B. L., Sahakian, B. J., & Thomas, S. M. (2008). The mental wealth of nations. *Nature*, *455*(7216), 1057–1060.

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, *456*(7219), 245–249.

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

Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *European Journal of Neuroscience*, *27*(7), 1823-1835.

Bellman, R. (1958). On a routing problem. *Quarterly of Applied Mathematics*, *16*(1), 87–90.

Bellman, R. (1966). Dynamic programming. *Science*, *153*(3731), 34–37.

Beltzer, M. L., Adams, S., Beling, P. A., & Teachman, B. A. (2019). Social anxiety and dynamic social reinforcement learning in a volatile environment. *Clinical Psychological Science*, *7*(6), 1372–1388.

Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, *93*(3), 491–507.

Bennett, D., Sutcliffe, K., Tan, N. P. J., Smillie, L. D., & Bode, S. (2021). Anxious and obsessive-compulsive traits are independently associated with valuation of noninstrumental information. *Journal of Experimental Psychology: General, 150*(4), 739.

Berggren, N., & Derakshan, N. (2013). Attentional control deficits in trait anxiety: why you see them and why you don't. *Biological Psychology*, *92*(3), 440–446.

Berke, J. D. (2009). Fast oscillations in cortical-striatal networks switch frequency following rewarding events and stimulant drugs. *The European Journal of Neuroscience*, 30(5), 848–859.

Bernardoni, F., Geisler, D., King, J. A., Javadi, A.-H., Ritschel, F., Murr, J., Reiter, A. M. F., Rössner, V., Smolka, M. N., Kiebel, S., & Ehrlich, S. (2018). Altered Medial Frontal Feedback Learning Signals in Anorexia Nervosa. *Biological Psychiatry*, *83*(3), 235–243.

Berns, G. S., Laibson, D., & Loewenstein, G. (2007). Intertemporal choice--toward an integrative framework. *Trends in Cognitive Sciences*, *11*(11), 482–488.

Berridge, K. C. (1996). Food reward: brain substrates of wanting and liking. *Neuroscience and Biobehavioral Reviews*, *20*(1), 1–25.

Berridge, K. C. (2000). Reward learning: Reinforcement, incentives, and expectations. In *Psychology of Learning and Motivation* (Vol. 40, pp. 223–278). Academic Press.

Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiology & Behavior*, *81*(2), 179–209.

Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology*, *191*(3), 391–431.

Berridge, K. C. (2012). From prediction error to incentive salience: mesolimbic computation of reward motivation. *The European Journal of Neuroscience*, *35*(7), 1124–1143.

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research. Brain Research Reviews*, *28*(3), 309–369.

Berridge, K. C., & Robinson, T. E. (2003). Parsing reward. *Trends in Neurosciences*, *26*(9), 507–513.

Betzel, R. F., & Bassett, D. S. (2017). Generative models for network neuroscience: prospects and promise. *Journal of the Royal Society, Interface / the Royal Society*, *14*(136). https://doi.org/10.1098/rsif.2017.0623

Bhatia, V., & Tandon, R. K. (2005). Stress and the gastrointestinal tract. *Journal of Gastroenterology and Hepatology*, *20*(3), 332–339.

Bielajew, C. H., & Harris, T. (1991). Self-stimulation: a rewarding decade. *Journal of Psychiatry & Neuroscience: JPN*, *16*(3), 109–114.

Birkett, M. A. (2011). The Trier Social Stress Test protocol for inducing psychological stress. *Journal of Visualized Experiments: JoVE*, *56*. https://doi.org/10.3791/3238

Birrell, J., Meares, K., Wilkinson, A., & Freeston, M. (2011). Toward a definition of intolerance of uncertainty: a review of factor analytical studies of the Intolerance of Uncertainty Scale. *Clinical Psychology Review*, *31*(7), 1198–1208.

Bishop, S., Duncan, J., Brett, M., & Lawrence, A. D. (2004). Prefrontal cortical function and anxiety: controlling attention to threat-related stimuli. *Nature Neuroscience*, *7*(2), 184–188.

Bishop, S., & Forster, S. (2013). Trait anxiety, neuroticism, and the brain basis of vulnerability to affective disorder. In J. Armony (Ed.), *The Cambridge handbook of*

human affective neuroscience , (pp (Vol. 667, pp. 553–574). Cambridge University Press, xii.

Bishop, S. J. (2007). Neurocognitive mechanisms of anxiety: an integrative account. *Trends in Cognitive Sciences*, *11*(7), 307–316.

Bishop, S. J. (2008). Neural mechanisms underlying selective attention to threat. *Annals of the New York Academy of Sciences*, *1129*, 141–152.

Bishop, S. J. (2009). Trait anxiety and impoverished prefrontal control of attention. *Nature Neuroscience*, *12*(1), 92–98.

Bishop, S. J., Duncan, J., & Lawrence, A. D. (2004). State anxiety modulation of the amygdala response to unattended threat-related stimuli. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *24*(46), 10364–10368.

Bishop, S. J., & Gagne, C. (2018). Anxiety, Depression, and Decision Making: A Computational Perspective. *Annual Review of Neuroscience*, *41*, 371–388.

Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *Journal of Psychosomatic Research*, *52*(2), 69–77.

Blanchard, D. C., & Blanchard, R. J. (1988). Ethoexperimental approaches to the biology of emotion. *Annual Review of Psychology*, *39*, 43–68.

Blanchette, I., & Richards, A. (2003). Anxiety and the interpretation of ambiguous information: beyond the emotion-congruent effect. *Journal of Experimental Psychology. General*, *132*(2), 294–309.

Bland, A. R., & Schaefer, A. (2012). Different varieties of uncertainty in human decision-making. *Frontiers in Neuroscience*, *6*, 85.

Bleil, M. E., Gianaros, P. J., Jennings, J. R., Flory, J. D., & Manuck, S. B. (2008). Trait negative affect: toward an integrated model of understanding psychological risk for impairment in cardiac autonomic function. *Psychosomatic Medicine*, *70*(3), 328–337.

Blokpoel, M., Kwisthout, J., & van Rooij, I. (2012). When Can Predictive Brains be Truly Bayesian? *Frontiers in Psychology*, *3*, 406.

Boesch, M., Sefidan, S., Ehlert, U., Annen, H., Wyss, T., Steptoe, A., & La Marca, R. (2014). Mood and autonomic responses to repeated exposure to the Trier Social Stress Test for Groups (TSST-G). *Psychoneuroendocrinology*, *43*, 41–51.

Bogdan, R., & Pizzagalli, D. A. (2006). Acute stress reduces reward responsiveness: implications for depression. *Biological Psychiatry*, *60*(10), 1147–1154.

Bonnelle, V., Veromann, K.-R., Burnett Heyes, S., Lo Sterzo, E., Manohar, S., & Husain, M. (2015). Characterization of reward and effort mechanisms in apathy. *Journal of Physiology, Paris*, *109*(1-3), 16–26.

Borkovec, T. D., Alcaine, O., & Behar, E. (2004). Avoidance theory of worry and generalized anxiety disorder. *Generalized Anxiety Disorder: Advances in Research and Practice*, *2004*. https://books.google.co.uk/books?hl=en&lr=&id=klrXMNzWhNMC&oi=fnd&pg=PA77&dq=avoidance+anxiety&ots=mvHrY2WSba&sig=sFx9z3q2dvJxHdSD7xyF309yCAg

Borkovec, T. D., Hazlett-Stevens, H., & Diaz, M. L. (1999). The role of positive beliefs about worry in generalized anxiety disorder and its treatment. *Clinical Psychology & Psychotherapy*, *6*(2), 126–138.

Botvinick, M., & An, J. (2009). Goal-directed decision making in prefrontal cortex: A computational framework. *Advances in Neural Information Processing Systems*, *21*, 169–176.

Botvinick, M. M., Niv, Y., & Barto, A. G. (2009). Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition*, *113*(3), 262–280.

Bouchard, S., Dumoulin, S., Robillard, G., Guitard, T., Klinger, É., Forget, H., Loranger, C., & Roucaut, F. X. (2017). Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: a three-arm randomised controlled trial. *The British Journal of Psychiatry: The Journal of Mental Science*, *210*(4), 276–283.

Bowers, J. S., & Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin*, *138*(3), 389–414.

Bradfield, L. A., Dezfouli, A., van Holstein, M., Chieng, B., & Balleine, B. W. (2015). Medial Orbitofrontal Cortex Mediates Outcome Retrieval in Partially Observable Task Situations. *Neuron*, *88*(6), 1268–1280.

Bray, S., Shimojo, S., & O'Doherty, J. P. (2010). Human medial orbitofrontal cortex is recruited during experience of imagined and real rewards. *Journal of neurophysiology*, *103*(5), 2506-2512.

Brazil, I. A., Mathys, C. D., Popma, A., Hoppenbrouwers, S. S., & Cohn, M. D. (2017). Representational Uncertainty in the Brain During Threat Conditioning and the Link With Psychopathic Traits. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *2*(8), 689–695.

Britton, J. C., Grillon, C., Lissek, S., Norcross, M. A., Szuhany, K. L., Chen, G., Ernst, M., Nelson, E. E., Leibenluft, E., Shechner, T., & Pine, D. S. (2013). Response to learned threat: An FMRI study in adolescent and adult anxiety. *The American Journal of Psychiatry*, *170*(10), 1195–1204.

Brodski, A., Paasch, G.-F., Helbling, S., & Wibral, M. (2015). The Faces of Predictive Coding. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(24), 8997–9006.

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*, *68*(5), 815–834.

Bromberg-Martin, E. S., Matsumoto, M., Hong, S., & Hikosaka, O. (2010). A pallidus-habenula-dopamine pathway signals inferred stimulus values. *Journal of Neurophysiology*, *104*(2), 1068–1076.

Broom, D. C., Jutkiewicz, E. M., Folk, J. E., Traynor, J. R., Rice, K. C., & Woods, J. H. (2002). Nonpeptidic δ-opioid Receptor Agonists Reduce Immobility in the Forced Swim Assay in Rats. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *26*(6), 744–755.

Broome, M. R., & de Cates, A. (2015). Choosing death in depression: a commentary on “Treatment-resistant major depressive disorder and assisted dying” [Review of

Choosing death in depression: a commentary on "Treatment-resistant major depressive disorder and assisted dying"]. *Journal of Medical Ethics*, *41*(8), 586–587. jme.bmj.com.

Brown, H. R., & Friston, K. J. (2012). Dynamic causal modelling of precision and synaptic gain in visual perception—an EEG study. *NeuroImage*, *63*(1), 223–231.

Brown, H. R., & Friston, K. J. (2013). The functional anatomy of attention: a DCM study. *Frontiers in Human Neuroscience*, *7*, 784.

Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience*, *18*(4), 590–596.

Browning, M., Holmes, E. A., Charles, M., Cowen, P. J., & Harmer, C. J. (2012). Using attentional bias modification as a cognitive vaccine against depression. *Biological Psychiatry*, *72*(7), 572–579.

Browning, M., Holmes, E. A., & Harmer, C. J. (2010). The modification of attentional bias to emotional information: A review of the techniques, mechanisms, and relevance to emotional disorders. *Cognitive, Affective & Behavioral Neuroscience*, *10*(1), 8–20.

Bublatzky, F., Alpers, G. W., & Pittig, A. (2017). From avoidance to approach: The influence of threat-of-shock on reward-based decision making. *Behaviour Research and Therapy*, *96*, 47–56.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*(2), 49–57.

Buffalo, E. A., Fries, P., & Landman, R. (2011). Laminar differences in gamma and alpha coherence in the ventral stream. *Proceedings of the*. https://www.pnas.org/content/108/27/11262.short

Bunzeck, N., Guitart-Masip, M., Dolan, R. J., & Düzel, E. (2011). Contextual novelty modulates the neural dynamics of reward anticipation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(36), 12816–12822.

Burgdorf, J., & Panksepp, J. (2006). The neurobiology of positive emotions. *Neuroscience and Biobehavioral Reviews*, *30*(2), 173–187.

Bush, G., Vogt, B. A., Holmes, J., Dale, A. M., Greve, D., Jenike, M. A., & Rosen, B. R. (2002). Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(1), 523–528.

Bush, R. R., & Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, *58*(5), 313–323.

Butler, G., & Mathews, A. (1983). Cognitive processes in anxiety. *Advances in Behaviour Research and Therapy*, *5*(1), 51–62.

Butler, G., & Mathews, A. (1987). Anticipatory anxiety and risk perception. *Cognitive Therapy and Research*, *11*(5), 551–565.

Butz, M. V., & Kutter, E. F. (2016). *How the Mind Comes into Being: Introducing Cognitive Science from a Functional and Computational Perspective*. Oxford University Press.

Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes. *Nature Reviews. Neuroscience*, *13*(6), 407–420.

Cahill, L., Prins, B., Weber, M., & McGaugh, J. L. (1994). Beta-adrenergic activation and memory for emotional events. *Nature*, *371*(6499), 702–704.

Calhoon, G. G., & Tye, K. M. (2015). Resolving the neural circuits of anxiety. *Nature Neuroscience*, *18*(10), 1394–1404.

Cameron, J., & Pierce, W. D. (1994). Reinforcement, Reward, and Intrinsic Motivation: A Meta-Analysis. *Review of Educational Research*, *64*(3), 363–423.

Campbell, M., Hoane, A. J., & Hsu, F.-H. (2002). Deep Blue. *Artificial Intelligence*, *134*(1), 57–83.

Cannon, C. M., & Palmiter, R. D. (2003). Reward without dopamine. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *23*(34), 10827–10831.

Cannon, W. B. (1932). Homeostasis. *The Wisdom of the Body. Norton, New York*. https://www.researchgate.net/profile/Kelvin_Rodolfo/publication/304496046_What_is_homeostsis/links/57714c6e08ae842225ac1402/What-is-homeostsis.pdf

Cao, L., Thut, G., & Gross, J. (2017). The role of brain oscillations in predicting self-generated sounds. *NeuroImage*, *147*, 895–903.

Carleton, R. N. (2016). Into the unknown: A review and synthesis of contemporary models involving uncertainty. *Journal of Anxiety Disorders*, *39*, 30–43.

Carleton, R. N., Mulvogue, M. K., Thibodeau, M. A., McCabe, R. E., Antony, M. M., & Asmundson, G. J. G. (2012). Increasingly certain about uncertainty: Intolerance of uncertainty across anxiety and depression. *Journal of Anxiety Disorders*, *26*(3), 468–479.

Carling, K. (2000). Resistant outlier rules and the non-Gaussian case. *Computational Statistics & Data Analysis*, *33*(3), 249–258.

Carpenter, R. H. S. (2004). Homeostasis: a plea for a unified approach. *Advances in Physiology Education*, *28*(1-4), 180–187.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, *67*(2), 319–333.

Casadesús, J., & D'Ari, R. (2002). Memory in bacteria and phage. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *24*(6), 512–518.

Cassell, M. D., & Wright, D. J. (1986). Topography of projections from the medial prefrontal cortex to the amygdala in the rat. *Brain Research Bulletin*, *17*(3), 321–333.

Cavanagh, J. F., Bismark, A. W., Frank, M. J., & Allen, J. J. B. (2019). Multiple Dissociations Between Comorbid Depression and Anxiety on Reward and Punishment Processing: Evidence From Computationally Informed EEG. *Computational Psychiatry (Cambridge, Mass.)*, *3*, 1–17.

Cavanagh, J. F., Frank, M. J., & Allen, J. J. B. (2011). Social stress reactivity alters reward and punishment learning. *Social Cognitive and Affective Neuroscience*, *6*(3), 311–320.

Cavanagh, J. F., Meyer, A., & Hajcak, G. (2017). Error-Specific Cognitive Control Alterations in Generalized Anxiety Disorder. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *2*(5), 413–420.

Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: meta-analytic evidence. *Journal of Physiology, Paris*, *109*(1-3), 3–15.

Chalmers, J. A., Quintana, D. S., Abbott, M. J.-A., & Kemp, A. H. (2014). Anxiety Disorders are Associated with Reduced Heart Rate Variability: A Meta-Analysis. *Frontiers in Psychiatry / Frontiers Research Foundation*, *5*, 80.

Chambers, J. A., Power, K. G., & Durham, R. C. (2004). The relationship between trait vulnerability and anxiety and depressive diagnoses at long-term follow-up of Generalized Anxiety Disorder. *Journal of Anxiety Disorders*, *18*(5), 587–607.

Chang, Y., Xu, J., Pang, X., Sun, Y., Zheng, Y., & Liu, Y. (2015). Mismatch negativity indices of enhanced preattentive automatic processing in panic disorder as measured by a multi-feature paradigm. *Biological Psychology*, *105*, 77–82.

Charpentier, C. J., Aylward, J., Roiser, J. P., & Robinson, O. J. (2017). Enhanced Risk Aversion, But Not Loss Aversion, in Unmedicated Pathological Anxiety. *Biological Psychiatry*, *81*(12), 1014–1022.

Chew, B., Hauser, T. U., Papoutsi, M., Magerkurth, J., Dolan, R. J., & Rutledge, R. B. (2019). Endogenous fluctuations in the dopaminergic midbrain drive behavioral choice variability. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1900872116

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: a global return on investment analysis. *The Lancet. Psychiatry*, *3*(5), 415–424.

Chumbley, J. R., Flandin, G., Bach, D. R., Daunizeau, J., Fehr, E., Dolan, R. J., & Friston, K. J. (2012). Learning and generalization under ambiguity: an fMRI study. *PLoS Computational Biology*, *8*(1), e1002346.

Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience*, *33*, 269–298.

Cisler, J. M., & Koster, E. H. W. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clinical Psychology Review*, *30*(2), 203–216.

Clark, A. (2013a). The many faces of precision (Replies to commentaries on “Whatever next? Neural prediction, situated agents, and the future of cognitive science”). *Frontiers in Psychology*, *4*, 270.

Clark, A. (2013b). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and Brain Sciences*, *36*(3), 181–204.

Clark, A. (2015). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press.

Clark, D. A., & Beck, A. T. (2011). *Cognitive Therapy of Anxiety Disorders: Science and Practice*. Guilford Press.

Clark, J. E., Watson, S., & Friston, K. J. (2018). What is mood? A computational perspective. *Psychological Medicine*, *48*(14), 2277–2284.

Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B., & Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature*, *482*(7383), 85–88.

Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *NeuroImage*, *35*(2), 968–978.

Cole, D. M., Diaconescu, A. O., Pfeiffer, U. J., Brodersen, K. H., Mathys, C. D., Julkowski, D., Ruhrmann, S., Schilbach, L., Tittgemeyer, M., Vogeley, K., & Stephan, K. E. (2020). Atypical processing of uncertainty in individuals at risk for psychosis. *NeuroImage. Clinical*, *26*, 102239.

Conant, R. C., & Ross Ashby, W. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systems Science*, *1*(2), 89–97.

Cooper, J. A., Arulpragasam, A. R., & Treadway, M. T. (2018). Anhedonia in depression: biological mechanisms and computational models. *Current Opinion in Behavioral Sciences*, *22*, 128–135.

Corbit, L. H., & Balleine, B. W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *25*(4), 962–970.

Corbit, L. H., Muir, J. L., & Balleine, B. W. (2001). The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *21*(9), 3251–3260.

Corcoran, C. M., & Cecchi, G. A. (2018). Computational Approaches to Behavior Analysis in Psychiatry. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *43*(1), 225–226.

Corlett, P. R., & Fletcher, P. C. (2014). Computational psychiatry: a Rosetta Stone linking the brain to mental illness. *The Lancet. Psychiatry*, *1*(5), 399–402.

Cornwell, B. R., Alvarez, R. P., Lissek, S., Kaplan, R., Ernst, M., & Grillon, C. (2011). Anxiety overrides the blocking effects of high perceptual load on amygdala reactivity to threat-related distractors. *Neuropsychologia*, *49*(5), 1363–1368.

Cornwell, B. R., Baas, J. M. P., Johnson, L., Holroyd, T., Carver, F. W., Lissek, S., & Grillon, C. (2007). Neural responses to auditory stimulus deviance under threat of electric shock revealed by spatially-filtered magnetoencephalography. *NeuroImage*, *37*(1), 282–289.

Cornwell, B. R., Garrido, M. I., Overstreet, C., Pine, D. S., & Grillon, C. (2017). The Unpredictive Brain Under Threat: A Neurocomputational Account of Anxious Hypervigilance. *Biological Psychiatry*, *82*(6), 447–454.

Costa, V. D., Dal Monte, O., Lucas, D. R., Murray, E. A., & Averbeck, B. B. (2016). Amygdala and Ventral Striatum Make Distinct Contributions to Reinforcement Learning. *Neuron*, *92*(2), 505–517.

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, *10*(7), 294–300.

Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron*, *29*(2), 537–545.

Croxson, P. L., Walton, M. E., O'Reilly, J. X., Behrens, T. E. J., & Rushworth, M. F. S. (2009). Effort-Based Cost–Benefit Valuation and the Human Brain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(14), 4531–4541.

Cumming, S. R., & Harris, L. M. (2001). The impact of anxiety on the accuracy of diagnostic decision-making. *Stress and Health: Journal of the International Society for the Investigation of Stress*, *17*(5), 281–286.

Cunillera, T., Fuentemilla, L., Periañez, J., Marco-Pallarès, J., Krämer, U. M., Càmara, E., Münte, T. F., & Rodríguez-Fornells, A. (2012). Brain oscillatory activity associated with task switching and feedback processing. *Cognitive, Affective & Behavioral Neuroscience*, *12*(1), 16–33.

Da Costa, L., Parr, T., Sajid, N., Veselic, S., Neacsu, V., & Friston, K. (2020). Active inference on discrete state-spaces: a synthesis. In *arXiv [q-bio.NC]*. arXiv. http://arxiv.org/abs/2001.07203

Dalgleish, H. W., Russell, L. E., Packer, A. M., Roth, A., Gauld, O. M., Greenstreet, F., Thompson, E. J., & Häusser, M. (2020). How many neurons are sufficient for perception of cortical activity? *eLife*, *9*. https://doi.org/10.7554/eLife.58889

Dalton, G. L., Wang, N. Y., Phillips, A. G., & Floresco, S. B. (2016). Multifaceted Contributions by Different Regions of the Orbitofrontal and Medial Prefrontal Cortex to Probabilistic Reversal Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(6), 1996–2006.

Damasio, A. R. (2003). *Looking for Spinoza: Joy, sorrow, and the feeling brain*. Houghton Mifflin Harcourt.

D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, *319*(5867), 1264–1267.

Darwin, C. (1956). The expression of the emotions in man and animals. *The American Journal of the Medical Sciences*, *232*(4), 477.

Dauer, W., & Przedborski, S. (2003). Parkinson's disease: mechanisms and models. *Neuron*, *39*(6), 889–909.

Daunizeau, J., den Ouden, H. E. M., Pessiglione, M., Kiebel, S. J., Friston, K. J., & Stephan, K. E. (2010). Observing the observer (II): deciding when to decide. *PloS One*, *5*(12), e15555.

Daunizeau, J., den Ouden, H. E. M., Pessiglione, M., Kiebel, S. J., Stephan, K. E., & Friston, K. J. (2010). Observing the observer (I): meta-bayesian models of learning and decision-making. *PloS One*, *5*(12), e15554.

David, N., Schneider, T. R., Peiker, I., Al-Jawahiri, R., Engel, A. K., & Milne, E. (2016). Variability of cortical oscillation patterns: A possible endophenotype in autism spectrum disorders? *Neuroscience and Biobehavioral Reviews*, *71*, 590–600.

Davidson, R. J. (2002). Anxiety and affective style: role of prefrontal cortex and amygdala. *Biological Psychiatry*, *51*(1), 68–80.

Davis, M. (1992). The role of the amygdala in fear and anxiety. *Annual Review of Neuroscience*, *15*(1), 353–375.

Davis, M. (2000). The role of the amygdala in conditioned and unconditioned fear and anxiety. *The Amygdala*, *2*, 213–287.

Davis, M., Walker, D. L., Miles, L., & Grillon, C. (2010). Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *35*(1), 105–135.

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature*, *446*(7137), 794–796.

Dawes, C. T., Loewen, P. J., Schreiber, D., Simmons, A. N., Flagan, T., McElreath, R., Bokemper, S. E., Fowler, J. H., & Paulus, M. P. (2012). Neural basis of egalitarian behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(17), 6479–6483.

Daw, N. D. (2003). *Reinforcement learning models of the dopamine system and their behavioral implications*. Carnegie Mellon University.

