Algorithms of anxiety:

Computational neuroscience approaches to translational measures of anxiety

Yumeya Yamamori

Submitted for the degree of *Doctor of Philosophy*

Prepared under the supervision of Professors Jonathan P. Roiser and Oliver J. Robinson

Institute of Cognitive Neuroscience, University College London

October 2024

Declaration

I, Yumeya Yamamori, confirm that the work presented in my thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Anxiety disorders are among the most prevalent mental health conditions worldwide, yet progress in developing effective treatments has stagnated. This thesis explores the potential of computational neuroscience approaches to enhance our understanding of anxiety-related cognition and behaviour, with a focus on improving translational measures between human and non-human animal research.

The overarching aim was to develop and validate computational models of anxiety-related behaviour that can be applied across species, framing them as algorithms that can apply to both humans and non-human animals, and by doing so potentially bridge the gap between preclinical and clinical research. This work focused on two key paradigms: conditioned tests involving learned approach-avoidance conflict, and unconditioned tests examining spontaneous spatial exploration in aversive environments.

Using large online samples, I first developed and validated a novel human version of classical rodent conditioned anxiety tests in Chapter 2. Computational modelling revealed that anxiety was associated with heightened sensitivity to punishments relative to rewards, which mediated avoidance behaviour. I then conducted two experimental studies manipulating anxiety levels to test the causal role of this computational mechanism in Chapter 3. Pharmacologically reducing anxiety with lorazepam shifted sensitivities towards rewards, while threat-of-shock shifted sensitivities towards punishment and exacerbated avoidance behaviour.

For unconditioned anxiety tests, I proposed a new computational model of spatial exploration in aversive environments, validated using existing data from a human version of the elevated plus maze, in Chapter 4. This model provided novel insights into the cognitive mechanisms underlying anxiolytic and anxiogenic drug effects on exploration behaviour.

Across these studies, the computational approach consistently revealed mechanistic insights that were not apparent from traditional model-agnostic behavioural analyses alone. The models developed here offer a more precise, quantitative framework for translating anxiety-related processes between species. This work demonstrates the potential of computational methods to enhance the validity and interpretability of translational anxiety research, potentially accelerating the development of novel treatments.

Impact statement

My thesis presents novel computational approaches to studying anxiety that have potential for impact both within academia and beyond. By developing and validating translational models of anxiety that bridge human and animal research, this work addresses a critical gap in anxiety research and anxiety treatment development.

Within academia, the computational models introduced here offer more precise ways to conceptualise and measure anxiety-related processes across species. This provides a common "language" for researchers studying anxiety in humans and animal models, potentially accelerating translational research. The approach-avoidance reinforcement learning task developed in Chapters 2-3 and the spatial exploration model in Chapter 4 give researchers new tools to probe the mechanisms underlying anxiety and avoidance behaviours. These computational frameworks allow for more rigorous comparisons between animal models and human anxiety, addressing long-standing issues of construct and predictive validity in preclinical anxiety research.

Beyond academia, this research has potential implications for drug development and clinical practice. The computational parameters identified as sensitive to anxiety manipulations, such as the reward-punishment sensitivity index, could serve as quantitative biomarkers in preclinical and clinical drug trials. This may help pharmaceutical companies identify promising anxiolytic compounds earlier in the development pipeline, potentially reducing the high failure rates and costs associated with anxiety drug development. For clinicians, the tasks and computational metrics developed here could be adapted into diagnostic tools or measures of treatment response, enabling more personalised approaches to anxiety treatment.

In the longer term, the naturalistic computational psychiatry approach demonstrated in Chapter 4 presents opportunities for studying anxiety in more ecologically valid settings. This could lead to digital mental health applications that assess anxiety through everyday behaviours captured by smartphones or wearable devices. Such tools could enable earlier detection of anxiety disorders or continuous monitoring of symptom trajectories.

The impact of this research could potentially be realised through multiple pathways. Within academia, dissemination of the theory via publications and analysis code via online repositories will make these computational frameworks and tools available to the wider research community. Collaboration with preclinical researchers to apply these models to animal data will be crucial for validating their cross-species utility. Engaging with pharmaceutical companies could lead to

incorporation of these computational metrics into drug screening pipelines. Adaptation of the tasks for clinical use would require partnerships with clinicians and possibly digital health companies.

While realising the full impact of this work will require many more years of continued research, it provides a foundation for a more mechanistic, translationally valid approach to studying anxiety. By bridging levels of analysis from computational models to behaviour to anxiety experienced by humans, this research contributes to the broader goal of developing more effective, personalised treatments for anxiety disorders.

Research paper declaration

Parts of the General Introduction draw upon ideas explored in a peer-reviewed review paper:

Yamamori, Y., & Robinson, O. J. (2023). Computational perspectives on human fear and anxiety. Neuroscience & Biobehavioral Reviews, 144, 104959. https://doi.org/10.1016/j.neubiorev.2022.104959

Chapter 2 has been published as a peer-reviewed paper:

Yumeya Yamamori, Oliver J Robinson, Jonathan P Roiser (2023) Approach-avoidance reinforcement learning as a translational and computational model of anxiety-related avoidance *eLife* 12:RP87720. https://doi.org/10.7554/eLife.87720.4

The foundational idea for Chapter 4 was introduced in a peer-reviewed commentary paper:

Yamamori, Y., & Robinson, O. J. (2024). Thinking computationally in translational psychiatry. A commentary on Neville et al. (2024). Cognitive, Affective, & Behavioral Neuroscience, 24(2), 384-387. https://doi.org/10.3758/s13415-024-01172-1

Acknowledgements

I've had the unique privilege of benefiting from two 'primary' supervisors, Professors Jonathan Roiser and Oliver Robinson. I am so grateful for the trust you placed in me by giving me the intellectual freedom and space to explore ideas, fail gracefully, make infamously oversized slide decks, and ultimately through this, develop the skills and confidence I have today. Thank you for your combined mentorship, under which I have grown and learned so much as a scientist - I hope this continues long into the future. Having known both of you prior to starting my PhD, I knew that you lead a lab not just with brilliant minds but with an incredibly friendly and supportive environment. This played a considerable part in my decision to join this lab over others, and nothing has made me regret this choice since, so thank you also for allowing me to experience only the best side of academia.

Thank you also to Professors Benedetto de Martino, Quentin Huys and Christina Carlisi for their mentorship in the early stages of my PhD.

I am grateful to the Wellcome Trust for their generous research funding, and the staff at the UCL Institute of Cognitive Neuroscience for their support in facilitating the day-to-day aspects of the research. A special note of appreciation goes to all the research participants who took part in my studies, which unfortunately often involved unpleasant anxiety-inducing tasks. Thank you all for allowing the research to happen.

Thanks to my many friends and colleagues at UCL, in the Neuroscience and Mental Health lab, but also beyond the lab. To Karel, Yannis, Alex, Hugo and Ani for the times pre-Covid when every Friday was a celebration of some unpredictable form and length. To more recent colleagues Carlos, Millie, Agatha, Ashley, Emily, Pete and Benedict for reviving the social side of the lab after Covid when we had forgotten it. And of course to members new and old of the 407 office, Harry, Sarah, Ella, Mari and Prady, for all of our discussions, intellectual or not, over our daily lunches and coffees. Thank you all for making my time at UCL such a fond memory.

I must also thank my tutor from my school days, Shinya Minegishi, who first taught me, beyond all of our lessons on mathematics and Japanese, the value of a PhD as something that was both challenging and deeply rewarding.

Finally, thank you to my mother, Toyoko, and my partner, Jocelyn, for sharing this journey with me and supporting me through it all. I could not have done it without you.