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, *16*(2), 199–204.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215.

Daw, N. D., & O'Doherty, J. P. (2014). *Multiple systems for value learning neuroeconomics: Decision making and the brain (pp. 393--410)*. Amsterdam, The Netherlands: Elsevier.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*. https://www.nature.com/articles/nature04766

Daw, N. D., & Tobler, P. N. (2014). Chapter 15 - Value Learning through Reinforcement: The Basics of Dopamine and Reinforcement Learning. In P. W. Glimcher & E. Fehr (Eds.), *Neuroeconomics (Second Edition)* (pp. 283–298). Academic Press.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711.

Dayan, P. (1993). Improving Generalization for Temporal Difference Learning: The Successor Representation. *Neural Computation*, *5*(4), 613–624.

Dayan, P. (2012a). Instrumental vigour in punishment and reward. *The European Journal of Neuroscience*, *35*(7), 1152–1168.

Dayan, P. (2012b). Twenty-five lessons from computational neuromodulation. *Neuron*, *76*(1), 240–256.

Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective & Behavioral Neuroscience*, *8*(4), 429–453.

Dayan, P., Hinton, G. E., Neal, R. M., & Zemel, R. S. (1995). The Helmholtz machine. *Neural Computation*, *7*(5), 889–904.

Dayan, P., & Huys, Q. J. M. (2008). Serotonin, inhibition, and negative mood. *PLoS Computational Biology*, *4*(2), e4.

Dayan, P., Kakade, S., & Montague, P. R. (2000). Learning and selective attention. *Nature Neuroscience*, *3 Suppl*, 1218–1223.

Dayan, P., & Long, T. (1998). Statistical Models of Conditioning. In M. I. Jordan, M. J. Kearns, & S. A. Solla (Eds.), *Advances in Neural Information Processing Systems 10* (pp. 117–123). MIT Press.

Dayan, P., & Yu, A. J. (2003). Uncertainty and Learning. *IETE Journal of Research*, *49*(2-3), 171–181.

Dayan, P., & Yu, A. J. (2006). Phasic norepinephrine: a neural interrupt signal for unexpected events. *Network*, *17*(4), 335–350.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *25*(50), 11730–11737.

de Berker, A. O. (2017). *Computation and representation in decision making and emotion*. Doctoral thesis, UCL (University College London). https://discovery.ucl.ac.uk/id/eprint/1553367/

de Berker, A. O., Kurth-Nelson, Z., Rutledge, R. B., Bestmann, S., & Dolan, R. J. (2019). Computing Value from Quality and Quantity in Human Decision-Making. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *39*(1), 163–176.

de Berker, A. O., Rutledge, R. B., Mathys, C., Marshall, L., Cross, G. F., Dolan, R. J., & Bestmann, S. (2016). Computations of uncertainty mediate acute stress responses in humans. *Nature Communications*, *7*, 10996.

De Boer, S. F., & Koolhaas, J. M. (2003). Defensive burying in rodents: ethology, neurobiology and psychopharmacology. *European Journal of Pharmacology*, *463*(1-3), 145–161.

Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268.

Deci, E. L., & Ryan, R. M. (2010). Intrinsic Motivation. In *The Corsini Encyclopedia of Psychology*. John Wiley & Sons, Inc. https://doi.org/10.1002/9780470479216.corpsy0467

Dedovic, K., Renwick, R., Mahani, N. K., Engert, V., Lupien, S. J., & Pruessner, J. C. (2005). The Montreal Imaging Stress Task: using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain. *Journal of Psychiatry & Neuroscience: JPN*, *30*(5), 319–325.

Dehaene, G. P., Coen-Cagli, R., & Pouget, A. (2021). Investigating the representation of uncertainty in neuronal circuits. *PLoS Computational Biology*, *17*(2), e1008138.

de Jong, P. J., Merckelbach, H., Bögels, S., & Kindt, M. (1998). Illusory correlation and social anxiety. *Behaviour Research and Therapy*, *36*(11), 1063–1073.

de la Mora, M. P., Gallegos-Cari, A., Arizmendi-García, Y., Marcellino, D., & Fuxe, K. (2010). Role of dopamine receptor mechanisms in the amygdaloid modulation of fear and anxiety: Structural and functional analysis. *Progress in Neurobiology*, *90*(2), 198–216.

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? *Trends in Cognitive Sciences*, *22*(9), 764–779.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Del Viva, M. M., Igliozzi, R., Tancredi, R., & Brizzolara, D. (2006). Spatial and motion integration in children with autism. *Vision Research*, *46*(8-9), 1242–1252.

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, *313*(5787), 684–687.

Deneve, S. (2008). Bayesian spiking neurons I: inference. *Neural Computation*, *20*(1), 91–117.

Deneve, S., Latham, P. E., & Pouget, A. (2001). Efficient computation and cue integration with noisy population codes. *Nature Neuroscience*, *4*(8), 826–831.

Denny-Brown, D. (1962). *The basal ganglia and their relation to disorders of movement*. Oxford University Press.

den Ouden, H. E. M., Friston, K. J., Daw, N. D., McIntosh, A. R., & Stephan, K. E. (2009). A Dual Role for Prediction Error in Associative Learning. *Cerebral Cortex* , *19*(5), 1175–1185.

den Ouden, H. E. M., Kok, P., & de Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology*, *3*, 548.

Derakshan, N., Ansari, T. L., Hansard, M., Shoker, L., & Eysenck, M. W. (2009). Anxiety, inhibition, efficiency, and effectiveness. An investigation using antisaccade task. *Experimental Psychology*, *56*(1), 48–55.

Der-Avakian, A., & Markou, A. (2012). The neurobiology of anhedonia and other reward-related deficits. *Trends in Neurosciences*, *35*(1), 68–77.

Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology*, *111*(2), 225–236.

Deserno, L., Boehme, R., Mathys, C., Katthagen, T., Kaminski, J., Stephan, K. E., Heinz, A., & Schlagenhauf, F. (2020). Volatility Estimates Increase Choice Switching and Relate to Prefrontal Activity in Schizophrenia. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *5*(2), 173–183.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.

de Visser, L., van der Knaap, L. J., van de Loo, A. J. A. E., van der Weerd, C. M. M., Ohl, F., & van den Bos, R. (2010). Trait anxiety affects decision-making differently in healthy men and women: towards gender-specific endophenotypes of anxiety. *Neuropsychologia*, *48*(6), 1598–1606.

de Vries, L., Fouquaet, I., Boets, B., Naulaers, G., & Steyaert, J. (2021). Autism spectrum disorder and pupillometry: A systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews*, *120*, 479–508.

Diaconescu, A. O., Hauke, D. J., & Borgwardt, S. (2019). Models of persecutory delusions: a mechanistic insight into the early stages of psychosis. *Molecular Psychiatry*, *24*(9), 1258–1267.

Diaconescu, A. O., Litvak, V., Mathys, C., Kasper, L., Friston, K. J., & Stephan, K. E. (2017). A computational hierarchy in human cortex. In *arXiv [q-bio.NC]*. arXiv. http://arxiv.org/abs/1709.02323

Diaconescu, A. O., Mathys, C., Weber, L. A. E., Daunizeau, J., Kasper, L., Lomakina, E. I., Fehr, E., & Stephan, K. E. (2014). Inferring on the intentions of others by hierarchical Bayesian learning. *PLoS Computational Biology*, *10*(9), e1003810.

Diaconescu, A. O., Mathys, C., Weber, L. A. E., Kasper, L., Mauer, J., & Stephan, K. E. (2017). Hierarchical prediction errors in midbrain and septum during social learning. *Social Cognitive and Affective Neuroscience*, *12*(4), 618–634.

Diaconescu, A. O., Wellstein, K. V., Kasper, L., Mathys, C., & Stephan, K. E. (2020). Hierarchical Bayesian models of social inference for probing persecutory delusional ideation. *Journal of Abnormal Psychology*, *129*(6), 556–569.

Dias-Ferreira, E., Sousa, J. C., Melo, I., Morgado, P., Mesquita, A. R., Cerqueira, J. J., Costa, R. M., & Sousa, N. (2009). Chronic stress causes frontostriatal reorganization and affects decision-making. *Science*, *325*(5940), 621–625.

Dickinson, A., & Balleine, B. (1994). Motivational control of goal-directed action. *Animal Learning & Behavior*, *22*(1), 1–18.

Dickinson, A., & Balleine, B. (2002). The role of learning in the operation of motivational systems. *Steven's Handbook of Experimental Psychology: Learning, Motivation and Emotion*, *3*, 497–534.

Dickinson, A., Bruyns-Haylett, M., Jones, M., & Milne, E. (2015). Increased peak gamma frequency in individuals with higher levels of autistic traits. *The European Journal of Neuroscience*, *41*(8), 1095–1101.

Dickinson, A., DiStefano, C., Senturk, D., & Jeste, S. S. (2018). Peak alpha frequency is a neural marker of cognitive function across the autism spectrum. *The European Journal of Neuroscience*, *47*(6), 643–651.

Diederen, K. M. J., & Schultz, W. (2015). Scaling prediction errors to reward variability benefits error-driven learning in humans. *Journal of Neurophysiology*, *114*(3), 1628–1640.

Diederen, K. M. J., Spencer, T., Vestergaard, M. D., Fletcher, P. C., & Schultz, W. (2016). Adaptive Prediction Error Coding in the Human Midbrain and Striatum Facilitates Behavioral Adaptation and Learning Efficiency. *Neuron*, *90*(5), 1127–1138.

Dishman, R. K., Nakamura, Y., Garcia, M. E., Thompson, R. W., Dunn, A. L., & Blair, S. N. (2000). Heart rate variability, trait anxiety, and perceived stress among physically fit men and women. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *37*(2), 121–133.

Disney, A. A., Aoki, C., & Hawken, M. J. (2007). Gain modulation by nicotine in macaque v1. *Neuron*. https://www.sciencedirect.com/science/article/pii/S0896627307007684

Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science*, *298*(5596), 1191–1194.

Doll, B. B., Bath, K. G., Daw, N. D., & Frank, M. J. (2016). Variability in Dopamine Genes Dissociates Model-Based and Model-Free Reinforcement Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(4), 1211–1222.

Doll, B. B., Hutchison, K. E., & Frank, M. J. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(16), 6188–6198.

Doll, B. B., Simon, D. A., & Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. *Current Opinion in Neurobiology*, *22*(6), 1075–1081.

Domenech, P., Rheims, S., & Koechlin, E. (2020). Neural mechanisms resolving exploitation-exploration dilemmas in the medial prefrontal cortex. *Science*, 369(6507).

Domjan, M. (2014). *The principles of learning and behavior*. Nelson Education.

Douglas, R. J., & Martin, K. A. (1991). A functional microcircuit for cat visual cortex. *The Journal of Physiology*, *440*, 735–769.

Dowd, E. C., Frank, M. J., Collins, A., Gold, J. M., & Barch, D. M. (2016). Probabilistic Reinforcement Learning in Patients With Schizophrenia: Relationships to Anhedonia and Avolition. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *1*(5), 460–473.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology*, *10*(6), 732–739.

Doya, K. (2008). Modulators of decision making. *Nature Neuroscience*, *11*(4), 410–416.

Doya, K., Ishii, S., Pouget, A., & Rao, R. P. (Eds.). (2007). *Bayesian brain: Probabilistic approaches to neural coding*. MIT press.

Druckman, J. N. (2001). Evaluating framing effects. *Journal Of Economic Psychology*, *22*(1), 91–101.

Dugas, M. J., Freeston, M. H., & Ladouceur, R. (1997). Intolerance of Uncertainty and Problem Orientation in Worry. *Cognitive Therapy and Research*, *21*(6), 593–606.

Dugas, M. J., Gagnon, F., Ladouceur, R., & Freeston, M. H. (1998). Generalized anxiety disorder: a preliminary test of a conceptual model. *Behaviour Research and Therapy*, *36*(2), 215–226.

Dugas, M. J., Hedayati, M., Karavidas, A., Buhr, K., Francis, K., & Phillips, N. A. (2005). Intolerance of Uncertainty and Information Processing: Evidence of Biased Recall and Interpretations. *Cognitive Therapy and Research*, *29*(1), 57–70.

Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., Van Den Hout, M. A., & Baas, J. M. P. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and Anxiety*, *32*(4), 239–253.

Duque, A., & Vázquez, C. (2015). Double attention bias for positive and negative emotional faces in clinical depression: evidence from an eye-tracking study. *Journal of Behavior Therapy and Experimental Psychiatry*, *46*, 107–114.

Duvarci, S., & Pare, D. (2014). Amygdala microcircuits controlling learned fear. *Neuron*, *82*(5), 966–980.

Düzel, E., Bunzeck, N., Guitart-Masip, M., Wittmann, B., Schott, B. H., & Tobler, P. N. (2009). Functional imaging of the human dopaminergic midbrain. *Trends in Neurosciences*, *32*(6), 321–328.

Dymond, S. (2019). Overcoming avoidance in anxiety disorders: The contributions of Pavlovian and operant avoidance extinction methods. *Neuroscience and Biobehavioral Reviews*, *98*, 61–70.

Edwards, E. J., Edwards, M. S., & Lyvers, M. (2015). Cognitive trait anxiety, situational stress, and mental effort predict shifting efficiency: Implications for attentional control theory. *Emotion* , *15*(3), 350–359.

Edwards, M. J., Adams, R. A., Brown, H., Pareés, I., & Friston, K. J. (2012). A Bayesian account of “hysteria.” *Brain: A Journal of Neurology*, *135*(Pt 11), 3495–3512.

Edwards, M. S., Burt, J. S., & Lipp, O. V. (2010). Selective attention for masked and unmasked emotionally toned stimuli: effects of trait anxiety, state anxiety, and test order. *British Journal of Psychology* , *101*(Pt 2), 325–343.

Ehlers, A., & Clark, D. M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy*, *38*(4), 319–345.

Eisenberg, I. W., Bissett, P. G., Zeynep Enkavi, A., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. *Nature Communications*, *10*(1), 2319.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, *221*(4616), 1208–1210.

Eldar, E., Cohen, J. D., & Niv, Y. (2013). The effects of neural gain on attention and learning. *Nature Neuroscience*, *16*(8), 1146–1153.

Eldar, E., Roth, C., Dayan, P., & Dolan, R. J. (2018). Decodability of Reward Learning Signals Predicts Mood Fluctuations. *Current Biology: CB*, *28*(9), 1433–1439.e7.

Eldar, E., Rutledge, R. B., Dolan, R. J., & Niv, Y. (2016). Mood as Representation of Momentum. *Trends in Cognitive Sciences*, *20*(1), 15–24.

Elektra Neuroscience. (2010). Elekta-Neuroscience MaxFilter User's Guide. *Helsinki, Finland: 2010. Software Version 2.2*.

Elias, P. (1955). Predictive coding--I. *IRE Transactions on Information Theory*, *1*(1), 16–24.

Ellwood, I. T., Patel, T., Wadia, V., Lee, A. T., Liptak, A. T., Bender, K. J., & Sohal, V. S. (2017). Tonic or Phasic Stimulation of Dopaminergic Projections to Prefrontal Cortex Causes Mice to Maintain or Deviate from Previously Learned Behavioral Strategies. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 37*(35), 8315–8329.

Endler, N. S., & Kocovski, N. L. (2001). State and trait anxiety revisited. *Journal of Anxiety Disorders*, *15*(3), 231–245.

Enel, P., Wallis, J. D., & Rich, E. L. (2020). Stable and dynamic representations of value in the prefrontal cortex. *eLife*, *9*. https://doi.org/10.7554/eLife.54313

Engel, A. K., & Fries, P. (2010). Beta-band oscillations—signalling the status quo? *Current Opinion in Neurobiology*, *20*(2), 156–165.

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

Enkavi, A. Z., Eisenberg, I. W., & Bissett, P. G. (2019). Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proceedings of the*. https://www.pnas.org/content/116/12/5472.short

Enriquez-Geppert, S., & Barceló, F. (2018). Multisubject Decomposition of Event-related Positivities in Cognitive Control: Tackling Age-related Changes in Reactive Control. Brain Topography, 31(1), 17–34.

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

Etkin, A. (2010). Functional neuroanatomy of anxiety: a neural circuit perspective. *Current Topics in Behavioral Neurosciences*, *2*, 251–277.

Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences*, *15*(2), 85–93.

Everitt, B. J., & Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nature Neuroscience*, *8*(11), 1481–1489.

Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews. Neuroscience*, *9*(4), 292–303.

Fan, L., Sun, Y.-B., Sun, Z.-K., Wang, N., Luo, F., Yu, F., & Wang, J.-Y. (2018). Modulation of auditory sensory memory by chronic clinical pain and acute experimental pain: a mismatch negativity study. *Scientific Reports*, *8*(1), 15673.

Fanselow, M. S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychonomic Bulletin & Review*, *1*(4), 429–438.

Farashahi, S., Donahue, C. H., Khorsand, P., Seo, H., Lee, D., & Soltani, A. (2017). Metaplasticity as a Neural Substrate for Adaptive Learning and Choice under Uncertainty. *Neuron*, *94*(2), 401–414.e6.

Fehr, E., & Schmidt, K. M. (1999). A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868.

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, *4*, 215.

Feldman, P. J., Cohen, S., Hamrick, N., & Lepore, S. J. (2004). Psychological stress, appraisal, emotion and Cardiovascular response in a public speaking task. *Psychology & Health*, *19*(3), 353–368.

Fellows, L. K. (2006). Deciding how to decide: ventromedial frontal lobe damage affects information acquisition in multi-attribute decision making. *Brain: A Journal of Neurology*, *129*(Pt 4), 944–952.

Ferguson, K. A., & Cardin, J. A. (2020). Mechanisms underlying gain modulation in the cortex. *Nature Reviews. Neuroscience*, *21*(2), 80–92.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. https://psycnet.apa.org/psycinfo/2004-21805-000/

Filimon, F., Nelson, J. D., Sejnowski, T. J., Sereno, M. I., & Cottrell, G. W. (2020). The ventral striatum dissociates information expectation, reward anticipation, and reward receipt. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(26), 15200–15208.

Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S., & Wyart, V. (2019). Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nature Neuroscience*, *22*(12), 2066–2077.

Findling, C., & Wyart, V. (2021). Computation noise in human learning and decision-making: origin, impact, function. *Current Opinion in Behavioral Sciences*, *38*, 124–132.

Fineberg, N. A., Haddad, P. M., Carpenter, L., Gannon, B., Sharpe, R., Young, A. H., Joyce, E., Rowe, J., Wellsted, D., Nutt, D. J., & Others. (2013). The size, burden and cost of disorders of the brain in the UK. *Journal of Psychopharmacology* , *27*(9), 761–770.

Fiorillo, C. D. (2013). Two dimensions of value: dopamine neurons represent reward but not aversiveness. *Science*, *341*(6145), 546–549.

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*(5614), 1898–1902.

Fisher, P. L., & Durham, R. C. (1999). Recovery rates in generalized anxiety disorder following psychological therapy: an analysis of clinically significant change in the STAI-T across outcome studies since 1990. *Psychological Medicine*, *29*(6), 1425–1434.

FitzGerald, T. H. B., Dolan, R. J., & Friston, K. (2015). Dopamine, reward learning, and active inference. *Frontiers in Computational Neuroscience*, *9*, 136.

FitzGerald, T. H. B., Dolan, R. J., & Friston, K. J. (2014). Model averaging, optimal inference, and habit formation. *Frontiers in Human Neuroscience*, *8*, 457.

FitzGerald, T. H. B., Friston, K. J., & Dolan, R. J. (2012). Action-specific value signals in reward-related regions of the human brain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(46), 16417–16423a.

FitzGerald, T. H. B., Moran, R. J., Friston, K. J., & Dolan, R. J. (2015). Precision and neuronal dynamics in the human posterior parietal cortex during evidence accumulation. *NeuroImage*, *107*, 219–228.

FitzGerald, T. H. B., Seymour, B., & Dolan, R. J. (2009). The role of human orbitofrontal cortex in value comparison for incommensurable objects. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(26), 8388–8395.

Flagel, S. B., Akil, H., & Robinson, T. E. (2009). Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction. *Neuropharmacology*, *56 Suppl 1*, 139–148.

Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., & Czuj, A. (2011). A selective role for dopamine in stimulus–reward learning. *Nature*. https://www.nature.com/articles/nature09588

Flandin, G., & Friston, K. J. (2019). Analysis of family-wise error rates in statistical parametric mapping using random field theory. *Human Brain Mapping*, *40*(7), 2052–2054.

Fletcher, P. C., Anderson, J. M., Shanks, D. R., Honey, R., Carpenter, T. A., Donovan, T., Papadakis, N., & Bullmore, E. T. (2001). Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. *Nature Neuroscience*, *4*(10), 1043–1048.

Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews. Neuroscience*, *10*(1), 48–58.

Forster, S., Nunez Elizalde, A. O., Castle, E., & Bishop, S. J. (2015). Unraveling the anxious mind: anxiety, worry, and frontal engagement in sustained attention versus off-task processing. *Cerebral Cortex* , *25*(3), 609–618.

Fouragnan, E., Retzler, C., & Philiastides, M. G. (2018). Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis. *Human Brain Mapping*, *39*(7), 2887–2906.

Franklin, N. T., & Frank, M. J. (2015). A cholinergic feedback circuit to regulate striatal population uncertainty and optimize reinforcement learning. *eLife*, *4*. https://doi.org/10.7554/eLife.12029

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience*, *17*(1), 51–72.

Frank, M. J. (2011). Computational models of motivated action selection in corticostriatal circuits. *Current Opinion in Neurobiology*, *21*(3), 381–386.

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(41), 16311–16316.

Frank, M. J., Seeberger, L. C., & O'reilly, R. C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science*, *306*(5703), 1940–1943.

Frässle, S., Aponte, E. A., Bollmann, S. A., & Brodersen, K. H. (2021). TAPAS: an open-source software package for Translational Neuromodeling and Computational Psychiatry. *bioRxiv*. https://www.biorxiv.org/content/10.1101/2021.03.12.435091v1.abstract

Frässle, S., Yao, Y., Schöbi, D., Aponte, E. A., Heinzle, J., & Stephan, K. E. (2018). Generative models for clinical applications in computational psychiatry. *Wiley Interdisciplinary Reviews. Cognitive Science*, *9*(3), e1460.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature*, *40*(2), 351–401.

Freud, S. (1961). The ego and the id. In *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XIX (1923-1925): The Ego and the Id and Other Works* (pp. 1–66). pep-web.org.

Freund, T. F., Powell, J. F., & Smith, A. D. (1984). Tyrosine hydroxylase-immunoreactive boutons in synaptic contact with identified striatonigral neurons, with particular reference to dendritic spines. *Neuroscience*, *13*(4), 1189-1215.

Friedman, B. H. (2007). An autonomic flexibility–neurovisceral integration model of anxiety and cardiac vagal tone. *Biological Psychology*, *74*(2), 185–199.

Friedman, B. H., & Thayer, J. F. (1998). Autonomic balance revisited: panic anxiety and heart rate variability. *Journal of Psychosomatic Research*, *44*(1), 133–151.

Fries, P., Schröder, J.-H., Roelfsema, P. R., Singer, W., & Engel, A. K. (2002). Oscillatory neuronal synchronization in primary visual cortex as a correlate of stimulus selection. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *22*(9), 3739–3754.

Friston, K. (2003). Learning and inference in the brain. *Neural Networks: The Official Journal of the International Neural Network Society*, *16*(9), 1325–1352.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 815–836.

Friston, K. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, *4*(11), e1000211.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews. Neuroscience*, *11*(2), 127–138.

Friston, K. (2012). The history of the future of the Bayesian brain. *NeuroImage*, *62*(2), 1230–1233.

Friston, K. (2013). Life as we know it. *Journal of the Royal Society, Interface / the Royal Society*, *10*(86), 20130475.

Friston, K., Brown, H. R., Siemerkus, J., & Stephan, K. E. (2016). The dysconnection hypothesis. *Schizophrenia Research*, *176*(2-3), 83–94.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active Inference: A Process Theory. *Neural Computation*, *29*(1), 1–49.

Friston, K. J. (2008). Variational filtering. *NeuroImage*, *41*(3), 747–766.

Friston, K. J. (2017). Precision Psychiatry [Review of *Precision Psychiatry*]. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *2*(8), 640–643. discovery.ucl.ac.uk.