Table of contents

1 General Introduction 10

1.1 The problem of anxiety 10

1.2 Defining anxiety: a brief history 11

1.3 From bench to bedside: cross-species translational research 13

1.4 The translational gap 16

1.5 What makes a valid translational test? 17

1.6 Anxiety as a problem of information processing 18

1.7 Bridging the translational gap: integrating translational and computational approaches 24

1.8 Outstanding questions 30

1.9 Aims and objectives of the present thesis 34

2 Approach-avoidance reinforcement learning as a translational and computational model of anxiety-related avoidance 37

2.1 Preface 37

2.2 Abstract 38

2.3 Introduction 39

2.4 Methods 42

2.5 Results 51

2.6 Discussion 57

2.7 Conclusion 62

2.8 Appendices 63

3 Computational effects of anxiolytic and anxiogenic manipulations on approach-avoidance reinforcement learning 81

3.1 Preface 81

3.2 Abstract 82

3.3 Introduction 83

3.4 Methods 85

3.5 Results 93

3.6 Discussion 102
3.7 Conclusion 107
3.8 Appendices 108
4 A computational model of spatial exploration in aversive environments 114
4.1 Preface 114
4.2 Abstract 115
4.3 Introduction 116
4.4 Methods 118
4.5 Results – model simulations 128
4.6 Results – application of the model to empirical data 133
4.7 Discussion 136
4.8 Conclusion 140
5 General Discussion 141
5.1 Overview 141
5.2 Summary of findings 141
5.3 Implications: returning to the outstanding questions 143
5.4 Broader implications 149
5.5 Limitations and future directions 151
5.6 Summary and concluding remarks 156
6 References 158

1 General Introduction

1.1 The problem of anxiety

Anxiety disorders are the most commonly reported mental health conditions worldwide, affecting over 300 million people globally as of 2019 - approximately 4% of the total population (Ferrari et al., 2022; Kessler et al., 2012). The prevalence and impact of anxiety have remained relatively stable since 1990 (Patel et al., 2016), with recent evidence indicating that the Covid-19 pandemic exacerbated global anxiety prevalence by up to 25% (Santomauro et al., 2021). Such widespread prevalence of anxiety disorders entails substantial economic costs. These conditions can be severely debilitating, to the point that they affect an individual's ability to function in the workplace (Deady et al., 2022). In the UK alone, approximately 18 million working days are lost annually due to work-related anxiety and depression (CIPD, 2023). On a global scale, anxiety disorders (in combination with depression) are estimated to cost the economy around $1 trillion annually, with projections suggesting this could rise to $6 trillion by 2030 (The Lancet Global, 2020).

Given these substantial personal and societal costs, there is an urgent need for more effective treatments for anxiety disorders. Currently, selective serotonin reuptake inhibitors (SSRIs) and cognitive-behavioural therapy (CBT) are the first-line pharmacological and psychological treatments for anxiety, respectively. While meta-analyses support the efficacy of both approaches (Bandelow et al., 2015; Hofmann & Smits, 2008; Koen & Stein, 2011), treatment outcomes remain suboptimal. Up to 40% of patients do not respond to treatment, and among those who do respond, only about half achieve a reduction in symptoms to minimal levels (Bystritsky, 2006). Other effective medications such as benzodiazepines exist, but while they are highly effective for short-term anxiety relief, their use is limited by side effects and the risk of dependence with long-term use.

To make things worse, despite the pressing need for improved treatments, there has been little innovation in pharmacological interventions for anxiety over the past few decades. The discovery that the most commonly used SSRIs were effective for anxiety symptoms was made in the late 1980s and early 1990s. However, this finding was largely serendipitous rather than the product of focussed anxiety research, which highlights the challenges in effectively driving innovation in this field.

The lack of progress is not due to a failure of effort or investment. The pharmaceutical industry invested heavily in the development of antidepressants and anxiolytics in the late 20th century. However, clinical trials for psychiatric drugs (and central nervous system (CNS) drugs more broadly) are costly and have lower success rates compared to non-CNS drugs (Cook et al., 2014; Kesselheim et al., 2015). These challenges have led most of the pharmaceutical industry to reduce or completely shut down their research in this field (Hyman, 2012; Miller, 2010).

The difficulties in developing new, effective treatments for anxiety disorders warrant novel approaches to understanding and treating these conditions. In the rest of the chapter, I will explore some avenues to improve this: namely the role of cross-species translational research in anxiety, its limitations, and the potential of computational neuroscience to address these limitations. But first, a brief look at how the conceptualisation of anxiety has changed over the years will be useful in contextualising later ideas.

1.2 Defining anxiety: a brief history

Anxiety is a complex phenomenon, involving various responses of the mind and body to threat. Physiologically, anxiety can manifest through increased heart rate, sweating, trembling, and muscle tension. Cognitively, it could involve excessive worry, racing thoughts, and difficulty concentrating. On the broader scale of behaviour, anxiety frequently leads to avoidance of threatening situations or objects. This complex, multi-faceted nature of anxiety has posed a problem for clearly defining and classifying anxiety disorders.

Before proceeding, it is important to distinguish between fear and anxiety, as these terms are often used interchangeably but are thought to represent distinct psychological and neurobiological processes (Perusini & Fanselow, 2015). Fear is typically characterised as an immediate, intense emotional response to a specific, identifiable threat. It involves a set of automatic physiological reactions aimed at dealing with imminent danger. Anxiety, on the other hand, refers to a more diffuse, future-oriented state of apprehension about potential threats. While fear is usually short-lived and dissipates once the threat is gone, anxiety can persist in the absence of a clear, immediate danger. However, it is also worth noting that the fear-anxiety distinction is contentious (Daniel-Watanabe & Fletcher, 2021), although it is beyond the scope of this thesis to discuss further here.

Historically, what we now recognise as anxiety disorders were classified under the broader category of neuroses or psychoneurotic disorders, as outlined in the early editions of the Diagnostic and Statistical Manual of Mental Disorders (DSM-I, DSM-II; American Psychiatric Association, 1952, 1968). These early classifications viewed anxiety symptoms as conscious reactions to inner, subconscious conflicts: "The chief characteristic of these disorders is 'anxiety' which may be directly felt and expressed or which may be unconsciously and automatically controlled by the utilization of various psychological defence mechanisms (depression, conversion, displacement, etc.)... 'Anxiety' in psychoneurotic disorders is a danger signal felt and perceived by the conscious portion of the personality" (American Psychiatric Association, 1952, p.31). This approach, however, lacked objective measurement criteria, leading to a crisis of diagnostic reliability. What one clinician diagnosed as 'anxiety neurosis', another might label 'obsessive-compulsive neurosis' (Wilson, 1993).

In response to these problems of reliability, the DSM-III (American Psychiatric Association, 1980) pivoted to a more symptom-focused approach, recommending diagnoses based on standardised criteria relating to observable symptoms. For instance, the DSM-III introduced specific duration requirements for diagnoses, such as symptoms persisting for at least six months in the case of generalised anxiety disorder. This revision aimed to develop a reliable system of classification, ensuring that clinicians would independently produce consistent diagnoses given a patient's history of symptoms. It was at this time that anxiety was classified as a unique category of psychiatric disorder, including diagnoses such as generalised anxiety disorder, social anxiety disorder, and panic disorder.

While this symptom-based approach improved diagnostic reliability, it faced criticism for lacking validity. The categorical nature of DSM diagnoses failed to capture the dimensional nature of anxiety symptoms (Shear et al., 2007). Moreover, the arbitrary thresholds for symptom presence/absence and the multiple possible combinations of criteria to qualify for an anxiety disorder led to high heterogeneity within diagnostic categories (Galatzer-Levy & Bryant, 2013). These issues with the DSM classification system point to a fundamental problem of validity in our conceptualisation of anxiety disorders. As the former director of the U.S. National Institute of Mental Health (NIMH), Steven Hyman argued, the DSM categories represented an 'epistemic prison' that limited progress in understanding and treating psychiatric disorders (Hyman, 2010).

In response to the limitations of categorical diagnostic approaches, there has been a shift towards more mechanistic, cognitively informed approaches to understanding anxiety. This perspective recognises that anxiety disorders likely arise from varying degrees of disruption in normative cognitive mechanisms, suggesting a dimensional relationship between adaptive and pathological anxiety (Dillon et al., 2014). By focussing on specific cognitive processes, such as attention, memory, and decision-making, the renewed aim was to identify the psychological mechanisms that may be disrupted across various anxiety disorders.

This shift towards more mechanistic, cognitively informed approaches gained momentum with the introduction of the Research Domain Criteria (RDoC) initiative (Insel et al., 2010). The RDoC initiative aimed to develop new ways of classifying mental disorders based on dimensions of observable behaviour and neurobiological measures, rather than traditional diagnostic categories. This framework provided a roadmap for integrating cognitive perspectives into translational anxiety research by focussing on processes that can be measured across multiple levels of analysis - from genes to neural circuits to behaviour. The RDoC approach represented a shift in psychiatric research, encouraging the field to look beyond symptom-based diagnoses and instead focus on understanding

the underlying mechanisms of mental disorders. These approaches have already yielded important insights into anxiety disorders. For example, research has identified attentional biases towards threat as a key feature across various anxiety disorders (Bar-Haim et al., 2007). Similarly, studies have suggested that threat learning and extinction processes are altered in anxiety (Duits et al., 2015), which have provided targets for exposure-based therapies (Craske et al., 2014).

It should be noted that the RDoC framework has faced criticism. Some argue that it overemphasises biological processes at the expense of psychosocial factors, and point to challenges in measurement and interpretation of neurobiological data (Lilienfeld, 2014). Alternative approaches, such as the Hierarchical Taxonomy of Psychopathology (HiTOP), have been proposed to address these limitations by focussing more on how psychopathological symptoms are organised empirically, compared to the RDoC's emphasis on genetic and neurobiological factors (Kotov et al., 2017). These approaches represent ongoing efforts to develop more comprehensive frameworks for understanding and researching mental health conditions.

The shift towards more mechanistic, cognitively informed approaches to anxiety presents new opportunities for research and treatment development. Focussing on specific, quantifiable cognitive processes could allow us to develop more precise operational definitions of anxiety, with far-reaching implications for how we describe, measure and ultimately treat anxiety in humans, but also how we think about studying anxiety in non-human animals.

1.3 From bench to bedside: cross-species translational research

1.3.1 Why we need animal models of anxiety

Non-human animal models of human diseases have long played a crucial role in medical innovation. For anxiety disorders and other diseases of the central nervous system, animal models are particularly important due to the complexity and inaccessibility of the brain, the presumed central organ in the aetiology of anxiety. While neuroimaging techniques such as magnetic resonance imaging and electro-/magnetoencephalography have made tremendous progress in measuring the live human brain with increasing temporal and spatial resolution, the most powerful imaging methods and many causal manipulation techniques are only possible in non-human species. Cross-species translational research serves two essential purposes in the field of anxiety studies. Firstly, it contributes to the theoretical development of anxiety aetiology by allowing us to investigate underlying neurobiological mechanisms in ways that are not feasible in human subjects. Secondly, animal models play a crucial role in the drug development process, providing a platform for initial testing of efficacy and safety of novel anxiolytic compounds before advancing to human trials. The most common species used as animal models of anxiety are rats and mice, due to their relatively

short lifespan, ease of genetic manipulation, and similarities in brain structure and function to humans. Therefore, the following discussion will focus primarily on rodent models of anxiety.

Before proceeding, it is important to clarify some terminology. In the literature, the terms 'translational model' and 'translational test' are often used interchangeably, but they have distinct meanings that are important for understanding the field. I will use the term ‘translational test' to refer to a specific experimental procedure or setup designed to measure behavioural or physiological responses of animals in potentially anxiogenic contexts. These are essentially specific instruments for quantifying anxiety-related phenomena. On the other hand, I will refer to ‘translational models’ as a broader concept that aligns more with a theory of how to measure anxiety in animals. A translational model might encompass many tests and provides a framework for interpreting the results of these tests in the context of anxiety.

Note on terminology. Usage of the word 'model', Part 1: animal models and computational models

At this point, I have discussed how animal models and translational models are used to study anxiety-like behaviours and test potential treatments across species. Later on in the thesis, I will also use the term 'model' to refer to mathematical representations of cognitive processes and behaviours. To avoid confusion, I will make every effort to disambiguate these terms where possible, but making their difference explicit here will hopefully aid readability.

1. Animal/translational models: These refer to behavioural setups and experimental paradigms (such as the elevated plus maze) designed to measure anxiety-like behaviours in non-human animals. These models provide a framework for interpreting results in the context of anxiety. They may encompass multiple specific behavioural tests.

2. Computational models: These are mathematical representations of cognitive processes or behaviours. In this thesis, computational models refer to algorithms and mathematical frameworks used to simulate or analyse decision-making, learning, and behaviour related to anxiety.

1.3.2 How is 'anxiety' modelled in non-human animals?

Numerous translational tests have been developed to measure anxiety in non-human animals. These can be broadly categorised into conditioned and unconditioned tests. A common feature across both types of tests is that they involve approach-avoidance conflicts, where the tests are structured such that key behavioural responses are implicated with both rewarding/appetitive features as well as aversive/punitive features, simultaneously. It is important to note that since we cannot model verbal aspects of anxiety like worries in animals, the aim has been to model behaviours observed in clinical anxiety, particularly avoidance behaviours which are key symptoms of anxiety disorders.

Conditioned tests

Conditioned tests, as their name implies, rely on either Pavlovian or instrumental conditioning (i.e. learning) to probe a subject's responses to rewards and punishments; in the field of anxiety research, a common form of instrumental test is known as the 'operant conflict test'. The first established form of this test was the Geller-Seifter test (Geller et al., 1962). The test involves a training phase in which food-deprived rodents are trained to press a lever that initially yields a food pellet as a reward. Once this association is learned, the test progresses to a phase where the same lever is paired with both a food pellet and an electric shock. The test measures 'anxiety-like avoidance' in this phase by measuring the number and latencies of lever presses, where fewer lever presses and longer latencies are thought to reflect more anxious behaviour. The test was later simplified into what is known as the Vogel conflict test (Vogel et al., 1971), which uses licks at a water spout to yield sugar-sweetened water instead of lever presses for food pellets as the 'approach' behaviour. In both of these tests, the key elements of the experimental design are that discrete actions can be simultaneously associated with both a rewarding outcome and an aversive one, thus inducing a conflict.

The Geller-Seifter and Vogel conflict tests were popular in early translational research due to their sensitivities to the effects of benzodiazepines. However, since other measures have been developed that require less intensive training, they have fallen out of favour (Griebel & Holmes, 2013).

It is worth noting that while Pavlovian fear conditioning paradigms are sometimes discussed in the context of anxiety research, they typically do not fall under the category of anxiety tests *per se*. It is commonly cited that fear conditioning primarily measures acute fear responses to specific cues or contexts, and involves learning associations between neutral stimuli and aversive outcomes, which is more closely related to fear acquisition and extinction processes. However, while this distinction between fear conditioning and anxiety tests is commonly made in the literature, I should note that the boundaries between these paradigms are not always clear-cut. Tests such as the Vogel conflict test, which I introduce here as an anxiety test, also involve elements of learned associations and responses to specific threats, similar to fear conditioning paradigms. The overlap between these paradigms highlights why the fear-anxiety distinction is contentious, as noted earlier, but again, for the purposes of this thesis, I will focus primarily on tests traditionally classified as 'anxiety tests' in the preclinical literature.

Unconditioned tests

Unconditioned tests have become increasingly popular due to their reliance on innate responses, eliminating the need for extensive training. These tests exploit the natural conflict between an animal's drive to explore novel environments and its instinct to avoid potentially threatening

situations. While numerous variants of unconditioned tests exist (Campos et al., 2013), three have come to dominate the literature over the past two decades: the elevated plus-maze, the open-field test, and the light/dark box.

The elevated plus-maze (Handley & Mithani, 1984) accounts for approximately half of all anxiety studies in non-human animals (Griebel & Holmes, 2013). The apparatus consists of a raised, plus-shaped platform where two opposing arms offer shelter with enclosing barriers, while the other two remain open and exposed. Rodents are placed onto the platform for a certain amount of time, during which they are free to move around the environment. Although a variety of behavioural readouts exist, the most common involve the time which rodents spend on the open arms vs the closed arms of the maze, where avoidance of the exposed, open arms is thought to indicate anxiety-like behaviour.