Friston, K. J. (2019). Waves of prediction [Review of *Waves of prediction*]. *PLoS Biology*, *17*(10), e3000426.

Friston, K. J., Lawson, R., & Frith, C. D. (2013). On hyperpriors and hypopriors: comment on Pellicano and Burr [Review of *On hyperpriors and hypopriors: comment on Pellicano and Burr*]. *Trends in Cognitive Sciences*, *17*(1), 1. cell.com.

Friston, K. J., Lin, M., Frith, C. D., Pezzulo, G., Hobson, J. A., & Ondobaka, S. (2017). Active Inference, Curiosity and Insight. *Neural Computation*, *29*(10), 2633–2683.

Friston, K. J., Parr, T., & de Vries, B. (2017). The graphical brain: Belief propagation and active inference. *Network Neuroscience (Cambridge, Mass.)*, *1*(4), 381–414.

Friston, K. J., Redish, A. D., & Gordon, J. A. (2017). Computational Nosology and Precision Psychiatry. *Computational Psychiatry*, *1*, 2–23.

Friston, K. J., Shiner, T., FitzGerald, T., Galea, J. M., Adams, R., Brown, H., Dolan, R. J., Moran, R., Stephan, K. E., & Bestmann, S. (2012). Dopamine, affordance and active inference. *PLoS Computational Biology*, *8*(1), e1002327.

Friston, K. J., Stephan, K. E., Montague, R., & Dolan, R. J. (2014). Computational psychiatry: the brain as a phantastic organ. *The Lancet. Psychiatry*, *1*(2), 148–158.

Friston, K., & Kiebel, S. (2009a). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1521), 1211–1221.

Friston, K., & Kiebel, S. (2009b). Cortical circuits for perceptual inference. *Neural Networks: The Official Journal of the International Neural Network Society*, *22*(8), 1093–1104.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology, Paris*, *100*(1-3), 70–87.

Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, *104*(1-2), 137–160.

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., & Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neuroscience*, *6*(4), 187–214.

Friston, K., Schwartenbeck, P., Fitzgerald, T., Moutoussis, M., Behrens, T., & Dolan, R. J. (2013). The anatomy of choice: active inference and agency. *Frontiers in Human Neuroscience*, *7*, 598.

Friston, K., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., & Dolan, R. J. (2014). The anatomy of choice: dopamine and decision-making. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1655). https://doi.org/10.1098/rstb.2013.0481

Fuller, B. F. (1992). The effects of stress-anxiety and coping styles on heart rate variability. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *12*(1), 81–86.

Gagliano, M., Vyazovskiy, V. V., Borbély, A. A., Grimonprez, M., & Depczynski, M. (2016). Learning by Association in Plants. *Scientific Reports*, *6*, 38427.

Gagne, C., Dayan, P., & Bishop, S. J. (2018). When planning to survive goes wrong: predicting the future and replaying the past in anxiety and PTSD. *Current Opinion in Behavioral Sciences*, *24*, 89–95.

Gallistel, C. R., Stellar, J. R., & Bubis, E. (1974). Parametric analysis of brain stimulation reward in the rat: I. The transient process and the memory-containing process. *Journal of Comparative and Physiological Psychology*, *87*(5), 848–859.

Garcés, P., López-Sanz, D., Maestú, F., & Pereda, E. (2017). Choice of Magnetometers and Gradiometers after Signal Space Separation. *Sensors* , *17*(12). https://doi.org/10.3390/s17122926

Gardner, M. P. H., Schoenbaum, G., & Gershman, S. J. (2018). Rethinking dopamine as generalized prediction error. *Proceedings. Biological Sciences / The Royal Society*, *285*(1891). https://doi.org/10.1098/rspb.2018.1645

Garfinkel, S. N., Seth, A. K., Barrett, A. B., Suzuki, K., & Critchley, H. D. (2015). Knowing your own heart: distinguishing interoceptive accuracy from interoceptive awareness. *Biological Psychology*, *104*, 65–74.

Garrido, M. I., Friston, K. J., Kiebel, S. J., Stephan, K. E., Baldeweg, T., & Kilner, J. M. (2008). The functional anatomy of the MMN: a DCM study of the roving paradigm. *NeuroImage*, *42*(2), 936–944.

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: a review of underlying mechanisms. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *120*(3), 453–463.

Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, *37*(7), 1297–1310.

Gega, L. (2017). The virtues of virtual reality in exposure therapy. *The British Journal of Psychiatry: The Journal of Mental Science*, *210*(4), 245–246.

Gehring, W. J., & Willoughby, A. R. (2004). Are all medial frontal negativities created equal? Toward a richer empirical basis for theories of action monitoring. *Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring*, *14*, 20.

Gené-Cos, N., Ring, H. A., Pottinger, R. C., & Barrett, G. (1999). Possible roles for mismatch negativity in neuropsychiatry. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *12*(1), 17–27.

Gershman, S., Cohen, J., Niv - Proceedings of the Cognitive, Y., & 2010. (2010). Learning to selectively attend. *Escholarship.org*. http://escholarship.org/uc/item/7bn072ds.pdf

Gershman, S. J. (2017). Dopamine, Inference, and Uncertainty. *Neural Computation*, *29*(12), 3311–3326.

Ge, Y., Wu, J., Sun, X., & Zhang, K. (2011). Enhanced mismatch negativity in adolescents with posttraumatic stress disorder (PTSD). *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *79*(2), 231–235.

Ghosh, S., Laxmi, T. R., & Chattarji, S. (2013). Functional connectivity from the amygdala to the hippocampus grows stronger after stress. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(17), 7234–7244.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review*, *103*(4), 650–669.

Gilbert, C. D. (1983). Microcircuitry of the visual cortex. *Annual Review of Neuroscience*, *6*, 217–247.

Gilbert, D. T., & Wilson, T. D. (2009). Why the brain talks to itself: sources of error in emotional prediction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1521), 1335–1341.

Gilovich, T., Griffin, D., Kahneman, D., & Cambridge University Press. (2002). *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge University Press.

Gil, Z., Connors, B. W., & Amitai, Y. (1997). Differential regulation of neocortical synapses by neuromodulators and activity. *Neuron*, *19*(3), 679–686.

Giorgetta, C., Grecucci, A., Zuanon, S., Perini, L., Balestrieri, M., Bonini, N., Sanfey, A. G., & Brambilla, P. (2012). Reduced risk-taking behavior as a trait feature of anxiety. *Emotion* , *12*(6), 1373–1383.

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595.

Glaser, J. I., Perich, M. G., Ramkumar, P., Miller, L. E., & Kording, K. P. (2018). Population coding of conditional probability distributions in dorsal premotor cortex. *Nature Communications*, *9*(1), 1788.

Glautier, S. (2013). Revisiting the learning curve (once again). *Frontiers in Psychology*, *4*, 982.

Glimcher, P. W. (2004). *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics*. MIT Press.

Glimcher, P. W., & Fehr, E. (2013). *Neuroeconomics: Decision Making and the Brain*. Academic Press.

Gnadt, J. W., & Andersen, R. A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *70*(1), 216–220.

Goldberger, A. L., Peng, C.-K., & Lipsitz, L. A. (2002). What is physiologic complexity and how does it change with aging and disease? *Neurobiology of Aging*, *23*(1), 23–26.

Gonzalez-Burgos, G., & Lewis, D. A. (2008). GABA neurons and the mechanisms of network oscillations: implications for understanding cortical dysfunction in schizophrenia. *Schizophrenia Bulletin*, *34*(5), 944–961.

Gorman, J. M., & Sloan, R. P. (2000). Heart rate variability in depressive and anxiety disorders. *American Heart Journal*, *140*(4 Suppl), 77–83.

Gormezano, I., & Moore, J. W. (1966). Classical conditioning. *Experimental Methods and Instrumentation in Psychology*, *1*, 385–420.

Gottlieb, J., Oudeyer, P.-Y., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends in Cognitive Sciences*, *17*(11), 585–593.

Gould, I. C., Rushworth, M. F., & Nobre, A. C. (2011). Indexing the graded allocation of visuospatial attention using anticipatory alpha oscillations. *Journal of Neurophysiology*, *105*(3), 1318–1326.

Grace, A. A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience*, *41*(1), 1–24.

Gray, J. A., & McNaughton, N. (2000). *The Neuropsychology of anxiety: an enquiry into the functions of the septo-hippocampal system*. Oxford: Oxford University Press.

Greenberg, P. E., Fournier, A.-A., Sisitsky, T., Pike, C. T., & Kessler, R. C. (2015). The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *The Journal of Clinical Psychiatry*, *76*(2), 155–162.

Green, L., & Myerson, J. (2004). A Discounting Framework for Choice With Delayed and Probabilistic Rewards. *Psychological Bulletin*, *130*(5), 769–792.

Gregory, R. L. (1980). Perceptions as hypotheses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *290*(1038), 181–197.

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: exploring representations and inductive biases. *Trends in Cognitive Sciences*, *14*(8), 357–364.

Grillon, C. (2008). Models and mechanisms of anxiety: evidence from startle studies. *Psychopharmacology*, *199*(3), 421–437.

Grillon, C., Ameli, R., Woods, S. W., Merikangas, K., & Davis, M. (1991). Fear-potentiated startle in humans: effects of anticipatory anxiety on the acoustic blink reflex. *Psychophysiology*, *28*(5), 588–595.

Grillon, C., & Charney, D. R. (2011). In the face of fear: anxiety sensitizes defensive responses to fearful faces. *Psychophysiology*, *48*(12), 1745–1752.

Grillon, C., Robinson, O. J., Cornwell, B., & Ernst, M. (2019). Modeling anxiety in healthy humans: a key intermediate bridge between basic and clinical sciences. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *44*(12), 1999–2010.

Grillon, C., Robinson, O. J., Krimsky, M., O'Connell, K., Alvarez, G., & Ernst, M. (2017). Anxiety-mediated facilitation of behavioral inhibition: Threat processing and defensive reactivity during a go/no-go task. *Emotion* , *17*(2), 259–266.

Grillon, C., Robinson, O. J., O'Connell, K., Davis, A., Alvarez, G., Pine, D. S., & Ernst, M. (2017). Clinical anxiety promotes excessive response inhibition. *Psychological Medicine*, *47*(3), 484–494.

Grissom, R. J., & Kim, J. J. (2012). *Effect sizes for research: Univariate and multivariate applications*. Routledge.

Gross, C., & Hen, R. (2004). The developmental origins of anxiety. *Nature Reviews. Neuroscience*, *5*(7), 545–552.

Grupe, D. W., & Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: an integrated neurobiological and psychological perspective. *Nature Reviews. Neuroscience*, *14*(7), 488–501.

Güntekin, B., & Başar, E. (2014). A review of brain oscillations in perception of faces and emotional pictures. *Neuropsychologia*, *58*, 33–51.

Gupta, A. S., van der Meer, M. A. A., Touretzky, D. S., & Redish, A. D. (2012). Segmentation of spatial experience by hippocampal theta sequences. *Nature Neuroscience*, *15*(7), 1032–1039.

Gu, Q. (2002). Neuromodulatory transmitter systems in the cortex and their role in cortical plasticity. *Neuroscience*, *111*(4), 815–835.

Gu, R., Wu, R., Broster, L. S., Jiang, Y., Xu, R., Yang, Q., Xu, P., & Luo, Y.-J. (2017). Trait Anxiety and Economic Risk Avoidance Are Not Necessarily Associated: Evidence from the Framing Effect. *Frontiers in Psychology*, *8*, 92.

Gustavson, D. E., Altamirano, L. J., Johnson, D. P., Whisman, M. A., & Miyake, A. (2017). Is set shifting really impaired in trait anxiety? Only when switching away from an effortfully established task set. *Emotion* , *17*(1), 88–101.

Gutiérrez-García, A. G., & Contreras, C. M. (2013). Anxiety: an adaptive emotion. *New Insights into Anxiety Disorders*, 21–37.

Haber, S. N. (2003). The primate basal ganglia: parallel and integrative networks. *Journal of Chemical Neuroanatomy*, *26*(4), 317–330.

Haber, S. N., & Fudge, J. L. (1997). The primate substantia nigra and VTA: integrative circuitry and function. *Critical Reviews in Neurobiology*, *11*(4), 323–342.

Haeusler, S., & Maass, W. (2007). A statistical analysis of information-processing properties of lamina-specific cortical microcircuit models. *Cerebral Cortex* , *17*(1), 149–162.

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*(2), 161–170.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: the feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*(6), 905–912.

HajiHosseini, A., Rodríguez-Fornells, A., & Marco-Pallarés, J. (2012). The role of beta-gamma oscillations in unexpected rewards processing. *NeuroImage*, *60*(3), 1678–1685.

Hamilton, J. P., Farmer, M., Fogelman, P., & Gotlib, I. H. (2015). Depressive Rumination, the Default-Mode Network, and the Dark Matter of Clinical Neuroscience. *Biological Psychiatry*, *78*(4), 224–230.

Hansen, P., Kringelbach, M., & Salmelin, R. (2010). *MEG: An Introduction to Methods*. Oxford University Press.

Happé, F. (1999). Autism: cognitive deficit or cognitive style? *Trends in Cognitive Sciences*, *3*(6), 216–222.

Happé, F. G. (1996). Studying weak central coherence at low levels: children with autism do not succumb to visual illusions. A research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *37*(7), 873–877.

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(22), 5623–5630.

Harlé, K. M., Guo, D., Zhang, S., Paulus, M. P., & Yu, A. J. (2017). Anhedonia and anxiety underlying depressive symptomatology have distinct effects on reward-based decision-making. *PloS One*, *12*(10), e0186473.

Harmer, C. J., & Cowen, P. J. (2013). "It"s the way that you look at it'--a cognitive neuropsychological account of SSRI action in depression. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1615), 20120407.

Harmer, C. J., Goodwin, G. M., & Cowen, P. J. (2009). Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. *The British Journal of Psychiatry: The Journal of Mental Science*, *195*(2), 102–108.

Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, *394*(6695), 780–784.

Harris, K. D., & Shepherd, G. M. G. (2015). The neocortical circuit: themes and variations. *Nature Neuroscience*, *18*(2), 170–181.

Harrison, L. M., Bestmann, S., Rosa, M. J., Penny, W., & Green, G. G. R. (2011). Time scales of representation in the human brain: weighing past information to predict future events. *Frontiers in Human Neuroscience*, *5*, 37.

Hart, A. S., Clark, J. J., & Phillips, P. E. M. (2015). Dynamic shaping of dopamine signals during probabilistic Pavlovian conditioning. *Neurobiology of Learning and Memory*, *117*, 84–92.

Hart, A. S., Rutledge, R. B., Glimcher, P. W., & Phillips, P. E. M. (2014). Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(3), 698–704.

Hartley, C. A., & Phelps, E. A. (2012). Anxiety and decision-making. *Biological Psychiatry*, *72*(2), 113–118.

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(5), 1726–1731.

Hassin, R. R., Aarts, H., Eitam, B., Custers, R., & Kleiman, T. (2009). Non-conscious goal pursuit and the effortful control of behavior. *Oxford Handbook of Human Action*, *442*, 457.

Hassler, U., Barreto, N. T., & Gruber, T. (2011). Induced gamma band responses in human EEG after the control of miniature saccadic artifacts. *NeuroImage*, *57*(4), 1411–1421.

Haumesser, J. K., Beck, M. H., Pellegrini, F., Kühn, J., Neumann, W.-J., Altschüler, J., Harnack, D., Kupsch, A., Nikulin, V. V., Kühn, A. A., & van Riesen, C. (2021). Subthalamic beta oscillations correlate with dopaminergic degeneration in experimental parkinsonism. *Experimental Neurology*, *335*, 113513.

Hauser, T. U., Fiore, V. G., Moutoussis, M., & Dolan, R. J. (2016). Computational Psychiatry of ADHD: Neural Gain Impairments across Marrian Levels of Analysis. *Trends in Neurosciences*, *39*(2), 63–73.

Hauser, T. U., Iannaccone, R., Ball, J., Mathys, C., Brandeis, D., Walitza, S., & Brem, S. (2014). Role of the medial prefrontal cortex in impaired decision making in juvenile attention-deficit/hyperactivity disorder. *JAMA Psychiatry* , *71*(10), 1165–1173.

Hayden, B. Y., Heilbronner, S. R., Pearson, J. M., & Platt, M. L. (2011). Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(11), 4178–4187.

Heeger, D. J. (2017). Theory of cortical function. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(8), 1773–1782.

Heilman, R. M., Crişan, L. G., Houser, D., Miclea, M., & Miu, A. C. (2010). Emotion regulation and decision making under risk and uncertainty. *Emotion* , *10*(2), 257–265.

Hein, T. P., de Fockert, J., & Ruiz, M. H. (2021). State anxiety biases estimates of uncertainty and impairs reward learning in volatile environments. *NeuroImage*, *224*, 117424.

Hein, T. P., & Ruiz, M. H. (2021). State anxiety alters the neural oscillatory correlates of predictions and prediction errors during reward learning. In *bioRxiv* (p. 2021.03.08.434415). https://doi.org/10.1101/2021.03.08.434415

Heinzle, J., Hepp, K., & Martin, K. A. C. (2007). A microcircuit model of the frontal eye fields. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(35), 9341–9353.

Heinzle, J., Koopmans, P. J., den Ouden, H. E. M., Raman, S., & Stephan, K. E. (2016). A hemodynamic model for layered BOLD signals. *NeuroImage*, *125*, 556–570.

Hellhammer, D. H., Wüst, S., & Kudielka, B. M. (2009). Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology*, *34*(2), 163–171.

Helmholtz, H. von. (1866). Concerning the perceptions in general. *Treatise on Physiological Optics,*.

Henco, L., Brandi, M.-L., Lahnakoski, J. M., Diaconescu, A. O., Mathys, C., & Schilbach, L. (2020). Bayesian modelling captures inter-individual differences in social belief computations in the putamen and insula. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*. https://doi.org/10.1016/j.cortex.2020.02.024

Henry, C., Sorbara, F., Lacoste, J., Gindre, C., & Leboyer, M. (2001). Antidepressant-induced mania in bipolar patients: identification of risk factors. *The Journal of Clinical Psychiatry*, *62*(4), 249–255.

Herry, C., & Johansen, J. P. (2014). Encoding of fear learning and memory in distributed neuronal circuits. *Nature Neuroscience*, *17*(12), 1644–1654.

Hilbert, M. (2012). Toward a synthesis of cognitive biases: how noisy information processing can bias human decision making. *Psychological Bulletin*, *138*(2), 211–237.

Hilker, M., Schwachtje, J., Baier, M., Balazadeh, S., Bäurle, I., Geiselhardt, S., Hincha, D. K., Kunze, R., Mueller-Roeber, B., Rillig, M. C., Rolff, J., Romeis, T., Schmülling, T., Steppuhn, A., van Dongen, J., Whitcomb, S. J., Wurst, S., Zuther, E., & Kopka, J. (2016). Priming and memory of stress responses in organisms lacking a nervous system. *Biological Reviews of the Cambridge Philosophical Society*, *91*(4), 1118–1133.

Hill, C. A. (2008). Rationality of Preference Construction (and the Irrationality of Rational Choice). *Minn. JL Sci. & Tech.*, *9*, 689.

Hinton, G. E., & Zemel, R. S. (1994). Autoencoders, Minimum Description Length and Helmholtz Free Energy. In J. D. Cowan, G. Tesauro, & J. Alspector (Eds.), *Advances in Neural Information Processing Systems 6* (pp. 3–10). Morgan-Kaufmann.

Hirsch, C. R., & Mathews, A. (2012). A cognitive model of pathological worry. *Behaviour Research and Therapy*, *50*(10), 636–646.

Hohwy, J. (2013). *The Predictive Mind*. Oxford University Press.

Hohwy, J. (2016). The self-evidencing brain. *Noûs*, *50*(2), 259–285.

Holland, P. C., & Gallagher, M. (1999). Amygdala circuitry in attentional and representational processes. *Trends in Cognitive Sciences*, *3*(2), 65–73.

Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*, *1*(4), 304–309.

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709.

Holroyd, C. B., & Coles, M. G. H. (2008). Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *44*(5), 548–559.

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, *44*(6), 913–917.

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport*, *14*(18), 2481–2484.

Homan, P., Levy, I., Feltham, E., Gordon, C., Hu, J., Li, J., Pietrzak, R. H., Southwick, S., Krystal, J. H., Harpaz-Rotem, I., & Schiller, D. (2019). Neural computations of threat in the aftermath of combat trauma. *Nature Neuroscience*, *22*(3), 470–476.

Hoogenboom, N., Schoffelen, J.-M., Oostenveld, R., Parkes, L. M., & Fries, P. (2006). Localizing human visual gamma-band activity in frequency, time and space. *NeuroImage*, *29*(3), 764–773.

Horvitz, J. C. (2002). Dopamine gating of glutamatergic sensorimotor and incentive motivational input signals to the striatum. *Behavioural Brain Research*, *137*(1-2), 65–74.

Hosokawa, T., Kennerley, S. W., Sloan, J., & Wallis, J. D. (2013). Single-neuron mechanisms underlying cost-benefit analysis in frontal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(44), 17385–17397.

Howard, J. D., Gottfried, J. A., Tobler, P. N., & Kahnt, T. (2015). Identity-specific coding of future rewards in the human orbitofrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(16), 5195–5200.

Howlett, J. R., Thompson, W. K., & Paulus, M. P. (2019). Computational Evidence for Underweighting of Current Error and Overestimation of Future Error in Anxious Individuals. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. https://doi.org/10.1016/j.bpsc.2019.12.011

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making [Review of *Neural systems responding to degrees of uncertainty in human decision-making*]. *Science*, *310*(5754), 1680–1683. science.sciencemag.org.

Huang, H., Thompson, W., & Paulus, M. P. (2017). Computational Dysfunctions in Anxiety: Failure to Differentiate Signal From Noise. *Biological Psychiatry*, *82*(6), 440–446.

Huang, Y., & Rao, R. P. N. (2011). Predictive coding. *Wiley Interdisciplinary Reviews. Cognitive Science*. https://onlinelibrary.wiley.com/doi/abs/10.1002/wcs.142

Huettel, S. A., Song, A. W., & McCarthy, G. (2005). Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *25*(13), 3304–3311.

Hull, C. L. (1943). *Principles of behavior* (Vol. 422). Appleton-century-crofts New York.

Hunt, L. T., Dolan, R. J., & Behrens, T. E. J. (2014). Hierarchical competitions subserving multi-attribute choice. *Nature Neuroscience*, *17*(11), 1613–1622.

Hunt, L. T., & Hayden, B. Y. (2017). A distributed, hierarchical and recurrent framework for reward-based choice. *Nature Reviews. Neuroscience*, *18*(3), 172–182.

Hunt, L. T., Kolling, N., Soltani, A., Woolrich, M. W., Rushworth, M. F. S., & Behrens, T. E. J. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nature Neuroscience*, *15*(3), 470–476, S1–S3.

Hunt, L. T., Nishantha Malalasekera, W. M., de Berker, A. O., Miranda, B., Farmer, S. F., Behrens, T. E. J., & Kennerley, S. W. (2017). Triple dissociation of attention and decision computations across prefrontal cortex. In *bioRxiv* (p. 171173). https://doi.org/10.1101/171173

Huys, Q. J. M., Browning, M., Paulus, M. P., & Frank, M. J. (2021). Advances in the computational understanding of mental illness. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *46*(1), 3–19.

Huys, Q. J. M., Daw, N. D., & Dayan, P. (2015). Depression: a decision-theoretic analysis. *Annual Review of Neuroscience*, *38*, 1–23.

Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, *19*(3), 404–413.

Huys, Q. J. M., Moutoussis, M., & Williams, J. (2011). Are computational models of any use to psychiatry? *Neural Networks: The Official Journal of the International Neural Network Society*, *24*(6), 544–551.

Iglesias, S., Kasper, L., Harrison, S. J., Manka, R., Mathys, C., & Stephan, K. E. (2021). Cholinergic and dopaminergic effects on prediction error and uncertainty responses during sensory associative learning. *NeuroImage*, *226*, 117590.

Iglesias, S., Mathys, C., Brodersen, K. H., Kasper, L., Piccirelli, M., den Ouden, H. E. M., & Stephan, K. E. (2013). Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron*, *80*(2), 519–530.