Building on similar principles, the open-field test (Hall, 1934) and light/dark box (Crawley & Goodwin, 1980) involve similar contrasts between safe and potentially dangerous locations. The open-field test involves observing the free movement of rodents in a large, walled enclosure, where it is commonly found that they avoid the central area of the field, a behaviour referred to as *thigmotaxis*, presumably due to its lack of protection from predators (relative to the walled perimeter which provides more protection). On the other hand, the light/dark box relies on rodents' aversion to bright light, by observing their movement in a spatial environment consisting of an aversive illuminated compartment, and a safer darkened compartment.

1.4 The translational gap

Despite the widespread use and initial promise of animal models in anxiety research, a translational gap has emerged between preclinical findings and clinical outcomes, which raises questions about the validity and utility of current animal models in anxiety research and drug development.

Initially, animal models of anxiety showed great promise as tools for drug development. Many of the behavioural tests described above demonstrated sensitivity to the effects of benzodiazepines, the first major pharmacological breakthrough in anxiety treatment. For example, these drugs reduced anxiety-like behaviour in rodents in tests such as the Geller-Seifter conflict test (Green & Hodges, 1991). These findings suggested that the tests truly tapped into something related to human anxiety, and so the tests could be, and indeed were, used to screen for other benzodiazepines which also had anxiolytic effects.

Unfortunately, this momentum did not last. The use of animal models of anxiety saw a surge from the 1980s onwards, but without a proportional yield in novel drug discoveries. The disconnect

between basic and clinical research became particularly evident with the development of non-benzodiazepine anxiolytics, such as serotonergic agents. Despite proving effective at reducing anxiety in humans, these newer drugs often showed null or inconsistent effects in animal testing (Griebel, 1995) – for example, a review on the effects of different anxiolytic drugs on the open field test suggested that only benzodiazepines and 5-HT1A receptor agonists reduce anxiety-like behaviour in the test, whereas other drugs (that are known to be anxiolytic in humans) are ineffective (Prut & Belzung, 2003). It thus became evident that the behavioural tests discussed above were primarily sensitive to the effects of benzodiazepines, but less reliable in predicting the efficacy of other classes of anxiolytic drugs.

This translational gap is now starkly illustrated by the high failure rate of candidate anxiolytic drugs in clinical trials (Stewart et al., 2015). Despite showing promise in preclinical studies and early clinical phases, many anxiolytic drugs fail at Phase III trials (large-scale studies involving hundreds to thousands of patients that aim to conclusively demonstrate a drug's safety and efficacy before seeking regulatory approval), often due to lack of efficacy in humans. This pattern suggests a fundamental issue with how we model and test anxiety in preclinical stages.

The persistent challenge in developing novel, effective anxiolytics may be partly attributed to this translational gap. The high attrition rate in late-stage clinical trials for anxiety medications points to the limitations of our current preclinical (i.e. animal) models in predicting human responses. This discrepancy raises pressing questions about the validity of our animal models: Are we truly modelling human anxiety, or are we instead measuring behaviours that are only tangentially related to it? These challenges suggest that we need to critically examine the foundations of our translational research. How can we ensure that our translational tests are truly representative of human anxiety? What criteria should we use to evaluate the validity of these models? Addressing these questions is crucial for bridging the gap between preclinical *promises* and real clinical outcomes.

1.5 What makes a valid translational test?

The basic premise of a valid translational test of anxiety is that it replicates relevant aspects of human anxiety, especially aspects which relate to anxiety disorders. Over the years, the field has converged on three key criteria for evaluating the validity of a test (Willner, 1986):

1. Face validity: Upholding face validity means maximising the outward similarity between the animal test and the human disease. While crucial for validating a disease model, face validity may be a red herring at times if observable behaviour is supported by different mechanisms across species.

2. Predictive validity: This criterion is met when interventions which are clinically effective in humans demonstrate similar effects in the translational test. However, achieving predictive validity is often difficult due to incomplete correlation between animal and human disease mechanisms, and the potential inability of approved human drugs to be active in the animal model species.
3. Construct validity: The final criterion calls for translational tests to engage the mechanisms relevant to the human disease – in other words, the target construct being measured should be identical. These mechanisms can include those at the molecular, cellular and/or cognitive levels. As noted earlier, the inaccessibility of the human brain has meant that these mechanisms, in the context of anxiety but also other psychiatric disorders, remain largely unknown, making the assessment of construct validity difficult.

While these criteria have been useful in guiding translational research, the persistent challenges in anxiety drug development suggest that they may be insufficient. To make progress, new approaches to validating and refining translational models may be fruitful. As we will discuss in the next sections, emerging perspectives from the field of computational neuroscience offer promising avenues for enhancing our understanding of validity in translational anxiety research.

1.6 Anxiety as a problem of information processing

While cross-species translational approaches have allowed us to investigate the underlying neurobiology and pathophysiology of anxiety, another field of research has focused instead on improving the understanding of the cognitive mechanisms of behaviour. Computational neuroscience starts from the premise that the brain's primary function is to process information, and that these information-processing mechanisms can be approximated with precise mathematical equations. Computational approaches allow us to model and infer latent, or hidden, cognitive processes that are not directly observable but are hypothesised to underlie behaviour. These latent mechanisms represent internal mental states, decision-making processes, or information processing steps that occur between environmental inputs and behavioural outputs. Adopting this perspective shifts our focus from neurobiological explanations to considering the computational and information-processing mechanisms that drive anxious thoughts and behaviours.

1.6.1 Computational psychiatry

If computational neuroscience is the study of the brain as an information-processing organ, with the goal of mapping its computational functions to their neural mechanisms, computational psychiatry (Huys et al., 2016; Montague et al., 2012) is the study of how information-processing can go awry, resulting in mental ill-health. In other words, psychiatric symptoms are interpreted as the product of

'abnormal' computations. It is worth noting that this description refers specifically to theory-driven computational psychiatry, which is commonly distinguished from data-driven approaches that involve the use of exploratory machine learning techniques to analyse high-dimensional data in psychiatry.

There are two powerful advantages of adopting mathematics as the language of describing neurocognitive processes. One is that it forces precision. Take, for example, the hypothesis that an organism's decision to escape a potentially dangerous situation depends on both the organism's state anxiety and how likely it thinks something threatening will appear. As stated in natural language, this hypothesis is ambiguous. Does state anxiety have a larger influence than the perceived likelihood of threat? Do these factors interact, and if so, how? At what point does the combination of anxiety and perceived threat trigger an escape response? These questions are left unanswered.

However, we can express this hypothesis much more precisely using mathematical statements:

$$P(escape) = f(A, T)$$

Where P(escape) is the probability of escaping, A represents state anxiety, T represents the perceived likelihood of threat, and f is a function that defines their relationship. We might further specify this function as:

$$P(escape) = \frac{1}{1+e^{\beta_A \cdot A + \beta_T \cdot T + C}}$$

This logistic function provides a specific mathematical relationship between anxiety, perceived threat, and the probability of escaping. The β coefficients represent the relative influence of each factor, while C is a constant that affects the baseline probability of escape. This mathematical formulation makes several aspects of the hypothesis explicit:

1. Both anxiety and perceived threat contribute positively to the probability of escaping.

2. Their effects are additive, implying that high levels of one factor can compensate for low levels of the other.

3. The logistic function ensures that the probability of escape is always between 0 and 1, and it allows for non-linear effects as the combined influence of A and T increases.

As we can see, the hypothesis is more precise, generating a series of predictions about future observations in relevant contexts. Importantly, this makes it more falsifiable.