Iglesias, S., Mathys, C., Brodersen, K. H., Kasper, L., Piccirelli, M., den Ouden, H. E. M., & Stephan, K. E. (2019). Editorial Note to: Hierarchical Prediction Errors in Midbrain and Basal Forebrain during Sensory Learning [Review of *Editorial Note to: Hierarchical Prediction Errors in Midbrain and Basal Forebrain during Sensory Learning*]. *Neuron*, *101*(6), 1195. narcis.nl.

Iglesias, S., Tomiello, S., Schneebeli, M., & Stephan, K. E. (2017). Models of neuromodulation for computational psychiatry. *Wiley Interdisciplinary Reviews. Cognitive Science*, *8*(3). https://doi.org/10.1002/wcs.1420

Iigaya, K. (2016). Adaptive learning and decision-making under uncertainty by metaplastic synapses guided by a surprise detection system. *eLife*, *5*. https://doi.org/10.7554/eLife.18073

Ikemoto, S., & Panksepp, J. (1999). The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Research. Brain Research Reviews*, *31*(1), 6–41.

Indovina, I., Robbins, T. W., Núñez-Elizalde, A. O., Dunn, B. D., & Bishop, S. J. (2011). Fear-conditioning mechanisms associated with trait vulnerability to anxiety in humans. *Neuron*, *69*(3), 563–571.

Inglis, J. B., Valentin, V. V., & Ashby, F. G. (2020). Modulation of Dopamine for Adaptive Learning: a Neurocomputational Model. *Computational Brain & Behavior*. https://doi.org/10.1007/s42113-020-00083-x

Ionescu, D. F., Niciu, M. J., Mathews, D. C., Richards, E. M., & Zarate, C. A., Jr. (2013). Neurobiology of anxious depression: a review. *Depression and Anxiety*, *30*(4), 374–385.

Ishikawa, S.-I., Okajima, I., Matsuoka, H., & Sakano, Y. (2007). Cognitive behavioural therapy for anxiety disorders in children and adolescents: A meta-analysis. *Child and Adolescent Mental Health*, *12*(4), 164–172.

Ivanov, I., Liu, X., Clerkin, S., Schulz, K., Friston, K., Newcorn, J. H., & Fan, J. (2012). Effects of motivation on reward and attentional networks: an fMRI study. *Brain and Behavior*, *2*(6), 741–753.

Izquierdo, A. (2017). Functional Heterogeneity within Rat Orbitofrontal Cortex in Reward Learning and Decision Making. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(44), 10529–10540.

Jahn, C. I., Gilardeau, S., Varazzani, C., Blain, B., Sallet, J., Walton, M. E., & Bouret, S. (2018). Correction to: Dual contributions of noradrenaline to behavioural flexibility and motivation. *Psychopharmacology*, *235*(10), 3081.

James, A. C., Reardon, T., Soler, A., James, G., & Creswell, C. (2020). Cognitive behavioural therapy for anxiety disorders in children and adolescents. *Cochrane Database of Systematic Reviews* , *11*, CD013162.

James, W. (1922). The emotions. In C. G. Lange (Ed.), *The emotions, Vol* (Vol. 1, pp. 93–135). Williams & Wilkins Co.

Jankovic, J., & Tolosa, E. (2007). *Parkinson's Disease and Movement Disorders*. Lippincott Williams & Wilkins.

JASP Team. (2019). *JASP (Version 0.9.2)[Computer software].https://jasp-stats.org/*.

Jávor-Duray, B. N., Vinck, M., van der Roest, M., Bezard, E., Berendse, H. W., Boraud, T., & Voorn, P. (2017). Alterations in Functional Cortical Hierarchy in Hemiparkinsonian Rats. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(32), 7669–7681.

Jaynes, E. T. (1957). Information Theory and Statistical Mechanics. *Physics Review*, *106*(4), 620–630.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press.

Jenkinson, N., & Brown, P. (2011). New insights into the relationship between dopamine, beta oscillations and motor function. *Trends in Neurosciences*, *34*(12), 611–618.

Jensen, O., Bonnefond, M., & VanRullen, R. (2012). An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends in Cognitive Sciences*, *16*(4), 200–206.

Jensen, O., & Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends in Cognitive Sciences*, *11*(7), 267–269.

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, *4*, 186.

Jepma, M., Murphy, P. R., Nassar, M. R., Rangel-Gomez, M., Meeter, M., & Nieuwenhuis, S. (2016). Catecholaminergic Regulation of Learning Rate in a Dynamic Environment. *PLoS Computational Biology*, *12*(10), e1005171.

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off: evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, *23*(7), 1587–1596.

Jiang, D., Zhang, D., Chen, Y., He, Z., Gao, Q., Gu, R., & Xu, P. (2018). Trait anxiety and probabilistic learning: Behavioral and electrophysiological findings. *Biological Psychology*, *132*, 17–26.

Jocham, G., Neumann, J., Klein, T. A., Danielmeier, C., & Ullsperger, M. (2009). Adaptive coding of action values in the human rostral cingulate zone. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(23), 7489–7496.

Joffily, M., & Coricelli, G. (2013). Emotional valence and the free-energy principle. *PLoS Computational Biology*, *9*(6), e1003094.

Johnson, A., & Redish, A. D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(45), 12176–12189.

Johnson, A., van der Meer, M. A. A., & Redish, A. D. (2007). Integrating hippocampus and striatum in decision-making. *Current Opinion in Neurobiology*, *17*(6), 692–697.

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, *89*(1), 221–234.

Jo, S., & Jung, M. W. (2016). Differential coding of uncertain reward in rat insular and orbitofrontal cortex. *Scientific Reports*, *6*, 24085.

Justin Kim, M., & Whalen, P. J. (2009). The Structural Integrity of an Amygdala–Prefrontal Pathway Predicts Trait Anxiety. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(37), 11614–11618.

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

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

Kael White, J., & Monosov, I. E. (2016). Neurons in the primate dorsal striatum signal the uncertainty of object–reward associations. *Nature Communications*, *7*(1), 1–8.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the Assumptions of Economics. *The Journal of Business*, *59*(4), S285–S300.

Kahneman, D., Slovic, S. P., Slovic, P., & Tversky, A. (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press.

Kahneman, D., & Tversky, A. (2013). Choices, values, and frames. *Handbook of the Fundamentals of Financial*. https://www.worldscientific.com/doi/abs/10.1142/9789814417358_0016

Kalisch, R., Wiech, K., Critchley, H. D., & Dolan, R. J. (2006). Levels of appraisal: a medial prefrontal role in high-level appraisal of emotional material. *NeuroImage*, *30*(4), 1458–1466.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, *82*(1), 35–45.

Kamin, L. J. (1967). Predictability, surprise, attention, and conditioning. In SYMP. ON PUNISHMENT.

Kanai, R., Komura, Y., Shipp, S., & Friston, K. (2015). Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1668). https://doi.org/10.1098/rstb.2014.0169

Kanuri, N., Taylor, C. B., Cohen, J. M., & Newman, M. G. (2015). Classification models for subthreshold generalized anxiety disorder in a college population: Implications for prevention. *Journal of Anxiety Disorders*, *34*, 43–52.

Kappes, A., Harvey, A. H., Lohrenz, T., Montague, P. R., & Sharot, T. (2020). Confirmation bias in the utilization of others' opinion strength. *Nature Neuroscience*, *23*(1), 130–137.

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773–795.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience*, *23*, 315–341.

Katthagen, T., Mathys, C., Deserno, L., Walter, H., Kathmann, N., Heinz, A., & Schlagenhauf, F. (2018). Modeling subjective relevance in schizophrenia and its relation to aberrant salience. *PLoS Computational Biology*, *14*(8), e1006319.

Kawachi, I., Sparrow, D., Vokonas, P. S., & Weiss, S. T. (1995). Decreased heart rate variability in men with phobic anxiety (data from the Normative Aging Study). *The American Journal of Cardiology*, *75*(14), 882–885.

Keller, G. B., & Mrsic-Flogel, T. D. (2018). Predictive Processing: A Canonical Cortical Computation. *Neuron*, *100*(2), 424–435.

Kennerley, S. W., Behrens, T. E. J., & Wallis, J. D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature Neuroscience*, *14*(12), 1581–1589.

Kennerley, S. W., Dahmubed, A. F., Lara, A. H., & Wallis, J. D. (2009). Neurons in the frontal lobe encode the value of multiple decision variables. *Journal of Cognitive Neuroscience*, *21*(6), 1162–1178.

Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, *9*(7), 940–947.

Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, *455*(7210), 227–231.

Keren, A. S., Yuval-Greenberg, S., & Deouell, L. Y. (2010). Saccadic spike potentials in gamma-band EEG: characterization, detection and suppression. *NeuroImage*, *49*(3), 2248–2263.

Kertz, S. J., McHugh, R. K., Lee, J., & Björgvinsson, T. (2014). Examining the latent structure of worry and generalized anxiety in a clinical sample. *Journal of Anxiety Disorders*, *28*(1), 8–15.

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Chatterji, S., Lee, S., Ormel, J., Üstün, T. B., & Wang, P. S. (2009). The global burden of mental disorders: an update from the WHO World Mental Health (WMH) surveys. *Epidemiologia E Psichiatria Sociale*, *18*(1), 23.

Kessler, R. C., Chiu, W. T., Demler, O., Merikangas, K. R., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 617–627.

Kiebel, S. J., & Friston, K. J. (2004a). Statistical parametric mapping for event-related potentials: I. Generic considerations. *NeuroImage*, *22*(2), 492–502.

Kiebel, S. J., & Friston, K. J. (2004b). Statistical parametric mapping for event-related potentials (II): a hierarchical temporal model. *NeuroImage*, *22*(2), 503–520.

Kiebel, S. J., Tallon-Baudry, C., & Friston, K. J. (2005). Parametric analysis of oscillatory activity as measured with EEG/MEG. *Human Brain Mapping*, *26*(3), 170–177.

Kieffer, B. L. (1999). Opioids: first lessons from knockout mice. *Trends in Pharmacological Sciences*, *20*(1), 19–26.

Kieffer, B. L., & Gavériaux-Ruff, C. (2002). Exploring the opioid system by gene knockout. *Progress in Neurobiology*, *66*(5), 285–306.

Kilner, J. M., & Friston, K. J. (2010). Topological inference for EEG and MEG. *The Annals of Applied Statistics*, *4*(3), 1272–1290.

Kilner, J. M., Kiebel, S. J., & Friston, K. J. (2005). Applications of random field theory to electrophysiology. *Neuroscience Letters*, *374*(3), 174–178.

Kim, A. J., Lee, D. S., & Anderson, B. A. (2020). The influence of threat on the efficiency of goal-directed attentional control. *Psychological Research*. https://doi.org/10.1007/s00426-020-01321-4

Kimble, D. P., & Kimble, R. J. (1970). The effect of hippocampal lesions on extinction and "hypothesis" behavior in rats. *Physiology & Behavior*, *5*(7), 735–738.

Kim, J. N., & Shadlen, M. N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience*, *2*(2), 176–185.

Kim, M. J., Shin, J., Taylor, J. M., Mattek, A. M., Chavez, S. J., & Whalen, P. J. (2017). Intolerance of uncertainty predicts increased striatal volume. *Emotion* , *17*(6), 895–899.

Kim, M., Kim, S., Lee, K.-U., & Jeong, B. (2020). Pessimistically biased perception in panic disorder during risk learning. *Depression and Anxiety*, *37*(7), 609–619.

Kircanski, K., Thompson, R. J., Sorenson, J., Sherdell, L., & Gotlib, I. H. (2015). Rumination and Worry in Daily Life: Examining the Naturalistic Validity of Theoretical Constructs. *Clinical Psychological Science*, *3*(6), 926–939.

Kirschbaum, C., Pirke, K.-M., & Hellhammer, D. H. (1993). The "Trier Social Stress Test"--a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*(1-2), 76–81.

Kitazawa, S., Kimura, T., & Yin, P. B. (1998). Cerebellar complex spikes encode both destinations and errors in arm movements. *Nature*, *392*(6675), 494–497.

Klein, E., Cnaani, E., Harel, T., Braun, S., & Ben-Haim, S. A. (1995). Altered heart rate variability in panic disorder patients. *Biological Psychiatry*, *37*(1), 18–24.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712–719.

Knyazev, G. G., Savostyanov, A. N., & Levin, E. A. (2004). Alpha oscillations as a correlate of trait anxiety. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *53*(2), 147–160.

Knyazev, G. G., & Slobodskaya, H. R. (2003). Personality trait of behavioral inhibition is associated with oscillatory systems reciprocal relationships. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *48*(3), 247–261.

Knyazev, G. G., Slobodskaya, H. R., & Wilson, G. D. (2002). Psychophysiological correlates of behavioural inhibition and activation. *Personality and Individual Differences*, *33*(4), 647–660.

Koblinger, Á., Fiser, J., & Lengyel, M. (2021). Representations of uncertainty: where art thou? *Current Opinion in Behavioral Sciences*, *38*, 150–162.

Kok, P., & de Lange, F. P. (2015). Predictive Coding in Sensory Cortex. In B. U. Forstmann & E.-J. Wagenmakers (Eds.), *An Introduction to Model-Based Cognitive Neuroscience* (pp. 221–244). Springer New York.

Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C., & de Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex* , *22*(9), 2197–2206.

Koller, W. C. (1992). *Handbook of Parkinson's disease* (Vol. 13). Marcel Dekker Incorporated.

Kolling, N., Behrens, T., Wittmann, M. K., & Rushworth, M. (2016). Multiple signals in anterior cingulate cortex. *Current Opinion in Neurobiology*, *37*, 36–43.

Kolling, N., Wittmann, M. K., Behrens, T. E. J., Boorman, E. D., Mars, R. B., & Rushworth, M. F. S. (2016). Value, search, persistence and model updating in anterior cingulate cortex. *Nature Neuroscience*, *19*(10), 1280–1285.

Kolossa, A., Kopp, B., & Fingscheidt, T. (2015). A computational analysis of the neural bases of Bayesian inference. *NeuroImage*, *106*, 222–237.

Koob, G. F., & Volkow, N. D. (2010). Neurocircuitry of addiction. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *35*(1), 217–238.

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology. General*, *139*(4), 665–682.

Körding, K. P. (2007). Decision Theory: What“ Should” the Nervous System Do? *Science*, *318*(5850), 606–610.

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

Koss, M. C. (1986). Pupillary dilation as an index of central nervous system alpha 2-adrenoceptor activation. *Journal of Pharmacological Methods*, *15*(1), 1–19.

Krain, A. L., Wilson, A. M., Arbuckle, R., Castellanos, F. X., & Milham, M. P. (2006). Distinct neural mechanisms of risk and ambiguity: a meta-analysis of decision-making. *NeuroImage*, *32*(1), 477–484.

Krishnamurthy, K., Nassar, M. R., Sarode, S., & Gold, J. I. (2017). Arousal-related adjustments of perceptual biases optimize perception in dynamic environments. *Nature Human Behaviour*, *1*(6), 0107.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 25* (pp. 1097–1105). Curran Associates, Inc.

Krolak-Salmon, P., Hénaff, M.-A., Vighetto, A., Bertrand, O., & Mauguière, F. (2004). Early amygdala reaction to fear spreading in occipital, temporal, and frontal cortex: a depth electrode ERP study in human. *Neuron*, *42*(4), 665–676.

Krugel, L. K., Biele, G., Mohr, P. N. C., Li, S.-C., & Heekeren, H. R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(42), 17951–17956.

Krypotos, A.-M., Effting, M., Kindt, M., & Beckers, T. (2015). Avoidance learning: a review of theoretical models and recent developments. *Frontiers in Behavioral Neuroscience*, *9*, 189.

Kudielka, B. M., Schommer, N. C., Hellhammer, D. H., & Kirschbaum, C. (2004). Acute HPA axis responses, heart rate, and mood changes to psychosocial stress (TSST) in humans at different times of day. *Psychoneuroendocrinology*, *29*(8), 983–992.

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, *139*(4), 917–940.

Kurniawan, I. T., Guitart-Masip, M., & Dolan, R. J. (2011). Dopamine and effort-based decision making. *Frontiers in Neuroscience*, *5*, 81.

Kurth-Nelson, Z., & Redish, A. D. (2009). Temporal-difference reinforcement learning with distributed representations. *PloS One*, *4*(10), e7362.

Kwisthout, J., & van Rooij, I. (2013). Predictive coding and the Bayesian brain: Intractability hurdles that are yet to be overcome. *CogSci*. http://www.dcc.ru.nl/~johank/material/cogsci13/paper.pdf

Labuschagne, I., Grace, C., Rendell, P., Terrett, G., & Heinrichs, M. (2019). An introductory guide to conducting the Trier Social Stress Test. *Neuroscience and Biobehavioral Reviews*, *107*, 686–695.

Lamba, A., Frank, M. J., & FeldmanHall, O. (2020). Anxiety Impedes Adaptive Social Learning Under Uncertainty. *Psychological Science*, 956797620910993.

Lammel, S., Ion, D. I., Roeper, J., & Malenka, R. C. (2011). Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli. *Neuron*, *70*(5), 855–862.

Lammel, S., Lim, B. K., Ran, C., Huang, K. W., Betley, M. J., Tye, K. M., Deisseroth, K., & Malenka, R. C. (2012). Input-specific control of reward and aversion in the ventral tegmental area. *Nature*, *491*(7423), 212–217.

Land, B. B., Bruchas, M. R., Lemos, J. C., Xu, M., Melief, E. J., & Chavkin, C. (2008). The Dysphoric Component of Stress Is Encoded by Activation of the Dynorphin κ-Opioid

System. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(2), 407–414.

Lang, M., Krátký, J., Shaver, J. H., Jerotijević, D., & Xygalatas, D. (2015). Effects of Anxiety on Spontaneous Ritualized Behavior. *Current Biology: CB*, *25*(14), 1892–1897.

Lang, P. J. (1995). The emotion probe. Studies of motivation and attention. *The American Psychologist*, *50*(5), 372–385.

Lang, P. J., Bradley, M. M., Cuthbert, B. N., & Others. (1997). Motivated attention: Affect, activation, and action. *Attention and Orienting: Sensory and Motivational Processes*, *97*, 135.

Larkum, M. E., Senn, W., & Lüscher, H.-R. (2004). Top-down dendritic input increases the gain of layer 5 pyramidal neurons. *Cerebral Cortex* , *14*(10), 1059–1070.

Larsen, R. S., & Waters, J. (2018). Neuromodulatory Correlates of Pupil Dilation. *Frontiers in Neural Circuits*, *12*, 21.

Latimer, K. W., Yates, J. L., Meister, M. L. R., Huk, A. C., & Pillow, J. W. (2015). NEURONAL MODELING. Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science*, *349*(6244), 184–187.

Lau, B., & Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron*, *58*(3), 451–463.

Laurent, V., Leung, B., Maidment, N., & Balleine, B. W. (2012). μ- and δ-opioid-related processes in the accumbens core and shell differentially mediate the influence of reward-guided and stimulus-guided decisions on choice. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(5), 1875–1883.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology. Human Perception and Performance*, *21*(3), 451–468.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology. General*, *133*(3), 339–354.

Lavín, C., San Martín, R., & Rosales Jubal, E. (2013). Pupil dilation signals uncertainty and surprise in a learning gambling task. *Frontiers in Behavioral Neuroscience*, *7*, 218.

Lavine, M., & Schervish, M. J. (1999). Bayes Factors: What They are and What They are Not. *The American Statistician*, *53*(2), 119–122.

Lavine, N., Reuben, M., & Clarke, P. B. (1997). A population of nicotinic receptors is associated with thalamocortical afferents in the adult rat: laminal and areal analysis. *Journal of Comparative Neurology*, *380*(2), 175-190.

Lavric, A., Rippon, G., & Gray, J. R. (2003). Threat-Evoked Anxiety Disrupts Spatial Working Memory Performance: An Attentional Account. *Cognitive Therapy and Research*, *27*(5), 489–504.

Lawson, R. P., Bisby, J., Nord, C. L., Burgess, N., & Rees, G. (2021). The computational, pharmacological, and physiological determinants of sensory learning under uncertainty. *Current Biology, 31*(1), 163-172.

Lawson, R. P., Mathys, C., & Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. *Nature Neuroscience*, *20*(9), 1293–1299.

Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience*, *8*, 302.

Le Bouc, R., & Pessiglione, M. (2013). Imaging social motivation: distinct brain mechanisms drive effort production during collaboration versus competition. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(40), 15894–15902.

Le Bouc, R., Rigoux, L., Schmidt, L., Degos, B., Welter, M.-L., Vidailhet, M., Daunizeau, J., & Pessiglione, M. (2016). Computational Dissection of Dopamine Motor and Motivational Functions in Humans. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(25), 6623–6633.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

LeDoux, J. E. (2012). Rethinking the emotional brain. *Neuron*, *73*(4), 653–676.

LeDoux, J. E. (2014). Coming to terms with fear. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(8), 2871–2878.

LeDoux, J. E. (2015). *Anxious: The Modern Mind in the Age of Anxiety*. Simon and Schuster.

LeDoux, J. E., & Pine, D. S. (2016). Using Neuroscience to Help Understand Fear and Anxiety: A Two-System Framework. *The American Journal of Psychiatry*, *173*(11), 1083–1093.

Lee, A., Dawes, H., & Chang, E. (2019). Limbic Beta Oscillations Are Evoked by Threat-provoking Cues and Correlate with Subjective Anxiety. *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation*, *12*(2), 442.

Lee, J. H., Whittington, M. A., & Kopell, N. J. (2013). Top-down beta rhythms support selective attention via interlaminar interaction: a model. *PLoS Computational Biology*, *9*(8), e1003164.

Lee, S. W., Shimojo, S., & O'Doherty, J. P. (2014). Neural computations underlying arbitration between model-based and model-free learning. *Neuron*, *81*(3), 687–699.

Lehrer, J. (2010). *How We Decide*. Houghton Mifflin Harcourt.

Le Merrer, J., Becker, J. A. J., Befort, K., & Kieffer, B. L. (2009). Reward processing by the opioid system in the brain. *Physiological Reviews*, *89*(4), 1379–1412.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology*, *66*, 799–823.

Levine, A. S., & Billington, C. J. (2004). Opioids as agents of reward-related feeding: a consideration of the evidence. *Physiology & behavior*, *82*(1), 57-61.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*, *22*(6), 1027–1038.

Liao, H. I., Yoneya, M., Kidani, S., Kashino, M., & Furukawa, S. (2016). Human pupillary dilation response to deviant auditory stimuli: Effects of stimulus properties and voluntary attention. *Frontiers in Neuroscience*, *10*, 43. https://www.frontiersin.org/articles/10.3389/fnins.2016.00043/full

Lieder, F., Daunizeau, J., Garrido, M. I., Friston, K. J., & Stephan, K. E. (2013). Modelling trial-by-trial changes in the mismatch negativity. *PLoS Computational Biology*, *9*(2), e1002911.

Li, J., Schiller, D., Schoenbaum, G., Phelps, E. A., & Daw, N. D. (2011). Differential roles of human striatum and amygdala in associative learning. *Nature Neuroscience*, *14*(10), 1250–1252.

Lin, C.-H., Chuang, S.-C., Kao, D. T., & Kung, C.-Y. (2006). The role of emotions in the endowment effect. *Journal Of Economic Psychology*, *27*(4), 589–597.

Lissek, S. (2012). Toward an account of clinical anxiety predicated on basic, neurally mapped mechanisms of Pavlovian fear-learning: the case for conditioned overgeneralization. *Depression and Anxiety*, *29*(4), 257–263.

Litvak, V., Jha, A., Flandin, G., & Friston, K. (2013). Convolution models for induced electromagnetic responses. *NeuroImage*, *64*, 388–398.

Litvak, V., Mattout, J., Kiebel, S., Phillips, C., Henson, R., Kilner, J., Barnes, G., Oostenveld, R., Daunizeau, J., Flandin, G., Penny, W., & Friston, K. (2011). EEG and MEG data analysis in SPM8. *Computational Intelligence and Neuroscience*, *2011*, 852961.

Liu, M., Dong, W., Qin, S., Verguts, T., & Chen, Q. (2021). Electrophysiological signatures of hierarchical learning. In *bioRxiv* (p. 2021.03.09.434666). https://doi.org/10.1101/2021.03.09.434666

Ljungberg, T., Apicella, P., & Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *Journal of neurophysiology*, *67*(1), 145-163. https://www.physiology.org/doi/abs/10.1152/jn.1992.67.1.145

Llinas, R. R. (2002). *I of the Vortex: From Neurons to Self*. MIT Press.