The second advantage of computational models of cognition is that, as *in silico* brains, they can generate data themselves. This is why theory-driven computational psychiatry is often referred to as

relying on ‘generative’ models. Although he likely intended its interpretation in a different way, the late Richard Feynman's quote left on his blackboard puts it well: "What I cannot create, I do not understand". By generating behaviour in accordance with one's best guess of how the brain solves a certain problem (such as escape), and comparing it to empirical natural behaviour, we might be able to move closer towards understanding the cognitive-computational mechanisms of anxiety.

1.6.2 Reinforcement learning

Among the computational frameworks used to study anxiety-related learning and decision-making processes, reinforcement learning (Sutton & Barto, 2018) has emerged as one of the most influential. At its core, reinforcement learning models describe how an agent (i.e., an organism) interacts with a potentially uncertain environment and uses knowledge based on past interactions with the environment to act in a value-efficient manner. Reinforcement learning encompasses a family of models, broadly categorised into model-free approaches, which learn directly from experience without an explicit model of the environment, and model-based approaches, which utilise an internal model of the environment to plan and make decisions. In the context of anxiety research, reinforcement learning models offer a powerful tool for quantifying how individuals act in contexts with potential threat, and how the mechanisms driving actions might be altered in anxiety disorders.

Note on terminology. Usage of the word 'model', Part 2: model-free and model-based reinforcement learning

Throughout this thesis, I will use the terms "model-free" and "model-based" specifically in the context of reinforcement learning algorithms. These usages are distinct from other appearances of the word "model" in this thesis, such as in " animal model" or "computational model". When the terms "model-free" or "model-based" appear, they will always refer to these specific types of reinforcement learning algorithms.

"Model-free" reinforcement learning refers to algorithms that learn directly from experience without explicitly representing the environment's dynamics. These methods update value estimates based on observed rewards without predicting future states.

"Model-based" reinforcement learning refers to algorithms that learn and use an internal representation (i.e. a ‘model’) of the environment to plan and make decisions. These methods explicitly represent how actions lead to future states and outcomes.

Model-free reinforcement learning

Model-free reinforcement learning provides a simple yet powerful framework for understanding how organisms learn to seek rewards and avoid punishments through trial-and-error experiences. This approach is particularly relevant to anxiety research as it can account for both active and passive avoidance behaviours. Active avoidance involves taking a specific action to prevent or escape from an aversive stimulus, such as pressing a lever to avoid a shock. Passive avoidance, on the other hand,

involves refraining from an action to avoid punishment, such as not entering an area associated with danger. Both types of avoidance are common in anxiety disorders, with active avoidance often manifesting as escape behaviours and passive avoidance as behavioural inhibition or withdrawal.

At the core of model-free reinforcement learning is the concept of learning from prediction errors - the discrepancy between expected and actual outcomes. This process captures how individuals incrementally refine their behaviour based on the observed outcomes of their actions, with ample evidence that it is supported by dopaminergic signalling in regions such as the striatum and frontal cortex (Schultz, 1998; Schultz et al., 1997). In the context of anxiety, perhaps the most interesting application of this approach is how it is used to understand how individuals learn to avoid negative events from occurring (e.g. Eldar et al., 2016; Li et al., 2011; Seymour et al., 2004). For example, if action X (like entering a crowded space) unexpectedly leads to a panic attack, the prediction error would be negative – the outcome was worse than expected. Consequently, an individual might update their expected value of performing action X such that they are more likely to expect a panic attack after action X. Over time, this could result in the individual avoiding action X (i.e. crowded spaces) altogether. This simple learning mechanism is thus proposed to explain how maladaptive avoidance behaviours, a hallmark of clinical anxiety, can develop and persist over time (Maia, 2010).

Computational models, including those of model-free reinforcement learning, typically include free parameters that can depend on the individual, and represent individual differences in learning and decision-making mechanisms. These parameters allow for the quantification of specific aspects of information processing. One key parameter in model-free reinforcement learning is the learning rate, which determines how quickly an individual updates their expectations based on new information. Previous research has suggested that individual differences in learning rates correlate with clinical anxiety in humans (Yamamori & Robinson, 2023). For example, one previous study found that individuals with anxiety disorders showed higher punishment learning rates compared to healthy controls (Aylward et al., 2019), in other words, that clinically anxious individuals tended to update their expectations about negative outcomes more rapidly. A meta-analysis of 27 studies has corroborated these findings, suggesting that pathologically anxious individuals consistently show higher punishment learning rates (Pike & Robinson, 2022).

Other studies point to more complex associations between anxiety and learning from negative events. For example, one study found that trait cognitive anxiety (reflecting cognitive symptoms such as worries) is associated with elevated threat learning rates, in other words, a bias for learning from threatening outcomes, while somatic anxiety (which reflect more physiological symptoms) was associated with elevated safety learning rates, i.e. a bias for learning from ‘safety’ outcomes, or a

lack of threat (Wise & Dolan, 2020). This double dissociation suggests that different forms of anxiety may be characterised by distinct learning biases. Other study has shown that the relationship between anxiety and learning rates can depend on the temporal properties of the environment (Browning et al., 2015). Specifically, the authors found that individuals with high trait anxiety show less flexibility compared to less anxious individuals in optimally adjusting their learning rates when the action-outcome associations are more unstable over time (Browning et al., 2015). This impaired ability to gauge when to update beliefs about action-outcome contingencies could lead to poor decision-making in anxious individuals and potentially exacerbate anxiety symptoms.

Model-based reinforcement learning

While model-free reinforcement learning provides a powerful framework for understanding relatively simple behaviours, model-based reinforcement learning offers a complementary approach that can capture more complex decision-making processes. Model-based reinforcement learning endows agents with knowledge of the environment's structure, allowing for more flexible evaluation of action consequences, though at the cost of increased computational demands.

The core principle of model-based reinforcement learning is that the agent maintains an internal representation of their environment. This internal representation encapsulates the agent's understanding of the causal relationships between the environment's states, actions, and outcomes. Unlike model-free reinforcement learning, which learn directly from trial-and-error experiences, this causality allows model-based agents to perform more flexible decision-making. One key advantage of model-based reinforcement learning methods lies in their adaptability to environmental changes without necessitating new experiences. For instance, if the environment's structure changes explicitly, a model-based agent can update its internal representation and adjust its behaviour accordingly. In contrast, model-free reinforcement learning methods would require accumulating new experiences through trial and error to adapt to such changes as it does not utilise an internal representation of the world. In the context of anxiety, model-based reinforcement learning might be particularly relevant for understanding complex decision-making processes such as how individuals plan over space and time to avoid potential threats.

The two-step task (Daw et al., 2011) is the classical paradigm for measuring model-free and model-based reinforcement learning in human decision-making. In this task, participants make two sequential choices to reach a final reward, with probabilistic transitions between stages. The task is designed to dissociate model-based and model-free reinforcement learning strategies by examining how participants adjust their first-stage choices following rewards or punishments in the previous trial. Model-free learners tend to repeat first-stage choices that led to rewards, regardless of whether

the transition to the second stage was expected or unexpected. In contrast, model-based learners take into account the transition structure, adjusting their choices based on both the outcome and whether the previous trial involved an expected or unexpected transition. This distinction allows for inference about the degree to which participants are employing model-based versus model-free reinforcement learning strategies.

When applied to understand threat avoidance, studies using the two-step task have shown that humans tend to use both model-free and model-based reinforcement learning strategies (Sebold et al., 2019; Wang et al., 2018) as we also do for reward pursuit (Daw et al., 2011). Some studies have suggested that the degree to which one engages in model-based control can depend on anxiety, for example, individuals appear to favour model-free reinforcement learning over model-based ones in aversive environments (when viewing aversive images) (Sebold et al., 2019) and under social stress (Park et al., 2017), potentially indicating that anxiety leads to preferences for simpler, but less flexible behavioural responses to the environment (i.e. 'quick and dirty solutions'). However, other research has shown that experimentally induced anxiety (via hypercapnic gas) and naturalistic anxiety (panic attacks/life stress) do not appear to impact model-based reinforcement learning (Gillan et al., 2021). These findings suggest a complex relationship between anxiety and model-based reinforcement learning, at least through the two-step task.

While the two-step task has been influential in studying the relative contributions of model-based versus model-free reinforcement learning, it remains a relatively simple task, with its name aptly reflecting its limited planning horizon of just two steps. In contrast, many real-world situations, particularly those relevant to anxiety, require planning over much broader temporal and spatial horizons. For instance, navigating through a city to avoid potential threats or planning a series of social interactions for someone with social anxiety involves considering many more steps and possible outcomes. However, applying model-based reinforcement learning frameworks to these more complex, spatially oriented scenarios has been less common, probably due to the increased complexity when modelling behaviour in such rich environments. Nonetheless, extending model-based reinforcement learning to these more naturalistic contexts represents an important opportunity in understanding how anxiety influences decision-making and behaviour in real-world situations.

Bach et al. (2014) developed a more naturalistic task that captures complex spatial behaviour under threat. In the task, participants collected tokens in a two-dimensional spatial environment while avoiding a virtual predator. This task design allowed for rich behavioural measures relevant to model-based control, such as the distance maintained from the predator or the time spent in safe areas of

space, and how they adapt to varying levels of threat. These measures provide a more comprehensive view of decision-making than the binary choices in the two-step task. Although Bach et al. did not go as far as specifying a generative computational model, their approach is a step towards capturing the complex nature of anxiety-related behaviour in environmentally rich contexts.

More recently, Zorowitz et al. (2020) proposed a computational framework that extends the concept of model-based reinforcement learning to spatial environments, to understand how anxiety might influence navigation and avoidance behaviours in broad state spaces. In their computational model, agents navigate through a spatial environment where some locations are associated with potential threats. Their key insight was that in optimal decision-making, the potential for future avoidance should be protective: if danger can be reliably avoided later, it poses less of a threat now. However, anxious individuals may have a pessimistic belief about their ability to avoid future threats, leading to excessive avoidance behaviours in the present, even when the current state is far removed from the actual source of danger. Simulation analyses of agents navigating through environments with potential threats demonstrated how such pessimistic biases can lead to maladaptive avoidance behaviour, even if threats could be straightforwardly avoided. This work highlights the potential for applying model-based reinforcement learning to more complex, naturalistic tasks.

1.7 Bridging the translational gap: integrating translational and computational approaches

So far, we have discussed the potential of animal models of anxiety for advancing our understanding of the underlying pathophysiology of anxiety, and as screening tools for anxiolytic drug development. We have also considered computational approaches as powerful techniques that can help us understand the latent, unobservable mechanisms that drive observable behaviour. Now, we will explore how integrating translational research with computational neuroscience methods may provide a bridge to cross the translational gap. Specifically, we will explore the implications of thinking computationally about translational anxiety models, and how this can help us improve their validity.

1.7.1 Construct validity refined: computational validity

Recent ideas in computational neuroscience offer a new perspective on how we can assess the validity of animal models in psychiatry. Redish et al. (2022) propose the concept of *computational validity* as a new criterion for evaluating translational models. This approach extends beyond traditional notions of construct validity by focussing on the underlying computational processes that drive behaviour, rather than solely on observable outcomes.

Computational validity posits that a valid translational model of anxiety should not only produce similar observable behaviours across species but should also engage analogous information processing mechanisms. In fact, these mechanisms should be the primary target of translation, rather than the observable behaviours. Our current definitions of the constructs measured by the translational tests of anxiety are vague, ranging from the broadest terms like 'anxiety', to 'anxiety-like behaviour' to more focused terms such as thigmotaxis (wall-following behaviour) for the open-field test, but rarely extend to the computational mechanisms of behaviour. This shift in focus from behaviour to mechanisms is important because the same mechanisms elicited in non-human animals may not necessarily result in behaviours that resemble human symptoms. Conversely, animal behaviour that is outwardly similar to human symptoms may arise from entirely different computational mechanisms.

The key principle of computational validity is that it aims for a more precise mapping of disease-relevant *mechanisms* across species. For anxiety research, adopting a computational validity framework could address some of the limitations of current animal models. By modelling the computational processes involved in translational tests of anxiety, we may be better placed to compare and translate findings across species.

1.7.2 Computations enable predictions

As discussed earlier, traditional translational approaches have struggled to predict changes in human anxiety based on preclinical data. The computational validity framework could provide a means to improve not only the construct validity, but also the predictive validity of translational models of anxiety.

Historically, the predictive validity of a test has been evaluated by considering whether the effects of a drug or intervention on human symptoms can be inferred by its effects on animal *behaviour*. But this approach neglects a critical intermediate step: the mechanism of action. A computational perspective enables us to evaluate predictive validity differently. Since fundamental cross-species ethological differences mean that similar neuro-computational mechanisms, or an intervention's effect on these mechanisms, could manifest as vastly different behaviours across species, relying solely on behavioural markers to infer whether findings from preclinical studies will translate to humans is inherently unreliable. For example, consider the light-dark box test. Elevated threat sensitivity, a computational mechanism relevant to anxiety, manifests as preference for dark areas in mice but preference for brightly lit areas in humans. This contrast in behaviour stems from species-specific evolutionary adaptations: mice, as nocturnal prey animals, associate darkness with safety, while humans, as diurnal animals, generally perceive well-lit environments as safer. Thus, an

anxiolytic drug that reduces threat sensitivity might decrease dark preference in mice but decrease bright area preference in humans - opposite behaviours driven by the same computational change.

A relevant advantage of computational modelling is that it allows for the "unmixing" of multiple computational processes that simultaneously contribute to observable behaviour. In any given behavioural task, especially complex ones, multiple processes often contribute to a single behavioural readout such as response accuracy, making it difficult to determine which specific process is affected by an intervention based on the behavioural readout alone. Computational models can help to disentangle these processes, providing a more nuanced understanding of how potential treatments affect cognition and anxiety.

Understanding which computational mechanisms mediate the effects of interventions on behaviour has the potential to reduce noise in translating findings across species. Take for example, a case where a preclinical study finds that a novel, potentially anxiolytic drug reduced avoidance behaviour in a behavioural test, but computational analyses showed that the effect was mediated by a computational parameter that has not reliably been associated with anxiety symptoms in humans, and conversely, the drug had no effect on a parameter with strong associations with human anxiety (say the punishment learning rate). A traditional interpretation, i.e. one based on the behavioural readout alone, would be that the drug could be effective in humans; but the computational results indicate that this may be a red herring.

On the other hand, consider a preclinical study that instead found that the drug had no effect on avoidance, as measured by traditional behavioural indices, but the computational analyses revealed that the drug dampened a computational mechanism that is thought to be exaggerated in human anxiety. Predicting a null effect on human anxiety based on these behavioural results might lead to a foregone opportunity, despite the promising effects seen via the computational analyses. This scenario is an extreme and hypothetical example of the potential disconnect between an intervention’s effect on behaviour and its mechanisms, but as noted earlier, the multiple parallel computational processes that govern behaviour can weaken the relationships between individual mechanisms and behavioural readouts. Therefore, taking both forms of evidence into account has the potential to strengthen the evaluation of predictive validity.

Experiments using anxiety manipulations, that is, experimental interventions designed to directly increase or decrease anxiety levels, in both animal and human samples, are necessary to investigate and validate these computational mechanisms. By observing how anxiogenic and anxiolytic interventions affect specific computational parameters across species, it may be possible to identify certain mechanisms that are robustly related to anxiety and most likely to translate between animal

models and human patients. This approach creates a more robust link between what we can measure in animal models and what we predict in relation to human treatment outcomes, potentially reducing the translational gap that has hindered anxiety research.

1.7.3 Examples of combined translational and computational approaches

To illustrate the potential of integrating computational approaches with translational research, I will examine some examples where similar computational models have provided parsimonious explanations of both human and animal behaviour in anxiety-related tasks.

Affectively biased evidence accumulation

One example of a behavioural anxiety task that has been used across species with evidence of similar cross-species computational mechanisms is one which involves learning about perceptually ambiguous stimuli that are associated with asymmetrical rewards (Aylward et al., 2020; Hales et al., 2016).