Lloyd, K., & Dayan, P. (2016). Safety out of control: dopamine and defence. *Behavioral and Brain Functions: BBF*, *12*(1), 15.

Lohani, S., Martig, A. K., Deisseroth, K., Witten, I. B., & Moghaddam, B. (2019). Dopamine Modulation of Prefrontal Cortex Activity Is Manifold and Operates at Multiple Temporal and Spatial Scales. *Cell Reports, 27*(1), 99–114.e6.

Lopes da Silva, F. (2013). EEG and MEG: relevance to neuroscience. *Neuron*, *80*(5), 1112–1128.

Lorian, C. N., & Grisham, J. R. (2011). Clinical implications of risk aversion: an online study of risk-avoidance and treatment utilization in pathological anxiety. *Journal of Anxiety Disorders*, *25*(6), 840–848.

Lübke, J., Egger, V., Sakmann, B., & Feldmeyer, D. (2000). Columnar organization of dendrites and axons of single and synaptically coupled excitatory spiny neurons in layer 4 of the rat barrel cortex. *Journal of Neuroscience*, *20*(14), 5300–5311.

Lucantonio, F., Gardner, M. P. H., Mirenzi, A., Newman, L. E., Takahashi, Y. K., & Schoenbaum, G. (2015). Neural Estimates of Imagined Outcomes in Basolateral Amygdala Depend on Orbitofrontal Cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(50), 16521–16530.

Lumaca, M., Trusbak Haumann, N., Brattico, E., Grube, M., & Vuust, P. (2019). Weighting of neural prediction error by rhythmic complexity: A predictive coding account using mismatch negativity. *European Journal of Neuroscience*, *49*(12), 1597-1609. https://onlinelibrary.wiley.com/doi/abs/10.1111/ejn.14329?casa_token=Rq2XPuvbztUAAAAA:lT1yYMRrF9GP3HulU1hYxFYW1ChZSb2HkgSq9XF1ifSuX7gGrXXnJd3d_cjtTl9yOVmRfPDFsA

Lundqvist, M., Bastos, A. M., & Miller, E. K. (2020). Preservation and Changes in Oscillatory Dynamics across the Cortical Hierarchy. *Journal of Cognitive Neuroscience*, *32*(10), 2024–2035.

Lundqvist, M., Herman, P., Warden, M. R., Brincat, S. L., & Miller, E. K. (2018). Gamma and beta bursts during working memory readout suggest roles in its volitional control. *Nature communications, 9*(1), 1-12. https://www.nature.com/articles/s41467-017-02791-8

Lundqvist, M., Rose, J., Herman, P., Brincat, S. L., Buschman, T. J., & Miller, E. K. (2016). Gamma and Beta Bursts Underlie Working Memory. *Neuron*, *90*(1), 152–164.

MacKay, D. J. C., & Mac, D. J. (2003). *Information Theory, Inference and Learning Algorithms*. Cambridge University Press.

Mackintosh, N. J. (1974). *The psychology of animal learning*. *730*. https://psycnet.apa.org/fulltext/1975-11296-000.pdf

MacLeod, C., & Mathews, A. (1988). Anxiety and the allocation of attention to threat. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *40*(4), 653–670.

Madan, C. R. (2017). Motivated Cognition: Effects of Reward, Emotion, and Other Motivational Factors Across a Variety of Cognitive Domains. *Collabra: Psychology*, *3*(1), 24.

Mager, R., Bullinger, A. H., Mueller-Spahn, F., Kuntze, M. F., & Stoermer, R. (2001). Real-time monitoring of brain activity in patients with specific phobia during exposure therapy, employing a stereoscopic virtual environment. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *4*(4), 465–469.

Maguire, E. A., & Hassabis, D. (2011). Role of the hippocampus in imagination and future thinking [Review of *Role of the hippocampus in imagination and future thinking*]. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(11), E39. National Acad Sciences.

Maheu, M., Dehaene, S., & Meyniel, F. (2019). Brain signatures of a multiscale process of sequence learning in humans. *eLife*, *8*. https://doi.org/10.7554/eLife.41541

Maia, T. V. (2010). Two-factor theory, the actor-critic model, and conditioned avoidance. *Learning & Behavior*, *38*(1), 50–67.

Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature Neuroscience*, *14*(2), 154–162.

Maia, T. V., Huys, Q. J. M., & Frank, M. J. (2017). Theory-Based Computational Psychiatry. *Biological Psychiatry*, *82*(6), 382–384.

Maner, J. K., & Schmidt, N. B. (2006). The role of risk avoidance in anxiety. *Behavior Therapy*, *37*(2), 181–189.

Manning, C., Kilner, J., Neil, L., Karaminis, T., & Pellicano, E. (2017). Children on the autism spectrum update their behaviour in response to a volatile environment. *Developmental Science*, *20*(5). https://doi.org/10.1111/desc.12435

Manucia, G. K., Baumann, D. J., & Cialdini, R. B. (1984). Mood influences on helping: Direct effects or side effects? *Journal of Personality and Social Psychology*, *46*(2), 357–364.

Marco-Pallares, J., Cucurell, D., Cunillera, T., García, R., Andrés-Pueyo, A., Münte, T. F., & Rodríguez-Fornells, A. (2008). Human oscillatory activity associated to reward processing in a gambling task. *Neuropsychologia*, *46*(1), 241–248.

Marin, R. S. (1990). Differential diagnosis and classification of apathy. *The American Journal of Psychiatry*, *147*(1), 22–30.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190.

Marković, D., & Kiebel, S. J. (2016). Comparative Analysis of Behavioral Models for Adaptive Learning in Changing Environments. *Frontiers in Computational Neuroscience*, *10*, 33.

Marr, D. (1969). A theory of cerebellar cortex. *The Journal of Physiology*, *202*(2), 437–470.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. http://papers.cumincad.org/cgi-bin/works/Show?fafa

Marr, D., & Poggio, T. (1976). *From Understanding Computation to Understanding Neural Circuitry*. https://dspace.mit.edu/handle/1721.1/5782?show=full

Marshall, L., Mathys, C., Ruge, D., de Berker, A. O., Dayan, P., Stephan, K. E., & Bestmann, S. (2016). Pharmacological Fingerprints of Contextual Uncertainty. *PLoS Biology*, *14*(11), e1002575.

Marshall, P. J., Reeb, B. C., & Fox, N. A. (2009). Electrophysiological responses to auditory novelty in temperamentally different 9-month-old infants. *Developmental Science*, *12*(4), 568–582.

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., & Bestmann, S. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(47), 12539–12545.

Martin, J. P. (1967). *The basal ganglia and posture*. Lippincott.

Mas-Herrero, E., Ripollés, P., HajiHosseini, A., Rodríguez-Fornells, A., & Marco-Pallarés, J. (2015). Beta oscillations and reward processing: Coupling oscillatory activity and hemodynamic responses. *NeuroImage*, *119*, 13–19.

Massi, B., Donahue, C. H., & Lee, D. (2018). Volatility Facilitates Value Updating in the Prefrontal Cortex. *Neuron*, *99*(3), 598–608.e4.

Mathews, A., & Mackintosh, B. (1998). A Cognitive Model of Selective Processing in Anxiety. *Cognitive Therapy and Research*, *22*(6), 539–560.

Mathews, A., Mackintosh, B., & Fulcher, E. P. (1997). Cognitive biases in anxiety and attention to threat. *Trends in Cognitive Sciences*, *1*(9), 340–345.

Mathews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annual Review of Clinical Psychology*, *1*, 167–195.

MathWorks, I. (2012). MATLAB and statistics toolbox release. *Natick, MA: The MathWorks*.

Mathys, C. (2012). *Hierarchical Gaussian filtering: Construction and variational inversion of a generic Bayesian model of individual learning under uncertainty*. Doctoral thesis. [ETH Zurich]. https://doi.org/10.3929/ethz-a-007595146

Mathys, C., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A bayesian foundation for individual learning under uncertainty. *Frontiers in Human Neuroscience*, *5*, 39.

Mathys, C. D., Lomakina, E. I., Daunizeau, J., Iglesias, S., Brodersen, K. H., Friston, K. J., & Stephan, K. E. (2014). Uncertainty in perception and the Hierarchical Gaussian Filter. *Frontiers in Human Neuroscience*, *8*, 825.

Mathys, C., & Weber, L. (2020). Hierarchical Gaussian Filtering of Sufficient Statistic Time Series for Active Inference. *Active Inference*, 52–58.

Matsumoto, H., Tian, J., Uchida, N., & Watabe-Uchida, M. (2016). Midbrain dopamine neurons signal aversion in a reward-context-dependent manner. *eLife*, *5*. https://doi.org/10.7554/eLife.17328

Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, *459*(7248), 837–841.

Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience*, *10*(5), 647–656.

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

Ma, W. J., & Jazayeri, M. (2014). Neural coding of uncertainty and probability. *Annual Review of Neuroscience*, *37*, 205–220.

Mayer, A., Schwiedrzik, C. M., Wibral, M., Singer, W., & Melloni, L. (2016). Expecting to See a Letter: Alpha Oscillations as Carriers of Top-Down Sensory Predictions. *Cerebral Cortex* , *26*(7), 3146–3160.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, *38*(2), 339–346.

McCorduck, P., & Cfe, C. (2004). *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. CRC Press.

McCoy, A. N., & Platt, M. L. (2005). Risk-sensitive neurons in macaque posterior cingulate cortex. *Nature Neuroscience*, *8*(9), 1220–1227.

McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y., & Schoenbaum, G. (2011). Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(7), 2700–2705.

McGuire, J. T., Nassar, M. R., Gold, J. I., & Kable, J. W. (2014). Functionally dissociable influences on learning rate in a dynamic environment. *Neuron, 84*(4), 870-881.https://www.sciencedirect.com/science/article/pii/S0896627314009118

McGuire, J. T., Nassar, M. R., Gold, J. I., & Kable, J. W. (2014). Functionally dissociable influences on learning rate in a dynamic environment. *Neuron*, *84*(4), 870–881.

Meacham, F., & T Bergstrom, C. (2016). Adaptive behavior can produce maladaptive anxiety due to individual differences in experience. *Evolution, Medicine, and Public Health*, *2016*(1), 270–285.

Meiran, N., Diamond, G. M., Toder, D., & Nemets, B. (2011). Cognitive rigidity in unipolar depression and obsessive compulsive disorder: examination of task switching, Stroop, working memory updating and post-conflict adaptation. *Psychiatry Research*, *185*(1-2), 149–156.

Mendl, M., Burman, O. H. P., & Paul, E. S. (2010). An integrative and functional framework for the study of animal emotion and mood. *Proceedings. Biological Sciences / The Royal Society*, *277*(1696), 2895–2904.

Merolla, P. A., Arthur, J. V., Alvarez-Icaza, R., Cassidy, A. S., Sawada, J., Akopyan, F., Jackson, B. L., Imam, N., Guo, C., Nakamura, Y., Brezzo, B., Vo, I., Esser, S. K., Appuswamy, R., Taba, B., Amir, A., Flickner, M. D., Risk, W. P., Manohar, R., & Modha, D. S. (2014). Artificial brains. A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science*, *345*(6197), 668–673.

Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J.-M., Kennedy, H., & Fries, P. (2016). Alpha-Beta and Gamma Rhythms Subserve Feedback and Feedforward Influences among Human Visual Cortical Areas. *Neuron*, *89*(2), 384–397.

Miller, E. K., Lundqvist, M., & Bastos, A. M. (2018). Working Memory 2.0. *Neuron*, *100*(2), 463–475.

Miller, N. E. (1948). Studies of fear as an acquirable drive fear as motivation and fear-reduction as reinforcement in the learning of new responses. *Journal of Experimental Psychology*, *38*(1), 89–101.

Milne, E., & Scope, A. (2008). Are children with autistic spectrum disorders susceptible to contour illusions? *The British Journal of Developmental Psychology*, *26*(1), 91–102.

Milne, E., Swettenham, J., & Campbell, R. (2005). Motion perception and autistic spectrum disorder: a review. *Current Psychology of Cognition = Cahiers de Psychologie Cognitive: CPC*, *23*(1/2), 3.

Milne, E., Swettenham, J., Hansen, P., Campbell, R., Jeffries, H., & Plaisted, K. (2002). High motion coherence thresholds in children with autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *43*(2), 255–263.

Mineka, S., & Zinbarg, R. (2006). A contemporary learning theory perspective on the etiology of anxiety disorders: it's not what you thought it was. *The American Psychologist*, *61*(1), 10–26.

Minsky, M. (1988). *Society Of Mind*. Simon and Schuster.

Mishkin, M., Malamut, B., & Bachevalier, J. (1984). *Neurobiology of human learning and memory*.

Mitte, K. (2007). Anxiety and risky decision-making: The role of subjective probability and subjective costs of negative events. *Personality and Individual Differences*, *43*(2), 243–253.

Miu, A. C., Heilman, R. M., & Houser, D. (2008). Anxiety impairs decision-making: psychophysiological evidence from an Iowa Gambling Task. *Biological Psychology*, *77*(3), 353–358.

Miu, A. C., Heilman, R. M., & Miclea, M. (2009). Reduced heart rate variability and vagal tone in anxiety: trait versus state, and the effects of autogenic training. *Autonomic Neuroscience: Basic & Clinical*, *145*(1-2), 99–103.

Miu, A. C., Miclea, M., & Houser, D. (2008). Anxiety and decision-making: toward a neuroeconomics perspective. *Advances in Health Economics and Health Services Research*, *20*, 55–84.

Miyazaki, K., Miyazaki, K. W., & Doya, K. (2012). The role of serotonin in the regulation of patience and impulsivity. *Molecular Neurobiology*, *45*(2), 213–224.

Mkrtchian, A., Aylward, J., Dayan, P., Roiser, J. P., & Robinson, O. J. (2017). Modeling Avoidance in Mood and Anxiety Disorders Using Reinforcement Learning. *Biological Psychiatry*, *82*(7), 532–539.

Mkrtchian, A., Roiser, J. P., & Robinson, O. J. (2017). Threat of shock and aversive inhibition: Induced anxiety modulates Pavlovian-instrumental interactions. *Journal of Experimental Psychology. General*, *146*(12), 1694–1704.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A.,

Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533.

Mobbs, D., Wise, T., Suthana, N., Guzmán, N., Kriegeskorte, N., & Leibo, J. Z. (2021). Promises and challenges of human computational ethology. *Neuron*, *109*(14), 2224–2238.

Mobini, S., Body, S., Ho, M.-Y., Bradshaw, C. M., Szabadi, E., Deakin, J. F. W., & Anderson, I. M. (2002). Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology*, *160*(3), 290–298.

Moens, V., & Zénon, A. (2019). Learning and forgetting using reinforced Bayesian change detection. *PLoS Computational Biology*, *15*(4), e1006713.

Mogenson, G. J., Jones, D. L., & Yim, C. Y. (1980). From motivation to action: functional interface between the limbic system and the motor system. *Progress in Neurobiology*, *14*(2-3), 69–97.

Monosov, I. E. (2017). Anterior cingulate is a source of valence-specific information about value and uncertainty. *Nature Communications*, *8*(1), 134.

Monroe, S. M., & Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. *Psychological Bulletin*, *110*(3), 406–425.

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

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *16*(5), 1936–1947.

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, *16*(1), 72–80.

Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, *431*(7010), 760–767.

Montague, R. (2006). *Why choose this book?: how we make decisions*. EP Dutton.

Moody, G. B., & Mark, R. G. (1982). Development and evaluation of a 2-lead ECG analysis program. *Computers in Cardiology*, *9*, 39–44.

Moran, R. J., Campo, P., Symmonds, M., Stephan, K. E., Dolan, R. J., & Friston, K. J. (2013). Free energy, precision and learning: the role of cholinergic neuromodulation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(19), 8227–8236.

Morgan, C. A., 3rd, & Grillon, C. (1999). Abnormal mismatch negativity in women with sexual assault-related posttraumatic stress disorder. *Biological Psychiatry*, *45*(7), 827–832.

Morris, G., Arkadir, D., Nevet, A., Vaadia, E., & Bergman, H. (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron*, *43*(1), 133–143.

Morris, G., Nevet, A., Arkadir, D., Vaadia, E., & Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. *Nature Neuroscience*, *9*(8), 1057–1063.

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., & Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature*, *383*(6603), 812–815.

Morrison, S. E., Saez, A., Lau, B., & Salzman, C. D. (2011). Different time courses for learning-related changes in amygdala and orbitofrontal cortex. *Neuron*, *71*(6), 1127–1140.

Morville, T., Friston, K., Burdakov, D., & Siebner, H. R. (2018). The homeostatic logic of reward. *bioRxiv*. https://www.biorxiv.org/content/10.1101/242974v1.abstract

Moscovitch, D. A. (2009). What Is the Core Fear in Social Phobia? A New Model to Facilitate Individualized Case Conceptualization and Treatment. *Cognitive and Behavioral Practice*, *16*(2), 123–134.

Moustafa, A. A., & Gluck, M. A. (2011). Computational cognitive models of prefrontal-striatal-hippocampal interactions in Parkinson's disease and schizophrenia. *Neural Networks: The Official Journal of the International Neural Network Society*, *24*(6), 575–591.

Moutoussis, M., Eldar, E., & Dolan, R. J. (2017). Building a New Field of Computational Psychiatry. *Biological Psychiatry*, *82*(6), 388–390.

Moutoussis, M., Fearon, P., El-Deredy, W., Dolan, R. J., & Friston, K. J. (2014). Bayesian inferences about the self (and others): A review. *Consciousness and Cognition*, *25*, 67–76.

Moutoussis, M., Shahar, N., Hauser, T. U., & Dolan, R. J. (2018). Computation in Psychotherapy, or How Computational Psychiatry Can Aid Learning-Based Psychological Therapies. *Computational Psychiatry (Cambridge, Mass.)*, *2*, 50–73.

Moutoussis, M., Story, G. W., & Dolan, R. J. (2015). The computational psychiatry of reward: broken brains or misguided minds? *Frontiers in Psychology*, *6*, 1445.

Mowrer, O. H. (1951). Two-factor learning theory: summary and comment. *Psychological Review*, *58*(5), 350–354.

Mucci, A., Dima, D., Soricelli, A., Volpe, U., Bucci, P., Frangou, S., Prinster, A., Salvatore, M., Galderisi, S., & Maj, M. (2015). Is avolition in schizophrenia associated with a deficit of dorsal caudate activity? A functional magnetic resonance imaging study during reward anticipation and feedback. *Psychological Medicine*, *45*(8), 1765–1778.

Mujica-Parodi, L. R., Korgaonkar, M., Ravindranath, B., Greenberg, T., Tomasi, D., Wagshul, M., Ardekani, B., Guilfoyle, D., Khan, S., Zhong, Y., & Others. (2009). Limbic dysregulation is associated with lowered heart rate variability and increased trait anxiety in healthy adults. *Human Brain Mapping*, *30*(1), 47–58.

Muller, T. H., Mars, R. B., Behrens, T. E., & O'Reilly, J. X. (2019). Control of entropy in neural models of environmental state. *eLife*, *8*. https://doi.org/10.7554/eLife.39404

Mumford, D. (1992). On the computational architecture of the neocortex. *Biological Cybernetics*, *66*(3), 241–251.

Mumford, J. A., Poline, J.-B., & Poldrack, R. A. (2015). Orthogonalization of regressors in FMRI models. *PloS One*, *10*(4), e0126255.

Munn, N. L. (1950). *Handbook of psychological research on the rat; an introduction to animal psychology*. *598*. https://psycnet.apa.org/fulltext/1950-05580-000.pdf

Näätänen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology*, *42*(1), 25–32.

Näätänen, R., Kujala, T., & Winkler, I. (2011). Auditory processing that leads to conscious perception: a unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology*, *48*(1), 4–22.

Nair, A., Rutledge, R. B., & Mason, L. (2020). Under the Hood: Using Computational Psychiatry to Make Psychological Therapies More Mechanism-Focused. *Frontiers in Psychiatry / Frontiers Research Foundation*, *11*, 140.

Nakahara, H. (2014). Multiplexing signals in reinforcement learning with internal models and dopamine. *Current Opinion in Neurobiology*, *25*, 123–129.

Nakahara, H., & Hikosaka, O. (2012). Learning to represent reward structure: a key to adapting to complex environments. *Neuroscience Research*, *74*(3-4), 177–183.

Narita, K., Murata, T., Hamada, T., Takahashi, T., Omori, M., Suganuma, N., Yoshida, H., & Wada, Y. (2007). Interactions among higher trait anxiety, sympathetic activity, and endothelial function in the elderly. *Journal of Psychiatric Research*, *41*(5), 418–427.

Nassar, M. R., Bruckner, R., & Frank, M. J. (2019). Statistical context dictates the relationship between feedback-related EEG signals and learning. *eLife*, *8*. https://doi.org/10.7554/eLife.46975

Nassar, M. R., Bruckner, R., Gold, J. I., Li, S.-C., Heekeren, H. R., & Eppinger, B. (2016). Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nature Communications*, *7*, 11609.

Nassar, M. R., McGuire, J. T., Ritz, H., & Kable, J. W. (2019). Dissociable Forms of Uncertainty-Driven Representational Change Across the Human Brain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *39*(9), 1688–1698.

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*, *15*(7), 1040–1046.

Nassar, M. R., Wilson, R. C., Heasly, B., & Gold, J. I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(37), 12366–12378.

Nawijn, L., van Zuiden, M., Frijling, J. L., Koch, S. B. J., Veltman, D. J., & Olff, M. (2015). Reward functioning in PTSD: a systematic review exploring the mechanisms underlying anhedonia. *Neuroscience and Biobehavioral Reviews*, *51*, 189–204.

Nelson, E. A., Lickel, J. J., Sy, J. T., Dixon, L. J., & Deacon, B. J. (2010). Probability and cost biases in social phobia: nature, specificity, and relationship to treatment outcome. *Journal of Cognitive Psychotherapy*, *24*(3), 213–228.

Newell, A., Shaw, J. C., & Simon, H. A. (1959). Report on a general problem solving program. *IFIP Congress*, *256*, 64.

Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. H. (2004). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neuroscience and Biobehavioral Reviews*, *28*(4), 441–448.

Nishi, A., Shirado, H., Rand, D. G., & Christakis, N. A. (2015). Inequality and visibility of wealth in experimental social networks. *Nature*, *526*(7573), 426–429.

Nitschke, J. B., Heller, W., Imig, J. C., McDonald, R. P., & Miller, G. A. (2001). Distinguishing Dimensions of Anxiety and Depression. *Cognitive Therapy and Research*, *25*(1), 1–22.

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154.

Niv, Y., Daw, N., & Dayan, P. (2005). How fast to work: Response vigor, motivation and tonic dopamine. *Advances in Neural Information Processing Systems*, *18*, 1019–1026.

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520.

Niv, Y., Duff, M. O., & Dayan, P. (2005). Dopamine, uncertainty and TD learning. *Behavioral and Brain Functions: BBF*, *1*, 6.

Niv, Y., Joel, D., & Dayan, P. (2006). A normative perspective on motivation. *Trends in Cognitive Sciences*, *10*(8), 375–381.

Noonan, M. P., Walton, M. E., Behrens, T. E. J., Sallet, J., Buckley, M. J., & Rushworth, M. F. S. (2010). Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(47), 20547–20552.

Nord, C. L., Prabhu, G., Nolte, T., Fonagy, P., Dolan, R., & Moutoussis, M. (2017). Vigour in active avoidance. *Scientific Reports*, *7*(1), 60.

Nour, M. M., Dahoun, T., Schwartenbeck, P., Adams, R. A., FitzGerald, T. H. B., Coello, C., Wall, M. B., Dolan, R. J., & Howes, O. D. (2018). Dopaminergic basis for signaling belief updates, but not surprise, and the link to paranoia. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(43), E10167–E10176.

Oatley, K., & Johnson-Laird, P. N. (2014). Cognitive approaches to emotions. *Trends in Cognitive Sciences*, *18*(3), 134–140.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*(5669), 452–454.

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, *4*(1), 95–102.

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current Opinion in Neurobiology*, *14*(6), 769–776.

O'Doherty, J. P. (2016). Multiple Systems for the Motivational Control of Behavior and Associated Neural Substrates in Humans. *Current Topics in Behavioral Neurosciences*, *27*, 291–312.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329–337.

O'keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Clarendon Press.