In the original rodent version of this task (Hales et al., 2016), rats learn to discriminate between high- or low-frequency tones using two levers, one associated with the high tone and the other with the low tone. Importantly, one tone is deterministically paired with more reward (food pellets) compared to the other, given a correct response. Once these associated are learned, the rats complete the test phase, where they are also presented with an ambiguous mid-tone which is exactly equidistant between the unambiguous stimuli, and can be reinforced with both high and low reward outcomes. If the animals respond with the lever associated with the large-reward/tone, they get a large reward exactly 50% of the time, and the converse is true for the small-reward lever. Therefore, the optimal response to this ambiguous stimulus is to press the lever associated with the large reward more often, in order to maximise reward. However, in practice, some subjects showed a 'pessimistic' response, by more frequently choosing the small-reward lever, and this was especially the case after the rats were administered an anxiogenic drug or subjected to chronic stress, suggesting that these interventions induce a negative affective bias in decision-making.

Behaviour on this task was modelled with a computational model, specifically the drift-diffusion model, which parameterises decision-making as a process of noisy accumulation of evidence (Ratcliff, 1978) and is able to model the decisions and reaction times to stimuli on two-alternative forced choice tasks. The negative choice bias following anxiogenic drug administration and chronic stress in rats was explained via the model by increased 'boundary separation' parameters (more information required in order to reach a decision), whereas reduced 'drift rate' (rate of information accumulation) parameters were seen following the pharmacological manipulation. Importantly, this same computational framework has been successfully applied to human versions of the task. Aylward et al.

(2020) translated the rodent paradigm into a human task and found that induced anxiety (through threat of shock) led to increased negative bias in ambiguous cue interpretation, mirroring the effects seen in stressed rats. Moreover, computational modelling revealed that this bias was also driven by changes in drift rates, indicating a common computational mechanism of anxiety across species. The congruent pattern of findings across these pair of studies provides an initial validation of the relevance of this task and how its computational mechanisms respond to external stressors across species.

Approach-avoidance conflict

Another example comes from attempts to explicitly translate classical animal tests of anxiety for human applications. Aupperle et al. (2011) developed an approach-avoidance conflict task aimed at emulating conditioned tests of anxiety such as the Geller-Seifter and Vogel conflict tests. The important decisions in the task involve a choice between i) points-based reward and the presentation of a negatively valenced affective image and sound (i.e., an affective punishment), or ii) no points/reward and the presentation of a positively valenced image and sound. These options create a conflict between approach (accrue points) and avoidance (avoid affective punishment). Studies using this task have found that avoidance behaviour is significantly correlated with anxiety experienced during the task, although not with clinical measures of anxiety (Aupperle et al., 2015; Aupperle et al., 2011) such as the State-Trait Anxiety Inventory (Spielberger et al., 1971) or the Behavioural Inhibition Scale (Carver & White, 1994).

A more recent study developed a computational model of how humans may solve this task (Smith et al., 2021), building on a Bayesian computational model of decision-making (active inference; Friston et al., 2017). The computational model formalised the approach-avoidance conflict task as a Markov decision process, in which participants choose actions (runway positions) based on their beliefs about their preferences for different outcomes. The model included two key parameters. The 'emotional conflict' parameter quantified the relative aversiveness of negative stimuli compared to the reward value of points, whereas the 'decision uncertainty' parameter reflected the inverse of expected policy precision -- in other words, the confidence about which outcome was more preferable, with higher decision uncertainty indicating less confidence in decision-making, leading to more variable choices. The model therefore described the process of balancing the pursuit of preferred outcomes with the consistency of decision-making. In keeping with the notion of the task as measuring anxiety-related behaviour, state anxiety during the task was correlated with greater emotional conflict, meaning that those who found the aversive outcomes more unpleasant were more anxious during the task. Importantly, this computational analysis revealed that the cognitive

processes most relevant to anxiety in this task are those relating to valuation of positive vs negative outcomes, rather than uncertainty.

While this particular study focused on human participants, the computational framework it provides could potentially be applied to animal versions of approach-avoidance conflict tasks, offering a way to more directly compare anxiety-related processes across species.

1.7.4 Opportunities for future research

These examples of combined translational and computational approaches have shown initial promise that it is possible to measure the same computational processes in both humans and non-human animals, a potentially crucial step towards improving how well we can translate research across species. However, several opportunities remain for future research.

First, while previous work has aimed to translate the conditioned tests of anxiety for human use (Aupperle et al., 2011), the human version of the task differed from the rodent version in an important way: it did not involve learning about outcome probabilities. Specifically, in the Geller-Seifter and Vogel conflict tests, subjects learn that a behavioural response which was previously only associated with reward (i.e., food pellets or water), is now associated with both reward and punishment (i.e., an electric shock), and they must decide based on this newly learned association whether to engage or avoid the behavioural response in question. The human translated test instead offered participants an explicit choice between varying levels of reward and punishment without them needing to learn these combinations. This dissimilarity potentially limits comparisons of behaviour across species.

Second, unconditioned tests, which leverage the natural inclination of rodents to explore spatial environments, are much more popular in the contemporary animal literature. However, given the difficulty of modelling movement in two-dimensional space compared to binary decisions, little effort has been made to computationally model these behaviours. Previous studies have aimed to build computational models of the elevated plus maze, but these adopted data-driven (i.e., descriptive) approaches with the focus on reproducing behavioural patterns (Arantes et al., 2013), rather than modelling the underlying cognitive processes/mechanisms that generate behaviour in line with the theory-driven, 'generative' models which are the focus of much of contemporary computational psychiatry. Therefore, the precise computational mechanisms elicited by the classical unconditioned tests of anxiety is as yet unclear, and whether these mechanisms are common across species remains an open question.

1.8 Outstanding questions

The literature reviewed thus far highlights important challenges in the animal modelling of anxiety, but also points to opportunities for novel approaches. Computational perspectives could allow for a better understanding of the mechanisms engaged in behavioural tests that approximate anxiety. Rather than discarding the current battery of translational anxiety tests, this thesis aims to examine these established paradigms through a contemporary computational lens. Fittingly, this approach follows the instruction of Ivan Pavlov, the pioneer of classical conditioning: "If you want new ideas, read old books." In this spirit, I will apply modern computational techniques to longstanding experimental paradigms, and in doing so, aim to bridge the gap between animal and human measures of anxiety. My research will aim to address three over-arching questions that emerged from the literature to date.

1.8.1 Are classical preclinical tests of anxiety relevant to human anxiety?

Despite a large body of previous research that has used these classical preclinical tests to study anxiety-related behaviours over the past few decades (Griebel & Holmes, 2013), little attention has been paid to how anxious humans behave in analogous situations (Kirlic et al., 2017). Instead, the relevance of these tests to human anxiety has been assumed based on theoretical principles and face validity. As discussed above, it was only recently that one study took an explicitly forward-translational approach to empirically validate conditioned tests against self-reported anxiety in humans - this study did indeed show that behaviour in this paradigm correlates with human anxiety (Aupperle et al., 2011).

This approach has been followed by translations of unconditioned, exploration-based tests. For example, Walz et al. (2016) developed a human open field test using GPS tracking in a large outdoor area, finding that humans exhibit thigmotaxis similar to rodents, especially those with high levels of agoraphobic fear. Biedermann et al. (2017) translated the elevated plus-maze to humans using a combination of virtual reality and a real wooden maze structure. They observed that humans, like rodents, showed avoidance of open arms, and this behaviour was sensitive to anxiolytic and anxiogenic drugs. These studies demonstrate that unconditioned anxiety-like behaviours can be measured in humans using paradigms similar to those used in rodents.

The scarcity of empirical validation in translating animal models of anxiety to humans highlights several unanswered questions in the field. While classical preclinical tests of anxiety have been widely used in animal research, their relevance to human anxiety remains largely assumptive rather than empirically validated. There is a pressing need for more forward-translational studies, in other words, efforts to create human versions of established animal models. The few studies conducted to

date (Aupperle et al., 2011; Biedermann et al., 2017; Walz et al., 2016) have provided valuable initial insights, but more robust empirical validation of these tests against human anxiety data is needed.