Olatunji, B. O., Broman-Fulks, J. J., Bergman, S. M., Green, B. A., & Zlomke, K. R. (2010). A taxometric investigation of the latent structure of worry: dimensionality and associations with depression, anxiety, and stress. *Behavior Therapy*, *41*(2), 212–228.

Olds, J., & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology*, *47*(6), 419–427.

Olds, M. E., & Fobes, J. L. (1981). The central basis of motivation: intracranial self-stimulation studies. *Annual Review of Psychology*, *32*, 523–574.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*(6583), 607–609.

Olvet, D. M., & Hajcak, G. (2008). The error-related negativity (ERN) and psychopathology: toward an endophenotype. *Clinical Psychology Review*, *28*(8), 1343–1354.

O'Neill, M., & Schultz, W. (2010). Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron*, *68*(4), 789–800.

Onge, J. R. S., Stopper, C. M., Zahm, D. S., & Floresco, S. B. (2012). Separate prefrontal-subcortical circuits mediate different components of risk-based decision making. *Journal of Neuroscience*, *32*(8), 2886-2899. https://www.jneurosci.org/content/32/8/2886.short

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 156869.

O'Reilly, J. X. (2013). Making predictions in a changing world—inference, uncertainty, and learning. *Frontiers in Neuroscience*, *7*. https://doi.org/10.3389/fnins.2013.00105

O'Reilly, J. X., Jbabdi, S., & Behrens, T. E. J. (2012). How can a Bayesian approach inform neuroscience? *The European Journal of Neuroscience*, *35*(7), 1169–1179.

O'Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(38), E3660–E3669.

Ostwald, D., Spitzer, B., Guggenmos, M., Schmidt, T. T., Kiebel, S. J., & Blankenburg, F. (2012). Evidence for neural encoding of Bayesian surprise in human somatosensation. *NeuroImage*, *62*(1), 177–188.

Oswald, M. E., & Grosjean, S. (2004). Confirmation bias. *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory*, *79*. https://books.google.co.uk/books?hl=en&lr=&id=MS5Fr8safgEC&oi=fnd&pg=PA79&dq=oswald+Cognitive+Illusions:+A+Handbook+on+Fallacies+and+Biases+in+Thinking,+Judgement+and+Memory.+&ots=9PizPK9LTe&sig=6hSkAVUH9VMzcnAQDHUd7WqXq1s

Oudeyer, P.-Y., & Kaplan, F. (2009). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics*, *1*, 6.

Ousdal, O. T., Andreassen, O. A., Server, A., & Jensen, J. (2014). Increased amygdala and visual cortex activity and functional connectivity towards stimulus novelty is associated with state anxiety. *PloS One*, *9*(4), e96146.

Overman, M. J., Sarrazin, V., Browning, M., & O'Shea, J. (2021). Stimulating human prefrontal cortex increases reward learning. *bioRxiv*. https://www.biorxiv.org/content/10.1101/2021.01.27.428488v1.abstract

Paavilainen, P., Jaramillo, M., Näätänen, R., & Winkler, I. (1999). Neuronal populations in the human brain extracting invariant relationships from acoustic variance. *Neuroscience Letters*, *265*(3), 179–182.

Padmala, S., & Pessoa, L. (2011). Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *Journal of Cognitive Neuroscience*, *23*(11), 3419–3432.

Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(44), 14004–14014.

Padoa-Schioppa, C. (2011). Neurobiology of economic choice: a good-based model. *Annual Review of Neuroscience*, *34*, 333–359.

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

Padoa-Schioppa, C., & Conen, K. E. (2017). Orbitofrontal Cortex: A Neural Circuit for Economic Decisions. *Neuron*, *96*(4), 736–754.

Paliwal, S., Mosley, P. E., Breakspear, M., Coyne, T., Silburn, P., Aponte, E., Mathys, C., & Stephan, K. E. (2019). Subjective estimates of uncertainty during gambling and impulsivity after subthalamic deep brain stimulation for Parkinson's disease. *Scientific Reports*, *9*(1), 14795.

Palmer, C. E., Auksztulewicz, R., Ondobaka, S., & Kilner, J. M. (2019). Sensorimotor beta power reflects the precision-weighting afforded to sensory prediction errors. *NeuroImage*, *200*, 59–71.

Palmer, C. J., Seth, A. K., & Hohwy, J. (2015). The felt presence of other minds: Predictive processing, counterfactual predictions, and mentalising in autism. *Consciousness and Cognition*, *36*, 376–389.

Palminteri, S., Lebreton, M., Worbe, Y., Grabli, D., Hartmann, A., & Pessiglione, M. (2009). Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(45), 19179–19184.

Parkkonen, L., Fujiki, N., & Mäkelä, J. P. (2009). Sources of auditory brainstem responses revisited: contribution by magnetoencephalography. *Human Brain Mapping*, *30*(6), 1772–1782.

Park, S.-C., & Kim, Y.-K. (2020). Anxiety Disorders in the DSM-5: Changes, Controversies, and Future Directions. *Advances in Experimental Medicine and Biology*, *1191*, 187–196.

Parr, T., Benrimoh, D. A., Vincent, P., & Friston, K. J. (2018). Precision and False Perceptual Inference. *Frontiers in Integrative Neuroscience*, *12*, 39.

Parr, T., & Friston, K. J. (2017). Uncertainty, epistemics and active inference. *Journal of the Royal Society, Interface / the Royal Society*, *14*(136). https://doi.org/10.1098/rsif.2017.0376

Parr, T., & Friston, K. J. (2018). The Anatomy of Inference: Generative Models and Brain Structure. *Frontiers in Computational Neuroscience*, *12*, 90.

Parr, T., Markovic, D., Kiebel, S. J., & Friston, K. J. (2019). Neuronal message passing using Mean-field, Bethe, and Marginal approximations. *Scientific Reports*, *9*(1), 1889.

Parr, T., Rees, G., & Friston, K. J. (2018). Computational Neuropsychology and Bayesian Inference. *Frontiers in Human Neuroscience*, *12*, 61.

Paulus, F. M., Rademacher, L., Schäfer, T. A. J., Müller-Pinzler, L., & Krach, S. (2015). Journal Impact Factor Shapes Scientists' Reward Signal in the Prospect of Publication. *PloS One*, *10*(11), e0142537.

Paulus, M. P., Feinstein, J. S., Simmons, A., & Stein, M. B. (2004). Anterior cingulate activation in high trait anxious subjects is related to altered error processing during decision making. *Biological Psychiatry*, *55*(12), 1179–1187.

Paulus, M. P., & Stein, M. B. (2006). An insular view of anxiety. *Biological Psychiatry*, *60*(4), 383–387.

Paulus, M. P., & Stein, M. B. (2010). Interoception in anxiety and depression. *Brain Structure & Function*, *214*(5-6), 451–463.

Paulus, M. P., & Yu, A. J. (2012). Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends in Cognitive Sciences*, *16*(9), 476–483.

Pavlenko, V. B., Chernyi, S. V., & Goubkina, D. G. (2009). EEG Correlates of Anxiety and Emotional Stability in Adult Healthy Subjects. *Neurophysiology*, *41*(5), 337–345.

Pavlov, I. P. (1927). Conditioned reflexes, translated by GV Anrep. *London: Oxford*.

Payzan-LeNestour, E., & Bossaerts, P. (2011). Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS Computational Biology*, *7*(1), e1001048.

Payzan-LeNestour, E., Dunne, S., Bossaerts, P., & O'Doherty, J. P. (2013). The neural representation of unexpected uncertainty during value-based decision making. *Neuron*, *79*(1), 191–201.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*(6), 532–552.

Pearl, J. (1984). Intelligent search strategies for computer problem solving. *Addision Wesley*. https://pdfs.semanticscholar.org/2a2d/5c2532cd2d1cf2b612a50baa2bb2fe1b5735.pdf

Pellicano, E., & Burr, D. (2012). When the world becomes "too real": a Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, *16*(10), 504–510.

Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (2011). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier.

Penny, W., & Holmes, A. (2007). Random effects analysis. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*, 156–165.

Perkins, A. M., & Corr, P. J. (2014). Anxiety as an adaptive emotion. In W. G. Parrott (Ed.), *The positive side of negative emotions , (pp* (Vol. 304, pp. 37–54). The Guilford Press, xvi.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042–1045.

Pessiglione, M., Vinckier, F., Bouret, S., Daunizeau, J., & Le Bouc, R. (2017). Why not try harder? Computational approach to motivation deficits in neuro-psychiatric diseases. *Brain: A Journal of Neurology*. https://doi.org/10.1093/brain/awx278

Pessoa, L. (2017). A Network Model of the Emotional Brain. *Trends in Cognitive Sciences*, *21*(5), 357–371.

Peters, A., McEwen, B. S., & Friston, K. (2017). Uncertainty and stress: Why it causes diseases and how it is mastered by the brain. *Progress in Neurobiology*, *156*, 164–188.

Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active Inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology*, *134*, 17–35.

Pezzulo, G., Rigoli, F., & Friston, K. J. (2018). Hierarchical Active Inference: A Theory of Motivated Control. *Trends in Cognitive Sciences*, *22*(4), 294–306.

Pfeiffer, B. E., & Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, *497*(7447), 74–79.

Phelps, E. A., Lempert, K. M., & Sokol-Hessner, P. (2014). Emotion and decision making: multiple modulatory neural circuits. *Annual Review of Neuroscience*, *37*, 263–287.

Philiastides, M. G., Biele, G., Vavatzanidis, N., Kazzer, P., & Heekeren, H. R. (2010). Temporal dynamics of prediction error processing during reward-based decision making. *NeuroImage*, *53*(1), 221–232.

Pickering, A. D., & Pesola, F. (2014). Modeling dopaminergic and other processes involved in learning from reward prediction error: contributions from an individual differences perspective. *Frontiers in Human Neuroscience*, *8*, 740.

Pine, D. S. (2017). Clinical Advances From a Computational Approach to Anxiety. *Biological Psychiatry*, *82*(6), 385–387.

Pinotsis, D. A., Loonis, R., Bastos, A. M., Miller, E. K., & Friston, K. J. (2019). Bayesian modelling of induced responses and neuronal rhythms. *Brain topography, 32*(4), 569-582.https://doi.org/10.1007/s10548-016-0526-y

Piray, P., & Daw, N. D. (2019). A transparent model for learning in volatile environments. In *bioRxiv* (p. 701466). https://doi.org/10.1101/701466

Piray, P., & Daw, N. D. (2020a). A simple model for learning in volatile environments. *PLoS Computational Biology*, *16*(7), e1007963.

Piray, P., & Daw, N. D. (2020b). Unpredictability vs. volatility and the control of learning. In *bioRxiv* (p. 2020.10.05.327007). https://doi.org/10.1101/2020.10.05.327007

Piray, P., Dezfouli, A., Heskes, T., Frank, M. J., & Daw, N. D. (2019). Hierarchical Bayesian inference for concurrent model fitting and comparison for group studies. *PLoS Computational Biology*, *15*(6), e1007043.

Piray, P., Ly, V., Roelofs, K., Cools, R., & Toni, I. (2019). Emotionally Aversive Cues Suppress Neural Systems Underlying Optimal Learning in Socially Anxious Individuals. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *39*(8), 1445–1456.

Pittig, A., Arch, J. J., Lam, C. W. R., & Craske, M. G. (2013). Heart rate and heart rate variability in panic, social anxiety, obsessive–compulsive, and generalized anxiety disorders at baseline and in response to relaxation and hyperventilation. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *87*(1), 19–27.

Pittig, A., Treanor, M., LeBeau, R. T., & Craske, M. G. (2018). The role of associative fear and avoidance learning in anxiety disorders: Gaps and directions for future research. *Neuroscience and Biobehavioral Reviews*, *88*, 117–140.

Plaisted, K., Swettenham, J., & Rees, L. (1999). Children with autism show local precedence in a divided attention task and global precedence in a selective attention task. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *40*(5), 733–742.

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

Plous, S. (1993). *The psychology of judgment and decision making*. Mcgraw-Hill Book Company.

Polezzi, D., Lotto, L., Daum, I., Sartori, G., & Rumiati, R. (2008). Predicting outcomes of decisions in the brain. *Behavioural Brain Research*, *187*(1), 116–122.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–2148.

Pomè, A., Binda, P., Cicchini, G. M., & Burr, D. C. (2020). Pupillometry correlates of visual priming, and their dependency on autistic traits. *Journal of Vision*, *20*(3), 3.

Porta-Casteràs, D., Fullana, M. A., Tinoco, D., Martínez-Zalacaín, I., Pujol, J., Palao, D. J., Soriano-Mas, C., Harrison, B. J., Via, E., & Cardoner, N. (2020). Prefrontal-amygdala connectivity in trait anxiety and generalized anxiety disorder: Testing the boundaries between healthy and pathological worries. *Journal of Affective Disorders*, *267*, 211–219.

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, *32*(1), 3-25.

Potes, C., Brunner, P., Gunduz, A., Knight, R. T., & Schalk, G. (2014). Spatial and temporal relationships of electrocorticographic alpha and gamma activity during auditory processing. *NeuroImage*, *97*, 188–195.

Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: knowns and unknowns. *Nature Neuroscience*, *16*(9), 1170–1178.

Powers, A. R., Mathys, C., & Corlett, P. R. (2017). Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science*, *357*(6351), 596–600.

Press, C., & Yon, D. (2019). Perceptual Prediction: Rapidly Making Sense of a Noisy World. *Current Biology: CB*, *29*(15), R751–R753.

Preuschoff, K., & Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. *Annals of the New York Academy of Sciences*, *1104*, 135–146.

Preuschoff, K., Bossaerts, P., & Quartz, S. R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron*, *51*(3), 381–390.

Preuschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(11), 2745–2752.

Preuschoff, K., 't Hart, B. M., & Einhäuser, W. (2011). Pupil Dilation Signals Surprise: Evidence for Noradrenaline's Role in Decision Making. *Frontiers in Neuroscience*, *5*, 115.

Prinzmetal, W., Zvinyatskovskiy, A., Gutierrez, P., & Dilem, L. (2009). Voluntary and involuntary attention have different consequences: the effect of perceptual difficulty. *Quarterly Journal of Experimental Psychology* , *62*(2), 352–369.

Pulcu, E., & Browning, M. (2017). Affective bias as a rational response to the statistics of rewards and punishments. *eLife*, *6*. https://doi.org/10.7554/eLife.27879

Pulcu, E., & Browning, M. (2019). The Misestimation of Uncertainty in Affective Disorders. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2019.07.007

Pury, C. L., & Mineka, S. (1997). Covariation bias for blood-injury stimuli and aversive outcomes. *Behaviour Research and Therapy*, *35*(1), 35–47.

Quintana, D. S., Alvares, G. A., & Heathers, J. A. J. (2016). Guidelines for Reporting Articles on Psychiatry and Heart rate variability (GRAPH): recommendations to advance research communication. *Translational Psychiatry*, *6*, e803.

Quirk, G. J., & Mueller, D. (2008). Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *33*(1), 56–72.

Radakovic, R., & Abrahams, S. (2014). Developing a new apathy measurement scale: Dimensional Apathy Scale. *Psychiatry Research*, *219*(3), 658–663.

Raghunathan, R., & Pham, M. T. (1999). All Negative Moods Are Not Equal: Motivational Influences of Anxiety and Sadness on Decision Making. *Organizational Behavior and Human Decision Processes*, *79*(1), 56–77.

Ramaswamy, S., & Markram, H. (2015). Anatomy and physiology of the thick-tufted layer 5 pyramidal neuron. *Frontiers in Cellular Neuroscience*, *9*, 233.

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

Rao, H. M., Mayo, J. P., & Sommer, M. A. (2016). Circuits for presaccadic visual remapping. *Journal of Neurophysiology*, *116*(6), 2624–2636.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87.

Rauch, S. L., Shin, L. M., & Wright, C. I. (2003). Neuroimaging studies of amygdala function in anxiety disorders. *Annals of the New York Academy of Sciences*, *985*, 389–410.

Raymond, J. G., Steele, J. D., & Seriès, P. (2017). Modeling Trait Anxiety: From Computational Processes to Personality. *Frontiers in Psychiatry / Frontiers Research Foundation*, *8*, 1.

Rebollo, I., Devauchelle, A.-D., Béranger, B., & Tallon-Baudry, C. (2018). Stomach-brain synchrony reveals a novel, delayed-connectivity resting-state network in humans. *eLife*, *7*. https://doi.org/10.7554/eLife.33321

Redish, A. D. (1999). Beyond the cognitive map: from place cells to episodic memory. MIT press.

Redish, A. D. (2013). *The Mind Within the Brain: How We Make Decisions and How Those Decisions Go Wrong*. OUP USA.

Redish, A. D., & Gordon, J. A. (2016). *Computational Psychiatry: New Perspectives on Mental Illness*. MIT Press.

Redish, A. D., Jensen, S., Johnson, A., & Kurth-Nelson, Z. (2007). Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling. *Psychological Review*, *114*(3), 784–805.

Redish, A. D., Jensen, S., & Johnson, A. (2008). Addiction as vulnerabilities in the decision process. *The Behavioral and Brain Sciences*, *31*(4), 461–487.

Reed, E. J., Uddenberg, S., Suthaharan, P., Mathys, C. D., Taylor, J. R., Groman, S. M., & Corlett, P. R. (2020). Paranoia as a deficit in non-social belief updating. *eLife*, *9*. https://doi.org/10.7554/eLife.56345

Reichert, D. P., Seriès, P., & Storkey, A. J. (2013). Charles Bonnet syndrome: evidence for a generative model in the cortex? *PLoS Computational Biology*, *9*(7), e1003134.

Reid, C. R., MacDonald, H., Mann, R. P., Marshall, J. A. R., Latty, T., & Garnier, S. (2016). Decision-making without a brain: how an amoeboid organism solves the two-armed bandit. *Journal of the Royal Society, Interface / the Royal Society*, *13*(119). https://doi.org/10.1098/rsif.2016.0030

Remmers, C., & Zander, T. (2018). Why you don't see the Forest for the trees when you are anxious: Anxiety impairs intuitive decision making. *Clinical Psychological Science*, *6*(1), 48–62.

Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin*, *72*(2), 77.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, *2*, 64–99.

Reynolds, J. N., Hyland, B. I., & Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature*, *413*(6851), 67–70.

Riceberg, J. S., & Shapiro, M. L. (2017). Orbitofrontal Cortex Signals Expected Outcomes with Predictive Codes When Stable Contingencies Promote the Integration of Reward History. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(8), 2010–2021.

Richards, A., & Millwood, B. (1989). Colour-identification of differentially valenced words in anxiety. *Cognition and Emotion*, *3*(2), 171–176.

Richards, H. J., Benson, V., Donnelly, N., & Hadwin, J. A. (2014). Exploring the function of selective attention and hypervigilance for threat in anxiety. *Clinical Psychology Review*, *34*(1), 1–13.

Rich, E. L., & Wallis, J. D. (2016). Decoding subjective decisions from orbitofrontal cortex. *Nature Neuroscience*, *19*(7), 973–980.

Ridderinkhof, K. R., van den Wildenberg, W. P. M., Segalowitz, S. J., & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition*, *56*(2), 129–140.

Roach, B. J., & Mathalon, D. H. (2008). Event-related EEG time-frequency analysis: an overview of measures and an analysis of early gamma band phase locking in schizophrenia. *Schizophrenia bulletin*, *34*(5), 907-926. https://academic.oup.com/schizophreniabulletin/article-abstract/34/5/907/1886959

Robbins, T. W., & Arnsten, A. F. T. (2009). The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annual Review of Neuroscience*, *32*, 267–287.

Robert, P. H., Clairet, S., Benoit, M., Koutaich, J., Bertogliati, C., Tible, O., Caci, H., Borg, M., Brocker, P., & Bedoucha, P. (2002). The apathy inventory: assessment of apathy and awareness in Alzheimer's disease, Parkinson's disease and mild cognitive impairment. *International Journal of Geriatric Psychiatry*, *17*(12), 1099–1105.

Roberts, M. J., Lowet, E., Brunet, N. M., Ter Wal, M., Tiesinga, P., Fries, P., & De Weerd, P. (2013). Robust gamma coherence between macaque V1 and V2 by dynamic frequency matching. *Neuron*, *78*(3), 523–536.

Robinson, O. J., Bond, R. L., & Roiser, J. P. (2015). The impact of threat of shock on the framing effect and temporal discounting: executive functions unperturbed by acute stress? *Frontiers in Psychology*, *6*, 1315.

Robinson, O. J., Charney, D. R., Overstreet, C., & Vytal, K. (2012). The adaptive threat bias in anxiety: amygdala–dorsomedial prefrontal cortex coupling and aversive amplification. *NeuroImage*. https://www.sciencedirect.com/science/article/pii/S1053811911014017

Robinson, O. J., Krimsky, M., & Grillon, C. (2013). The impact of induced anxiety on response inhibition. *Frontiers in Human Neuroscience*, *7*, 69.

Robinson, O. J., Letkiewicz, A. M., Overstreet, C., Ernst, M., & Grillon, C. (2011). The effect of induced anxiety on cognition: threat of shock enhances aversive processing in healthy individuals. *Cognitive, Affective & Behavioral Neuroscience*, *11*(2), 217–227.

Robinson, O. J., Overstreet, C., Charney, D. R., Vytal, K., & Grillon, C. (2013). Stress increases aversive prediction error signal in the ventral striatum. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(10), 4129–4133.

Robinson, O. J., Pike, A. C., Cornwell, B., & Grillon, C. (2019). The translational neural circuitry of anxiety. *Journal of Neurology, Neurosurgery, and Psychiatry*, *90*(12), 1353–1360.

Robinson, O. J., Vytal, K., Cornwell, B. R., & Grillon, C. (2013). The impact of anxiety upon cognition: perspectives from human threat of shock studies. *Frontiers in Human Neuroscience*, *7*, 203.

Robinson, T. E., & Flagel, S. B. (2009). Dissociating the predictive and incentive motivational properties of reward-related cues through the study of individual differences. *Biological Psychiatry*, *65*(10), 869–873.

Roelofs, K., Bakvis, P., Hermans, E. J., van Pelt, J., & van Honk, J. (2007). The effects of social stress and cortisol responses on the preconscious selective attention to social threat. *Biological Psychology*, *75*(1), 1–7.

Roesch, M. R., Calu, D. J., Esber, G. R., & Schoenbaum, G. (2010). Neural correlates of variations in event processing during learning in basolateral amygdala. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(7), 2464–2471.

Roesch, M. R., Esber, G. R., Li, J., Daw, N. D., & Schoenbaum, G. (2012). Surprise! Neural correlates of Pearce--Hall and Rescorla--Wagner coexist within the brain. *The European Journal of Neuroscience*, *35*(7), 1190–1200.

Rogers, R. D. (2011). The roles of dopamine and serotonin in decision making: evidence from pharmacological experiments in humans. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *36*(1), 114–132.

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

Rolls, E. T., & Grabenhorst, F. (2008). The orbitofrontal cortex and beyond: from affect to decision-making. *Progress in Neurobiology*, *86*(3), 216–244.

Ronconi, L., Vitale, A., Federici, A., Pini, E., Molteni, M., & Casartelli, L. (2020). Altered neural oscillations and connectivity in the beta band underlie detail-oriented visual processing in autism. *NeuroImage: Clinical*, *28*, 102484.

Rosenblueth, A., & Wiener, N. (1945). The Role of Models in Science. *Philosophy of Science*, *12*(4), 316–321.

Rosen, J. B., & Schulkin, J. (1998). From normal fear to pathological anxiety. *Psychological Review*, *105*(2), 325–350.

Roxburgh, A. D., White, D. J., & Cornwell, B. R. (2020). Anxious arousal alters prefrontal cortical control of stopping. *The European Journal of Neuroscience*. https://doi.org/10.1111/ejn.14976

Rudebeck, P. H., Ripple, J. A., Mitz, A. R., Averbeck, B. B., & Murray, E. A. (2017). Amygdala Contributions to Stimulus–Reward Encoding in the Macaque Medial and Orbital Frontal Cortex during Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(8), 2186–2202.

Rudebeck, P. H., Saunders, R. C., Lundgren, D. A., & Murray, E. A. (2017). Specialized Representations of Value in the Orbital and Ventrolateral Prefrontal Cortex: Desirability versus Availability of Outcomes. *Neuron*, *95*(5), 1208–1220.e5.

Ruiz, M. H., Koelsch, S., & Bhattacharya, J. (2009). Decrease in early right alpha band phase synchronization and late gamma band oscillations in processing syntax in music. *Human Brain Mapping*, *30*(4), 1207–1225.

Ruscio, J., & Mullen, T. (2012). Confidence Intervals for the Probability of Superiority Effect Size Measure and the Area Under a Receiver Operating Characteristic Curve. *Multivariate Behavioral Research*, *47*(2), 201–223.

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, *11*(4), 389–397.

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

Rushworth, M. F. S., Noonan, M. P., Boorman, E. D., Walton, M. E., & Behrens, T. E. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron*, *70*(6), 1054–1069.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161.

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition and Emotion*, *23*(7), 1259–1283.

Rutledge, R. B., Dean, M., Caplin, A., & Glimcher, P. W. (2010). Testing the reward prediction error hypothesis with an axiomatic model. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(40), 13525–13536.

Rutledge, R. B., Lazzaro, S. C., Lau, B., Myers, C. E., Gluck, M. A., & Glimcher, P. W. (2009). Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(48), 15104–15114.

Rutledge, R. B., Moutoussis, M., Smittenaar, P., Zeidman, P., Taylor, T., Hrynkiewicz, L., Lam, J., Skandali, N., Siegel, J. Z., Ousdal, O. T., Prabhu, G., Dayan, P., Fonagy, P., & Dolan, R. J. (2017). Association of Neural and Emotional Impacts of Reward Prediction Errors With Major Depression. *JAMA Psychiatry* , *74*(8), 790–797.

Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2014). A computational and neural model of momentary subjective well-being. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(33), 12252–12257.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54–67.

Sadacca, B. F., Jones, J. L., & Schoenbaum, G. (2016). Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. *eLife*, *5*. https://doi.org/10.7554/eLife.13665

Saez, A., Rigotti, M., Ostojic, S., Fusi, S., & Salzman, C. D. (2015). Abstract Context Representations in Primate Amygdala and Prefrontal Cortex. *Neuron*, *87*(4), 869–881.

Saez, R. A., Saez, A., Paton, J. J., Lau, B., & Salzman, C. D. (2017). Distinct Roles for the Amygdala and Orbitofrontal Cortex in Representing the Relative Amount of Expected Reward. *Neuron*, *95*(1), 70–77.e3.

Sailer, U., Robinson, S., Fischmeister, F. P. S., König, D., Oppenauer, C., Lueger-Schuster, B., Moser, E., Kryspin-Exner, I., & Bauer, H. (2008). Altered reward processing in the nucleus accumbens and mesial prefrontal cortex of patients with posttraumatic stress disorder. *Neuropsychologia*, *46*(11), 2836–2844.

Salamone, J. D., & Correa, M. (2002). Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behavioural Brain Research*, *137*(1-2), 3–25.

Salamone, J. D., Correa, M., Farrar, A. M., Nunes, E. J., & Pardo, M. (2009). Dopamine, behavioral economics, and effort. *Frontiers in Behavioral Neuroscience*, *3*, 13.

Salamone, J. D., Correa, M., Mingote, S., & Weber, S. M. (2003). Nucleus accumbens dopamine and the regulation of effort in food-seeking behavior: implications for studies of natural motivation, psychiatry, and drug abuse. *The Journal of Pharmacology and Experimental Therapeutics*, *305*(1), 1–8.

Salamone, J. D., Cousins, M. S., & Bucher, S. (1994). Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response selection in a T-maze cost/benefit procedure. *Behavioural Brain Research*, *65*(2), 221–229.

Salamone, J. D., Pardo, M., Yohn, S. E., López-Cruz, L., SanMiguel, N., & Correa, M. (2016). Mesolimbic Dopamine and the Regulation of Motivated Behavior. *Current Topics in Behavioral Neurosciences*, *27*, 231–257.

Salamone, J. D., Yohn, S. E., López-Cruz, L., San Miguel, N., & Correa, M. (2016). Activational and effort-related aspects of motivation: neural mechanisms and implications for psychopathology. *Brain: A Journal of Neurology*, *139*(Pt 5), 1325–1347.

Sales, A. C., Friston, K. J., Jones, M. W., Pickering, A. E., & Moran, R. J. (2019). Locus Coeruleus tracking of prediction errors optimises cognitive flexibility: An Active Inference model. *PLoS Computational Biology*, *15*(1), e1006267.

Salmaso, L. (2003). Synchronized Permutation Tests in 2 k Factorial Designs. *Communications in Statistics - Theory and Methods*, *32*(7), 1419–1437.

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235.

Samejima, K., Ueda, Y., Doya, K., & Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science*, *310*(5752), 1337–1340.

Sara, S. J., & Bouret, S. (2012). Orienting and reorienting: the locus coeruleus mediates cognition through arousal. *Neuron*, *76*(1), 130–141.

Savine, A. C., & Braver, T. S. (2010). Motivated cognitive control: reward incentives modulate preparatory neural activity during task-switching. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(31), 10294–10305.

Schacter, D. L., & Addis, D. R. (2009). On the nature of medial temporal lobe contributions to the constructive simulation of future events. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1521), 1245–1253.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews. Neuroscience*, *8*(9), 657–661.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2008). Episodic simulation of future events: concepts, data, and applications. *Annals of the New York Academy of Sciences*, *1124*, 39–60.

Scheeringa, R., & Fries, P. (2019). Cortical layers, rhythms and BOLD signals. *NeuroImage*, *197*, 689–698.

Scherer, K. R., Schorr, A., & Johnstone, T. (2001). *Appraisal Processes in Emotion: Theory, Methods, Research*. Oxford University Press.

Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., & Pessiglione, M. (2012). Neural mechanisms underlying motivation of mental versus physical effort. *PLoS Biology*, *10*(2), e1001266.

Schmidt, R., Herrojo Ruiz, M., Kilavik, B. E., Lundqvist, M., Starr, P. A., & Aron, A. R. (2019). Beta Oscillations in Working Memory, Executive Control of Movement and Thought, and Sensorimotor Function. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *39*(42), 8231–8238.

Schmitz, A., & Grillon, C. (2012). Assessing fear and anxiety in humans using the threat of predictable and unpredictable aversive events (the NPU-threat test). *Nature Protocols*, *7*(3), 527–532.

Schneider, T. R., Hipp, J. F., Domnick, C., Carl, C., Büchel, C., & Engel, A. K. (2018). Modulation of neuronal oscillatory activity in the beta- and gamma-band is associated with current individual anxiety levels. *NeuroImage*, *178*, 423–434.

Schoffelen, J.-M., Oostenveld, R., & Fries, P. (2005). Neuronal coherence as a mechanism of effective corticospinal interaction. *Science*, *308*(5718), 111–113.

Scholl, J., & Klein-Flügge, M. (2018). Understanding psychiatric disorder by capturing ecologically relevant features of learning and decision-making. *Behavioural Brain Research*, *355*, 56–75.

Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron*, *91*(6), 1402–1412.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*(1), 1–27.

Schultz, W. (2015). Neuronal Reward and Decision Signals: From Theories to Data. *Physiological Reviews*, *95*(3), 853–951.

Schultz, W. (2016). Dopamine reward prediction-error signalling: a two-component response. *Nature Reviews. Neuroscience*, *17*(3), 183–195.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599.

Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience*, *23*, 473–500.

Schultz, W., Tremblay, L., & Hollerman, J. R. (1998). Reward prediction in primate basal ganglia and frontal cortex. *Neuropharmacology*, *37*(4-5), 421–429.

Schultz, W., Tremblay, L., & Hollerman, J. R. (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral Cortex* , *10*(3), 272–284.

Schutter, D. J. L. G. (2016). A Cerebellar Framework for Predictive Coding and Homeostatic Regulation in Depressive Disorder. *Cerebellum* , *15*(1), 30–33.

Schwabe, L., & Wolf, O. T. (2009). Stress prompts habit behavior in humans. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(22), 7191–7198.

Schwartenbeck, P., Fitzgerald, T., Dolan, R. J., & Friston, K. (2013). Exploration, novelty, surprise, and free energy minimization. *Frontiers in Psychology*, *4*, 710.

Schwartenbeck, P., FitzGerald, T. H. B., Mathys, C., Dolan, R., & Friston, K. (2015). The Dopaminergic Midbrain Encodes the Expected Certainty about Desired Outcomes. *Cerebral Cortex* , *25*(10), 3434–3445.

Schwartenbeck, P., FitzGerald, T. H. B., Mathys, C., Dolan, R., Kronbichler, M., & Friston, K. (2015). Evidence for surprise minimization over value maximization in choice behavior. *Scientific Reports*, *5*, 16575.

Schwertman, N. C., Owens, M. A., & Adnan, R. (2004). A simple more general boxplot method for identifying outliers. *Computational Statistics & Data Analysis*, *47*(1), 165–174.

Sedley, W., Gander, P. E., Kumar, S., Kovach, C. K., Oya, H., Kawasaki, H., Howard, M. A., & Griffiths, T. D. (2016). Neural signatures of perceptual inference. *eLife*, *5*, e11476.

Segal, D. L. (2010). Diagnostic and statistical manual of mental disorders (DSM-IV-TR). In *The Corsini Encyclopedia of Psychology*. John Wiley & Sons, Inc. https://doi.org/10.1002/9780470479216.corpsy0271

Sengupta, B., & Friston, K. (2017). Approximate Bayesian inference as a gauge theory. In *arXiv [q-bio.NC]*. arXiv. http://arxiv.org/abs/1705.06614

Sengupta, B., Tozzi, A., Cooray, G. K., Douglas, P. K., & Friston, K. J. (2016). Towards a Neuronal Gauge Theory. *PLoS Biology*, *14*(3), e1002400.

Seo, H., & Lee, D. (2008). Cortical mechanisms for reinforcement learning in competitive games. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1511), 3845–3857.

Seo, M., Lee, E., & Averbeck, B. B. (2012). Action selection and action value in frontal-striatal circuits. *Neuron*, *74*(5), 947–960.

Sequeira, P., Melo, F. S., & Paiva, A. (2011). Emotion-Based Intrinsic Motivation for Reinforcement Learning Agents. *Affective Computing and Intelligent Interaction*, 326–336.

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, *17*(11), 565–573.

Seth, A. K., & Friston, K. J. (2016). Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1708). https://doi.org/10.1098/rstb.2016.0007

Sevgi, M., Diaconescu, A. O., Tittgemeyer, M., & Schilbach, L. (2016). Social Bayes: Using Bayesian Modeling to Study Autistic Trait–Related Differences in Social Cognition. *Biological Psychiatry*, *80*(2), 112–119.

Seymour, B., O'Doherty, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., Friston, K. J., & Frackowiak, R. S. (2004). Temporal difference models describe higher-order learning in humans. *Nature*, *429*(6992), 664–667.

Shachter, R. D., & Peot, M. A. (1992). Decision Making Using Probabilistic Inference Methods. In D. Dubois, M. P. Wellman, B. D'Ambrosio, & P. Smets (Eds.), *Uncertainty in Artificial Intelligence* (pp. 276–283). Morgan Kaufmann.

Shackman, A. J., Maxwell, J. S., McMenamin, B. W., Greischar, L. L., & Davidson, R. J. (2011). Stress potentiates early and attenuates late stages of visual processing. *The*

Journal of Neuroscience: The Official Journal of the Society for Neuroscience, *31*(3), 1156–1161.

Shackman, A. J., Sarinopoulos, I., Maxwell, J. S., Pizzagalli, D. A., Lavric, A., & Davidson, R. J. (2006). Anxiety selectively disrupts visuospatial working memory. *Emotion* , *6*(1), 40–61.

Shah, A., & Frith, U. (1983). An islet of ability in autistic children: a research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *24*(4), 613–620.

Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the block design task? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *34*(8), 1351–1364.

Shapiro, C., Carl, S., & Varian, H. R. (1998). *Information rules: a strategic guide to the network economy*. Harvard Business Press. https://books.google.ca/books?hl=en&lr=&id=aE_J4Iv_PVEC&oi=fnd&pg=PP15&ots=o2LmXXicKY&sig=mDEI0pHXyZpGHjr52y9XvBcIPMQ

Sharpe, M. J., & Schoenbaum, G. (2016). Back to basics: Making predictions in the orbitofrontal–amygdala circuit. *Neurobiology of Learning and Memory*, *131*, 201–206.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a Rational and Mechanistic Account of Mental Effort. *Annual Review of Neuroscience*, *40*, 99–124.

Shiba, Y., Santangelo, A. M., & Roberts, A. C. (2016). Beyond the Medial Regions of Prefrontal Cortex in the Regulation of Fear and Anxiety. *Frontiers in Systems Neuroscience*, *10*, 12.

Shine, J. M., Müller, E. J., Munn, B., Cabral, J., Moran, R. J., & Breakspear, M. (2021). Computational models link cellular mechanisms of neuromodulation to large-scale neural dynamics. *Nature Neuroscience*. https://doi.org/10.1038/s41593-021-00824-6

Shin, H., Law, R., Tsutsui, S., Moore, C. I., & Jones, S. R. (2017). The rate of transient beta frequency events predicts behavior across tasks and species. *eLife*, *6*. https://doi.org/10.7554/eLife.29086

Shipp, S. (2005). The importance of being agranular: a comparative account of visual and motor cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 797–814.

Shipp, S. (2016). Neural Elements for Predictive Coding. *Frontiers in Psychology*, *7*, 1792.

Shipp, S., Adams, R. A., & Friston, K. J. (2013). Reflections on agranular architecture: predictive coding in the motor cortex. *Trends in Neurosciences*, *36*(12), 706–716.

Shizgal, P. (1997). Neural basis of utility estimation. *Current Opinion in Neurobiology*, *7*(2), 198–208.

Shohamy, D., Myers, C. E., Grossman, S., Sage, J., & Gluck, M. A. (2005). The role of dopamine in cognitive sequence learning: evidence from Parkinson's disease. *Behavioural Brain Research*, *156*(2), 191–199.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T., & Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature*, *550*(7676), 354–359.

Simmel, G. (2004). *The philosophy of money*. Routledge.

Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., & Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision Research*, *49*(22), 2705–2739.

Simon, D. A., & Daw, N. D. (2011). Environmental statistics and the trade-off between model-based and TD learning in humans. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 24* (pp. 127–135). Curran Associates, Inc.

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99–118.

Simonson, I., & Tversky, A. (1992). Choice in Context: Tradeoff Contrast and Extremeness Aversion. *JMR, Journal of Marketing Research*, *29*(3), 281–295.

Simpson, C. A., Diaz-Arteche, C., Eliby, D., Schwartz, O. S., Simmons, J. G., & Cowan, C. S. M. (2021). The gut microbiota in anxiety and depression - A systematic review. *Clinical Psychology Review*, *83*, 101943.

Simpson, J. R., Jr, Drevets, W. C., Snyder, A. Z., Gusnard, D. A., & Raichle, M. E. (2001). Emotion-induced changes in human medial prefrontal cortex: II. During anticipatory anxiety. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 688–693.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica: Journal of the Econometric Society*, *48*(1), 1–48.

Singh, I. L., Dwivedi, C. B., & Sinha, M. M. (1979). Effects of anxiety on vigilance. *Perceptual and Motor Skills*, *49*(1), 142.

Singh, S. P. (2014). Magnetoencephalography: Basic principles. *Annals of Indian Academy of Neurology*, *17*(Suppl 1), S107–S112.

Slee, A., Nazareth, I., Bondaronek, P., Liu, Y., Cheng, Z., & Freemantle, N. (2019). Pharmacological treatments for generalised anxiety disorder: a systematic review and network meta-analysis. *The Lancet*, *393*(10173), 768–777.

Smith, R., Badcock, P., & Friston, K. J. (2021). Recent advances in the application of predictive coding and active inference models within clinical neuroscience. *Psychiatry and Clinical Neurosciences*, *75*(1), 3–13.

Smith, R., Kuplicki, R., Feinstein, J., Forthman, K. L., Stewart, J. L., Paulus, M. P., Tulsa 1000 investigators, & Khalsa, S. S. (2020). A Bayesian computational model reveals a failure to adapt interoceptive precision estimates across depression, anxiety, eating, and substance use disorders. *PLoS Computational Biology*, *16*(12), e1008484.

Smith, R., Parr, T., & Friston, K. J. (2019). Simulating Emotions: An Active Inference Model of Emotional State Inference and Emotion Concept Learning. *Frontiers in Psychology*, *10*, 2844.

Smittenaar, P., FitzGerald, T. H. B., Romei, V., Wright, N. D., & Dolan, R. J. (2013). Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. *Neuron*, *80*(4), 914–919.

Soch, J., & Allefeld, C. (2018). MACS--a new SPM toolbox for model assessment, comparison and selection. *Journal of Neuroscience Methods*. https://www.sciencedirect.com/science/article/pii/S0165027018301468

Soch, J., Haynes, J.-D., & Allefeld, C. (2016). How to avoid mismodelling in GLM-based fMRI data analysis: cross-validated Bayesian model selection. *NeuroImage*, *141*, 469–489.

Sockeel, P., Dujardin, K., Devos, D., Denève, C., Destée, A., & Defebvre, L. (2006). The Lille apathy rating scale (LARS), a new instrument for detecting and quantifying apathy: validation in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *77*(5), 579–584.

Sokolowska, E., & Hovatta, I. (2013). Anxiety genetics--findings from cross-species genome-wide approaches. *Biology of Mood & Anxiety Disorders*, *3*(1), 1–8.

Soltani, A., & Izquierdo, A. (2019). Adaptive learning under expected and unexpected uncertainty. *Nature Reviews. Neuroscience*, *20*(10), 635–644.

Solway, A., & Botvinick, M. M. (2012). Goal-directed decision making as probabilistic inference: a computational framework and potential neural correlates. *Psychological Review*, *119*(1), 120–154.

Spielberger, C. D. (1979). *Understanding stress and anxiety*. Harper & Row.

Spielberger, C. D. (1983a). *Manual for the State-Trait Anxiety Inventory; Palo Alto, CA, Ed*. Consulting Psychologists Press, Inc.: Columbia, MO, USA.

Spielberger, C. D. (1983b). *Manual for the State-Trait Anxiety Inventory STAI (form Y)(" self-evaluation questionnaire")*. https://ubir.buffalo.edu/xmlui/handle/10477/1873

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1968). *State-Trait Anxiety Inventory (STAI): Test Manual for Form X*. Consulting Psychologists Press.

Spielberger, C. D., Gorsuch, R. L., Lushene, R. E., Vagg, P. R., & Jacobs, G. A. (1970). State-Trait Anxiety Inventory (STAI) for Adults-Manual. *Mind Garden, Inc. , Menlo Park, CA*.

Spinhoven, P., van Hemert, A. M., & Penninx, B. W. J. H. (2017). Experiential Avoidance and Bordering Psychological Constructs as Predictors of the Onset, Relapse and Maintenance of Anxiety Disorders: One or Many? *Cognitive Therapy and Research*, *41*(6), 867–880.

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

Sporn, S., Hein, T., & Herrojo Ruiz, M. (2020). Alterations in the amplitude and burst rate of beta oscillations impair reward-dependent motor learning in anxiety. *eLife*, *9*. https://doi.org/10.7554/eLife.50654

Spratling, M. W. (2017). A review of predictive coding algorithms. *Brain and Cognition*, *112*(Supplement C), 92–97.

Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*(2), 195–231.

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

Staddon, J. E. R., & Cerutti, D. T. (2003). Operant conditioning. *Annual Review of Psychology*, *54*, 115–144.

Starkstein, S. E., Mayberg, H. S., Preziosi, T. J., Andrezejewski, P., Leiguarda, R., & Robinson, R. G. (1992). Reliability, validity, and clinical correlates of apathy in Parkinson's disease. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *4*(2), 134–139.

Stefanics, G., Heinzle, J., Horváth, A. A., & Stephan, K. E. (2018). Visual Mismatch and Predictive Coding: A Computational Single-Trial ERP Study. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *38*(16), 4020–4030.

Stefanics, G., Kremláček, J., & Czigler, I. (2014). Visual mismatch negativity: a predictive coding view. *Frontiers in Human Neuroscience*, *8*, 666.

Steimer, T. (2002). The biology of fear-and anxiety-related behaviors. *Dialogues in Clinical Neuroscience*, *4*(3), 231.

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, *16*(7), 966–973.

Steiner, A. P., & Redish, A. D. (2012). The road not taken: neural correlates of decision making in orbitofrontal cortex. *Frontiers in Neuroscience*, *6*, 131.

Stein, M. B., & Craske, M. G. (2017). Treating Anxiety in 2017: Optimizing Care to Improve Outcomes. *JAMA: The Journal of the American Medical Association*, *318*(3), 235–236.

Stephan, K. E., Bach, D. R., Fletcher, P. C., Flint, J., Frank, M. J., Friston, K. J., Heinz, A., Huys, Q. J. M., Owen, M. J., Binder, E. B., Dayan, P., Johnstone, E. C., Meyer-Lindenberg, A., Montague, P. R., Schnyder, U., Wang, X.-J., & Breakspear, M. (2016). Charting the landscape of priority problems in psychiatry, part 1: classification and diagnosis. *The Lancet. Psychiatry*, *3*(1), 77–83.

Stephan, K. E., Iglesias, S., Heinzle, J., & Diaconescu, A. O. (2015). Translational Perspectives for Computational Neuroimaging. *Neuron*, *87*(4), 716–732.

Stephan, K. E., Manjaly, Z. M., Mathys, C. D., Weber, L. A. E., Paliwal, S., Gard, T., Tittgemeyer, M., Fleming, S. M., Haker, H., Seth, A. K., & Petzschner, F. H. (2016). Allostatic Self-efficacy: A Metacognitive Theory of Dyshomeostasis-Induced Fatigue and Depression. *Frontiers in Human Neuroscience*, *10*, 550.

Stephan, K. E., & Mathys, C. (2014). Computational approaches to psychiatry. *Current Opinion in Neurobiology*, *25*, 85–92.

Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *NeuroImage*, *46*(4), 1004–1017.

Stephan, K. E., Petzschner, F. H., Kasper, L., Bayer, J., Wellstein, K. V., Stefanics, G., Pruessmann, K. P., & Heinzle, J. (2019). Laminar fMRI and computational theories of brain function. *NeuroImage*, *197*, 699–706.

Stöber, J. (1997). Trait anxiety and pessimistic appraisal of risk and chance. *Personality and Individual Differences*, *22*(4), 465–476.

Stolyarova, A., & Izquierdo, A. (2017). Complementary contributions of basolateral amygdala and orbitofrontal cortex to value learning under uncertainty. *eLife*, *6*. https://doi.org/10.7554/eLife.27483

Stolyarova, A., Rakhshan, M., Hart, E. E., O'Dell, T. J., Peters, M. A. K., Lau, H., Soltani, A., & Izquierdo, A. (2019). Contributions of anterior cingulate cortex and basolateral amygdala to decision confidence and learning under uncertainty. *Nature Communications*, *10*(1), 4704.

Stopper, C. M., Tse, M. T. L., Montes, D. R., Wiedman, C. R., & Floresco, S. B. (2014). Overriding phasic dopamine signals redirects action selection during risk/reward decision making. *Neuron*, *84*(1), 177–189.

Strange, B. A., & Dolan, R. J. (2004). β-Adrenergic modulation of emotional memory-evoked human amygdala and hippocampal responses. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(31), 11454–11458.

Suarez, J. A., Howard, J. D., Schoenbaum, G., & Kahnt, T. (2019). Sensory prediction errors in the human midbrain signal identity violations independent of perceptual distance. *eLife*, *8*. https://doi.org/10.7554/eLife.43962

Summerfield, C., Behrens, T. E., & Koechlin, E. (2011). Perceptual classification in a rapidly changing environment. *Neuron*, *71*(4), 725–736.

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

Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., & Hirsch, J. (2006). Predictive codes for forthcoming perception in the frontal cortex. *Science*, *314*(5803), 1311–1314.

Summerfield, C., & Koechlin, E. (2008). A neural representation of prior information during perceptual inference. *Neuron*, *59*(2), 336–347.

Sutton, R. S. (1992). Gain adaptation beats least squares. *Proceedings of the 7th Yale Workshop on Adaptive and Learning Systems*, *161-168*.

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychological Review*, *88*(2), 135.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Sutton, R. S., Barto, A. G., & Others. (1998). *Introduction to reinforcement learning* (Vol. 135). MIT press Cambridge.

Szabó, M., & Lovibond, P. F. (2006). Worry episodes and perceived problem solving: A diary-based approach. *Anxiety, Stress, and Coping*, *19*(2), 175–187.

Takahashi, T., Ikeda, K., Ishikawa, M., Kitamura, N., Tsukasaki, T., Nakama, D., & Kameda, T. (2005). Anxiety, reactivity, and social stress-induced cortisol elevation in humans. *Neuro Endocrinology Letters*, *26*(4), 351–354.

Tan, H., Wade, C., & Brown, P. (2016). Post-Movement Beta Activity in Sensorimotor Cortex Indexes Confidence in the Estimations from Internal Models. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(5), 1516–1528.

Tappin, B. M., & Gadsby, S. (2019). Biased belief in the Bayesian brain: A deeper look at the evidence. *Consciousness and Cognition*, *68*, 107–114.

Tappin, B. M., van der Leer, L., & McKay, R. T. (2017). The heart trumps the head: Desirability bias in political belief revision. *Journal of Experimental Psychology. General*, *146*(8), 1143–1149.

Taylor, J. A. (1953). A personality scale of manifest anxiety. *Journal of Abnormal Psychology*, *48*(2), 285–290.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: statistics, structure, and abstraction. *Science*, *331*(6022), 1279–1285.

Thayer, J. F., Friedman, B. H., & Borkovec, T. D. (1996). Autonomic characteristics of generalized anxiety disorder and worry. *Biological Psychiatry*, *39*(4), 255–266.

Thomson, A. M., & Bannister, A. P. (2003). Interlaminar connections in the neocortex. *Cerebral Cortex* , *13*(1), 5–14.

Thorndike, E. L. (1898). Animal intelligence: an experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements*, *2*(4), i.

Thorndike, E. L. (1927). The Law of Effect. *The American Journal of Psychology*, *39*(1/4), 212–222.

Thorndike, E. L. (1932). *The fundamentals of learning*. https://psycnet.apa.org/record/2006-04535-000

Tobler, P. N., Dickinson, A., & Schultz, W. (2003). Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *23*(32), 10402–10410.

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, *307*(5715), 1642–1645.

Todorovic, A., Schoffelen, J.-M., van Ede, F., Maris, E., & de Lange, F. P. (2015). Temporal expectation and attention jointly modulate auditory oscillatory activity in the beta band. *PloS One*, *10*(3), e0120288.

Todorovic, A., van Ede, F., Maris, E., & de Lange, F. P. (2011). Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(25), 9118–9123.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, *315*(5811), 515–518.

Tovote, P., Fadok, J. P., & Lüthi, A. (2015). Neuronal circuits for fear and anxiety. *Nature Reviews. Neuroscience*, *16*(6), 317–331.

Treadway, M. T., Bossaller, N. A., Shelton, R. C., & Zald, D. H. (2012). Effort-based decision-making in major depressive disorder: a translational model of motivational anhedonia. *Journal of Abnormal Psychology*, *121*(3), 553–558.

Treadway, M. T., & Zald, D. H. (2011). Reconsidering anhedonia in depression: lessons from translational neuroscience. *Neuroscience and Biobehavioral Reviews*, *35*(3), 537–555.

Tremblay, L., Hollerman, J. R., & Schultz, W. (1998). Modifications of reward expectation-related neuronal activity during learning in primate striatum. *Journal of Neurophysiology*, *80*(2), 964–977.

Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature*, *398*(6729), 704–708.

Treynor, W., Gonzalez, R., & Nolen-Hoeksema, S. (2003). Rumination Reconsidered: A Psychometric Analysis. *Cognitive Therapy and Research*, *27*(3), 247–259.

Tsien, J. Z. (2015). Principles of Intelligence: On Evolutionary Logic of the Brain. *Frontiers in Systems Neuroscience*, *9*, 186.

Tsutsui, K.-I., Grabenhorst, F., Kobayashi, S., & Schultz, W. (2016). A dynamic code for economic object valuation in prefrontal cortex neurons. *Nature Communications*, *7*, 12554.

Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2). Reading, MA.

Tversky, A., & Kahneman, D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, *47*(2), 263–291.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458.

Tye, K. M., Prakash, R., Kim, S.-Y., Fenno, L. E., Grosenick, L., Zarabi, H., Thompson, K. R., Gradinaru, V., Ramakrishnan, C., & Deisseroth, K. (2011). Amygdala circuitry mediating reversible and bidirectional control of anxiety. *Nature*, *471*(7338), 358–362.

Uhlhaas, P. J., & Singer, W. (2010). Abnormal neural oscillations and synchrony in schizophrenia. *Nature Reviews. Neuroscience*, *11*(2), 100–113.

Uhlhaas, P. J., & Singer, W. (2012). Neuronal dynamics and neuropsychiatric disorders: toward a translational paradigm for dysfunctional large-scale networks. *Neuron*, *75*(6), 963–980.

Valton, V., Ahn, W. Y., Bond, R. L., & Dayan, P. (2019). Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nature Human*. https://www.nature.com/articles/s41562-019-0628-0

Van de Cruys, S., Van der Hallen, R., & Wagemans, J. (2017). Disentangling signal and noise in autism spectrum disorder. *Brain and Cognition*, *112*, 78–83.

van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., van Kesteren, E.-J., Derks, K., Dablander, F., Gronau, Q. F., Kucharský, Š., Komarlu Narendra Gupta, A., Sarafoglou, A., Voelkel, J. G., Stefan, A., Ly, A., Hinne, M., Matzke, D., & Wagenmakers, E.-J. (2020). A Tutorial on Conducting and Interpreting a Bayesian ANOVA in JASP. *L'annee psychologique*, *120*(1), 73–96.

van der Meer, M. A. A., Johnson, A., Schmitzer-Torbert, N. C., & Redish, A. D. (2010). Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. *Neuron*, *67*(1), 25–32.

van der Meer, M. A. A., & Redish, A. D. (2009). Covert Expectation-of-Reward in Rat Ventral Striatum at Decision Points. *Frontiers in Integrative Neuroscience*, *3*, 1.

Van Der Meer, M., Kurth-Nelson, Z., & Redish, A. D. (2012). Information processing in decision-making systems. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *18*(4), 342–359.

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharský, Š., Ly, A., Marsman, M., Matzke, D., Gupta, A. R. K. N., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E.-J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, *28*(3), 813–826.

van Ede, F., de Lange, F., Jensen, O., & Maris, E. (2011). Orienting attention to an upcoming tactile event involves a spatially and temporally specific modulation of sensorimotor alpha- and beta-band oscillations. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(6), 2016–2024.

Vanhove, J. (2020). Collinearity isn't a disease that needs curing. In *PsyArXiv*. https://doi.org/10.31234/osf.io/mv2wx

Van Kerkoerle, T., Self, M. W., & Dagnino, B. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the*. https://www.pnas.org/content/111/40/14332.short

van Marle, H. J. F., Hermans, E. J., Qin, S., & Fernández, G. (2009). From specificity to sensitivity: how acute stress affects amygdala processing of biologically salient stimuli. *Biological Psychiatry*, *66*(7), 649–655.

van Pelt, S., Heil, L., Kwisthout, J., Ondobaka, S., van Rooij, I., & Bekkering, H. (2016). Beta- and gamma-band activity reflect predictive coding in the processing of causal events. *Social Cognitive and Affective Neuroscience*, *11*(6), 973–980.

van Schalkwyk, G. I., Volkmar, F. R., & Corlett, P. R. (2017). A Predictive Coding Account of Psychotic Symptoms in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *47*(5), 1323–1340.

Vernon, G. (2020). *Decision making in tennis: exploring the use of kinematic and contextual information during anticipatory performance* [Victoria University]. http://vuir.vu.edu.au/40723/1/VERNON%20Georgina-thesis_nosignature.pdf

Vidaurre, C., Haufe, S., Jorajuría, T., Müller, K.-R., & Nikulin, V. V. (2020). Sensorimotor Functional Connectivity: A Neurophysiological Factor Related to BCI Performance. *Frontiers in Neuroscience*, *14*, 575081.

Vincent, P., Parr, T., Benrimoh, D., & Friston, K. J. (2019). With an eye on uncertainty: Modelling pupillary responses to environmental volatility. *PLoS Computational Biology*, *15*(7), e1007126.

Vlaev, I., Chater, N., Stewart, N., & Brown, G. D. A. (2011). Does the brain calculate value? *Trends in Cognitive Sciences*, *15*(11), 546–554.

Vo, K., Rutledge, R. B., Chatterjee, A., & Kable, J. W. (2014). Dorsal striatum is necessary for stimulus-value but not action-value learning in humans. *Brain: A Journal of Neurology*, *137*(Pt 12), 3129–3135.

Volz, K. G., Schubotz, R. I., & von Cramon, D. Y. (2003). Predicting events of varying probability: uncertainty investigated by fMRI. *NeuroImage*, *19*(2 Pt 1), 271–280.

von Neumann, J., Morgenstern, O., & Kuhn, H. W. (2007). *Theory of Games and Economic Behavior (Commemorative Edition)*. Princeton University Press.

Voss, A., Schroeder, R., Heitmann, A., Peters, A., & Perz, S. (2015). Short-term heart rate variability--influence of gender and age in healthy subjects. *PloS One*, *10*(3), e0118308.

Vossel, S., Bauer, M., Mathys, C., Adams, R. A., Dolan, R. J., Stephan, K. E., & Friston, K. J. (2014). Cholinergic stimulation enhances Bayesian belief updating in the deployment of spatial attention. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(47), 15735–15742.

Vossel, S., Mathys, C., Stephan, K. E., & Friston, K. J. (2015). Cortical Coupling Reflects Bayesian Belief Updating in the Deployment of Spatial Attention. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(33), 11532–11542.

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585–594.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron*, *30*(3), 829–841.

Vytal, K., Cornwell, B., Arkin, N., & Grillon, C. (2012). Describing the interplay between anxiety and cognition: from impaired performance under low cognitive load to reduced anxiety under high load. *Psychophysiology*, *49*(6), 842–852.

Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, *412*(6842), 43–48.

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). *Journal of Personality and Social Psychology*, *100*(3), 426–432.

Wager, T. D., Kang, J., Johnson, T. D., Nichols, T. E., Satpute, A. B., & Barrett, L. F. (2015). A Bayesian model of category-specific emotional brain responses. *PLoS Computational Biology*, *11*(4), e1004066.

Wallis, J. D. (2011). Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nature Neuroscience*, *15*(1), 13–19.

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, *36*(8), 1870–1884.

Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E. M., & Rushworth, M. F. S. (2006). Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. *Neural Networks: The Official Journal of the International Neural Network Society*, *19*(8), 1302–1314.

Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., Hassabis, D., & Botvinick, M. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nature Neuroscience*, *21*(6), 860–868.

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

Wang, X.-J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews*, *90*(3), 1195–1268.

Wassum, K. M., Cely, I. C., Balleine, B. W., & Maidment, N. T. (2011). μ-Opioid receptor activation in the basolateral amygdala mediates the learning of increases but not decreases in the incentive value of a food reward. *Journal of Neuroscience*, *31*(5), 1591–1599.

Wassum, K. M., & Izquierdo, A. (2015). The basolateral amygdala in reward learning and addiction. *Neuroscience and Biobehavioral Reviews*, *57*, 271–283.

Wassum, K. M., Ostlund, S. B., Balleine, B. W., & Maidment, N. T. (2011). Differential dependence of Pavlovian incentive motivation and instrumental incentive learning processes on dopamine signaling. *Learning & Memory* , *18*(7), 475–483.

Watabe-Uchida, M., Zhu, L., Ogawa, S. K., Vamanrao, A., & Uchida, N. (2012). Whole-brain mapping of direct inputs to midbrain dopamine neurons. *Neuron*, *74*(5), 858–873.

Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine Learning*, *8*(3), 279–292.

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin*, *134*(2), 163–206.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.

Watson, D., Weber, K., Assenheimer, J. S., Clark, L. A., Strauss, M. E., & McCormick, R. A. (1995). Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. *Journal of Abnormal Psychology*, *104*(1), 3–14.

Watson, J. B. (1907). Kinæsthetic and organic sensations: Their role in the reactions of the white rat to the maze. *The Psychological Review: Monograph Supplements*, *8*(2), i.

Weber, L. (2020). *Perception as Hierarchical Bayesian Inference - Toward noninvasive readouts of exteroceptive and interoceptive processing*. ETH Zurich, Switzerland.

Weber, L. A., Diaconescu, A. O., Mathys, C., Schmidt, A., Kometer, M., Vollenweider, F., & Stephan, K. E. (2020). Ketamine Affects Prediction Errors about Statistical Regularities: A Computational Single-Trial Analysis of the Mismatch Negativity. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *40*(29), 5658–5668.

Weber, L., Diaconescu, A., Tomiello, S., Schöbi, D., Iglesias, S., Mathys, C., Haker, H., Stefanics, G., Schmidt, A., Kometer, M., Vollenweider, F. X., & Stephan, K. E. (2018). F157. HIERARCHICAL PREDICTION ERRORS DURING AUDITORY MISMATCH

UNDER PHARMACOLOGICAL MANIPULATIONS: A COMPUTATIONAL SINGLE-TRIAL EEG ANALYSIS. *Schizophrenia Bulletin*, *44*(Suppl 1), S281.

Weilnhammer, V. A., Stuke, H., Sterzer, P., & Schmack, K. (2018). The Neural Correlates of Hierarchical Predictions for Perceptual Decisions. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *38*(21), 5008–5021.

Wessel, J. R., & Aron, A. R. (2017). On the Globality of Motor Suppression: Unexpected Events and Their Influence on Behavior and Cognition. *Neuron*, *93*(2), 259–280.

Wessel, J. R., Danielmeier, C., Morton, J. B., & Ullsperger, M. (2012). Surprise and error: common neuronal architecture for the processing of errors and novelty. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(22), 7528–7537.

Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PloS One*, *8*(7), e68210.

Wetzel, M. C. (1963). Self-stimulation aftereffects and runway performance in the rat. *Journal of Comparative and Physiological Psychology*, *56*(4), 673–678.

Whalen, P. J. (2007). The uncertainty of it all. *Trends in Cognitive Sciences*, *11*(12), 499–500.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *18*(1), 411–418.

Whelan, R., Watts, R., Orr, C. A., Althoff, R. R., Artiges, E., Banaschewski, T., Barker, G. J., Bokde, A. L. W., Büchel, C., Carvalho, F. M., Conrod, P. J., Flor, H., Fauth-Bühler, M., Frouin, V., Gallinat, J., Gan, G., Gowland, P., Heinz, A., Ittermann, B., … IMAGEN Consortium. (2014). Neuropsychosocial profiles of current and future adolescent alcohol misusers. *Nature*, *512*(7513), 185–189.

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Anxiety enhances threat processing without competition among multiple inputs: a diffusion model analysis. *Emotion* , *10*(5), 662–677.

White, S. F., Geraci, M., Lewis, E., Leshin, J., Teng, C., Averbeck, B., Meffert, H., Ernst, M., Blair, J. R., Grillon, C., & Blair, K. S. (2017). Prediction Error Representation in Individuals With Generalized Anxiety Disorder During Passive Avoidance. *The American Journal of Psychiatry*, *174*(2), 110–117.

White, T. L., & Depue, R. A. (1999). Differential association of traits of fear and anxiety with norepinephrine- and dark-induced pupil reactivity. *Journal of Personality and Social Psychology*, *77*(4), 863–877.

Whittington, J. C. R., & Bogacz, R. (2017). An Approximation of the Error Backpropagation Algorithm in a Predictive Coding Network with Local Hebbian Synaptic Plasticity. *Neural Computation*, *29*(5), 1229–1262.

Wiecki, T. V., & Frank, M. J. (2010). Neurocomputational models of motor and cognitive deficits in Parkinson's disease. *Progress in Brain Research*, *183*, 275–297.

Wiecki, T. V., Poland, J., & Frank, M. J. (2015). Model-Based Cognitive Neuroscience Approaches to Computational Psychiatry: Clustering and Classification. *Clinical Psychological Science*, *3*(3), 378–399.

Wiedemann, K. (2015). *Anxiety and anxiety disorders*.

Williams, B. K., Sanders, R. H., Ryu, J. H., Graham-Smith, P., & Sinclair, P. J. (2020). The kinematic differences between accurate and inaccurate squash forehand drives for athletes of different skill levels. *Journal of Sports Sciences*, *38*(10), 1115–1123.

Williams, L. E., Oler, J. A., Fox, A. S., McFarlin, D. R., Rogers, G. M., Jesson, M. A. L., Davidson, R. J., Pine, D. S., & Kalin, N. H. (2015). Fear of the unknown: uncertain anticipation reveals amygdala alterations in childhood anxiety disorders. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *40*(6), 1428–1435.

Williams, L. M. (2016). Precision psychiatry: a neural circuit taxonomy for depression and anxiety. *The Lancet. Psychiatry*, *3*(5), 472–480.

Wilson, C. G., Nusbaum, A. T., Whitney, P., & Hinson, J. M. (2018). Trait anxiety impairs cognitive flexibility when overcoming a task acquired response and a preexisting bias. *PloS One*, *13*(9), e0204694.

Wilson, F. N., Macleod, A. G., & Barker, P. S. (1931). The potential variations produced by the heart beat at the apices of Einthoven's triangle. *American Heart Journal*, *7*(2), 207–211.

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *eLife*, *8*. https://doi.org/10.7554/eLife.49547

Wilson, R. C., Nassar, M. R., & Gold, J. I. (2010). Bayesian online learning of the hazard rate in change-point problems. *Neural Computation*, *22*(9), 2452–2476.

Wilson, R. C., Nassar, M. R., & Gold, J. I. (2013). A mixture of delta-rules approximation to bayesian inference in change-point problems. *PLoS Computational Biology*, *9*(7), e1003150.

Wilson, R. C., & Niv, Y. (2011). Inferring relevance in a changing world. *Frontiers in Human Neuroscience*, *5*, 189.

Winecoff, A., Clithero, J. A., Carter, R. M., Bergman, S. R., Wang, L., & Huettel, S. A. (2013). Ventromedial prefrontal cortex encodes emotional value. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(27), 11032–11039.

Winkielman, P., Knutson, B., Paulus, M., & Trujillo, J. L. (2007). Affective Influence on Judgments and Decisions: Moving towards Core Mechanisms. *Review of General Psychology: Journal of Division 1, of the American Psychological Association*, *11*(2), 179–192.

Winkler, I. (2007). Interpreting the Mismatch Negativity. *Journal of Psychophysiology*, *21*(3-4), 147–163.

Winstanley, C. A., & Floresco, S. B. (2016). Deciphering Decision Making: Variation in Animal Models of Effort- and Uncertainty-Based Choice Reveals Distinct Neural

Circuitries Underlying Core Cognitive Processes. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(48), 12069–12079.

Wise, R. A. (2004). Dopamine, learning and motivation. *Nature Reviews. Neuroscience*, *5*(6), 483–494.

Wise, R. A., & Rompre, P. P. (1989). Brain dopamine and reward. *Annual Review of Psychology*, *40*, 191–225.

Wise, R. A., Spindler, J., DeWit, H., & Gerberg, G. J. (1978). Neuroleptic-induced“ anhedonia” in rats: pimozide blocks reward quality of food. *Science*, *201*(4352), 262–264.

Wise, R. A., Spindler, J., & Legault, L. (1978). Major attenuation of food reward with performance-sparing doses of pimozide in the rat. *Canadian Journal of Psychology*, *32*(2), 77–85.

Wise, T., & Dolan, R. J. (2020). Associations between aversive learning processes and transdiagnostic psychiatric symptoms in a general population sample. *Nature Communications*, *11*(1), 4179.

Wise, T., Michely, J., Dayan, P., & Dolan, R. J. (2019). A computational account of threat-related attentional bias. *PLoS Computational Biology*, *15*(10), e1007341.

Wolf, O. T., Schulte, J. M., Drimalla, H., Hamacher-Dang, T. C., Knoch, D., & Dziobek, I. (2015). Enhanced emotional empathy after psychosocial stress in young healthy men. *Stress* , *18*(6), 631–637.

Wolpert, D. M., & Landy, M. S. (2012). Motor control is decision-making. *Current Opinion in Neurobiology*, *22*(6), 996–1003.

Womelsdorf, T., Fries, P., Mitra, P. P., & Desimone, R. (2006). Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature*, *439*(7077), 733–736.

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, *4*(1), 58–73.

Wright, N. D., Symmonds, M., & Dolan, R. J. (2013). Distinct encoding of risk and value in economic choice between multiple risky options. *NeuroImage*, *81*, 431–440.

Wunderlich, K., Beierholm, U. R., Bossaerts, P., & O'Doherty, J. P. (2011). The human prefrontal cortex mediates integration of potential causes behind observed outcomes. *Journal of Neurophysiology*, *106*(3), 1558–1569.

Wunderlich, K., Dayan, P., & Dolan, R. J. (2012). Mapping value based planning and extensively trained choice in the human brain. *Nature Neuroscience*, *15*(5), 786–791.

Wunderlich, K., Rangel, A., & O'Doherty, J. P. (2009). Neural computations underlying action-based decision making in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(40), 17199–17204.

Wunderlich, K., Smittenaar, P., & Dolan, R. J. (2012). Dopamine enhances model-based over model-free choice behavior. *Neuron*, *75*(3), 418–424.

Wunderlich, K., Symmonds, M., Bossaerts, P., & Dolan, R. J. (2011). Hedging your bets by learning reward correlations in the human brain. *Neuron*, *71*(6), 1141–1152.

Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research*, *1286*, 114–122.

Xia, L., Xu, P., Yang, Z., Gu, R., & Zhang, D. (2021). Impaired probabilistic reversal learning in anxiety: Evidence from behavioral and ERP findings. *NeuroImage. Clinical*, *31*, 102751.

Xing, D., Yeh, C.-I., Burns, S., & Shapley, R. M. (2012). Laminar analysis of visually evoked activity in the primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(34), 13871–13876.

Xu, P., Gu, R., Broster, L. S., Wu, R., Van Dam, N. T., Jiang, Y., Fan, J., & Luo, Y.-J. (2013). Neural basis of emotional decision making in trait anxiety. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(47), 18641–18653.

Yager, L. M., Garcia, A. F., Wunsch, A. M., & Ferguson, S. M. (2015). The ins and outs of the striatum: role in drug addiction. *Neuroscience*, *301*, 529-541. https://www.sciencedirect.com/science/article/pii/S0306452215005746?casa_token=SN

pnpCGTlncAAAAA:BBysQAc_RHVMvGd7nAbYmsYb9fX3kfuZFrWKBY3WIhkSV6xQ2x1R5Zj2nW9xWbQML-q49g

Yang, T., & Shadlen, M. N. (2007). Probabilistic reasoning by neurons. *Nature*, *447*(7148), 1075–1080.

Yasuda, A., Sato, A., Miyawaki, K., Kumano, H., & Kuboki, T. (2004). Error-related negativity reflects detection of negative reward prediction error. *Neuroreport*, *15*(16), 2561–2565.

Yerkes, R. M., & Morgulis, S. (1909). The method of Pawlow in animal psychology. *Psychological Bulletin*, *6*(8), 257.

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex* , *15*(5), 535–544.

Yon, D., de Lange, F. P., & Press, C. (2019). The Predictive Brain as a Stubborn Scientist. *Trends in Cognitive Sciences*, *23*(1), 6–8.

Yon, D., Gilbert, S. J., de Lange, F. P., & Press, C. (2018). Action sharpens sensory representations of expected outcomes. *Nature Communications*, *9*(1), 4288.

Yu, A. J. (2007). Adaptive Behavior: Humans Act as Bayesian Learners. *Current Biology: CB*, *17*(22), R977–R980.

Yu, A. J., & Dayan, P. (2002). Acetylcholine in cortical inference. *Neural Networks: The Official Journal of the International Neural Network Society*, *15*(4), 719–730.

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

Zarr, N., & Brown, J. W. (2016). Hierarchical error representation in medial prefrontal cortex. *NeuroImage*, *124*, 238–247.

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

Zénon, A. (2019). Eye pupil signals information gain. *Proceedings. Biological Sciences / The Royal Society*, *286*(1911), 20191593.

Zhang, Z., Cordeiro Matos, S., Jego, S., Adamantidis, A., & Séguéla, P. (2013). Norepinephrine drives persistent activity in prefrontal cortex via synergistic α1 and α2 adrenoceptors. *PloS One*, *8*(6), e66122.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361–370.

Zorowitz, S., Momennejad, I., & Daw, N. D. (2020). Anxiety, avoidance, and sequential evaluation. Computational Psychiatry (Cambridge, Mass.), 4(0), 1.