1.8.2 What constructs do these tests measure?

Historically, preclinical tests of anxiety were designed to measure a unitary construct of 'anxiety' or 'anxiety-like behaviour'. However, as our understanding of anxiety has evolved, so too must our interpretation of what these tests actually measure. Rather than viewing them as direct analogues of human anxiety disorders, we should consider the specific cognitive and computational mechanisms they engage. For example, in the elevated plus maze, the construct being measured is not simply 'anxiety', but rather the balance between exploratory drive and threat avoidance. Computationally, this could be conceptualised as a decision-making process that weighs the potential rewards of exploration against the perceived risks of open spaces.

Similarly, in conditioned tests like the Geller-Seifter or Vogel conflict tests, the primary construct is not anxiety *per se*, but rather the cognitive processes involved in balancing reward (food/water) pursuit with punishment (shock) avoidance. From a computational perspective, these tests may be tapping into processes such as value computation and learning. Value computation refers to how animals calculate and compare the relative values of rewarding and punishing outcomes. Learning, on the other hand, reflect how animals update their expectations based on new information, which could potentially differ for positive and negative outcomes.

The constructs measured by the classical conditioned and unconditioned tests, in computational terms, is yet unclear. These behaviours are clearly problems of information processing – subjects must decide whether or not to approach a lever for food at risk of shock, or to explore a potentially dangerous region of space. Specifying these processes computationally may allow for better precision in terms of understanding what is being measured, and how behaviour might differ across species; any differences might indicate room for bringing the human and non-human animal behaviours closer together. This potential is worth exploring to lessen the translational gap.

1.8.3 How do these constructs respond to changes in human anxiety?

To establish the theoretical relevance of the computational mechanisms engaged in these translational tests, it is essential to validate them against experimentally induced changes in human anxiety. This process of validation is key to ensuring that the computational indices we derive from animal models are indeed relevant to human anxiety. By understanding how these computational constructs respond to controlled changes in human anxiety, we can more confidently interpret similar measures in animal studies and strengthen the translational bridge between preclinical and clinical research.

To test these relationships causally, human neuroscientists have typically relied on two main approaches: pharmacological interventions and threat-of-shock paradigms. These methods enable the controlled manipulation of anxiety levels in humans, allowing for the observation of subsequent changes in both observable behaviour and underlying computational parameters. This approach provides a precise way to identify which aspects of the translational models are most sensitive to and relevant for human anxiety.

Pharmacological manipulations

Pharmacological interventions provide a powerful means to experimentally manipulate anxiety across species. This approach involves two complementary strategies: the use of anxiolytic drugs to reduce anxiety, and anxiogenic drugs to induce or exacerbate anxiety.

Benzodiazepines, such as lorazepam, are widely used as anxiolytic drugs in experimental anxiety research given their fast-acting properties and reliable anxiolytic effects in humans (Kaplan & DuPont, 2005). Their use in animal studies serves two purposes: firstly, it helps validate translational models of anxiety-like behaviour – we discussed earlier how these drugs were crucial in establishing the face validity of many preclinical anxiety tests. Secondly, it allows for detailed investigation of the neurobiological mechanisms underlying anxiolysis at the circuit and cellular levels. This line of research has shown that benzodiazepines exert their anxiolytic effects primarily by enhancing the activity of gamma-aminobutyric acid (GABA), the main inhibitory neurotransmitter in the central nervous system (Chebib & Johnston, 1999). They bind to specific sites on GABA-A receptors, increasing the frequency of channel opening and thereby amplifying the inhibitory effect of GABA (Nutt & Malizia, 2001). This enhanced GABAergic transmission leads to reduced neuronal excitability in key brain regions associated with anxiety, such as the amygdala (Paulus et al., 2005).

Investigating the effects of benzodiazepines in humans using computational approaches offers a unique opportunity to validate the relevance of computational mechanisms engaged in translational tests of anxiety. However, few studies have adopted the computational framework in understanding the cognitive effects of benzodiazepines. Regardless, traditional cognitive studies have provided insights into how benzodiazepines affect anxiety-related information processing. For instance, cognitive studies have shown that benzodiazepines reduce attentional bias towards threat-related stimuli (Murphy et al., 2008), decrease the recognition of negative facial expressions (Garcez et al., 2020), and dampen startle responses to predictable and unpredictable punishments (Grillon et al., 2006). These findings suggest that benzodiazepines modulate how threat-related information is processed, but to establish the translational validity of the computational models of behaviour, it is necessary to examine how these cognitive changes map onto specific computational parameters.

How might this approach operate in practice? For instance, one could use computational modelling to examine how benzodiazepines affect parameters such as threat sensitivity, reward processing, or the balance between model-based and model-free decision-making. For example, the previously observed reduction in threat-related attentional bias following benzodiazepine administration (Murphy et al., 2008) might be reflected in decreased threat sensitivity parameters in computational models. Similarly, the benzodiazepine-dampened startle response (Grillon et al., 2006) could manifest as altered learning rates for negative outcomes. By examining how these computational parameters change in response to benzodiazepine administration, we can directly test whether the mechanisms identified in animal models truly reflect anxiety-related processes in humans.

Conversely, anxiogenic drugs provide a method to experimentally increase anxiety levels, offering another way to validate the computational mechanisms identified in translational models. One commonly used anxiogenic drug in human studies is yohimbine, an α2-adrenergic receptor antagonist. Yohimbine has been shown to induce anxiety-like states in both humans and animals (Bhattacharya et al., 1997; Charney et al., 1983), making it a valuable tool for translational research. This approach is particularly useful for identifying which computational mechanisms contribute to heightened anxiety states. For instance, yohimbine administration might lead to increased threat sensitivity parameters or altered learning rates for negative outcomes in computational models.

The use of both anxiolytic and anxiogenic drugs allows us to examine how our computational constructs respond to bidirectional changes in anxiety. This approach helps to pinpoint which parameters in our models are most closely linked to anxiety states in humans. Importantly, if we observe similar parameter changes in animal studies using analogous tasks and pharmacological manipulations, it strengthens our confidence in the translational validity of these animal models. An important secondary benefit of this computational psychopharmacology approach (Robbins & Cardinal, 2019) is that it can help elucidate the cognitive and computational mechanisms by which psychiatric drugs exert their therapeutic effects, given that despite their widespread use, our understanding of how they work mechanistically remains limited.

The threat-of-shock paradigm

Complementing the pharmacological approaches, the threat-of-shock paradigm (Robinson et al., 2013; Schmitz & Grillon, 2012) offers another method to experimentally manipulate anxiety in human participants. In threat-of-shock studies, participants are informed that they may receive mild, but aversive electric shocks during certain periods of a task, creating a state of anticipatory anxiety. This method has been validated as an effective way to induce anxiety-like states in humans (Grillon et al., 2019).

Studies using this paradigm have demonstrated that induced anxiety can lead to increased avoidance behaviours. For instance, one study found that acute anxiety induced by threat-of-shock exacerbated Pavlovian biases in instrumental learning (Mkrtchian et al., 2017). Specifically, under threat-of-shock conditions, individuals were more likely to make inappropriate responses driven by Pavlovian mechanisms rather than instrumental learning. This suggests that anxiety can disrupt the balance between different learning systems, potentially leading to maladaptive behaviours - a finding that can be captured through computational parameters representing the relative strengths of these learning systems. At the same time, another study found that while threat-of-shock successfully induced anxiety in participants, it did not significantly alter their performance or individual computational parameters on a translational measure of negative affective bias (Aylward et al., 2020). These mixed findings highlight the current uncertainty around which computational constructs are most robustly and consistently affected by anxiety manipulations.

Understanding how these computational constructs respond to experimentally induced changes in human anxiety will be an important step going forwards for bridging the gap between animal models and clinical anxiety. By identifying which parameters are consistently modulated by both anxiolytic and anxiogenic manipulations in humans, we can more confidently interpret similar measures in animal studies. This approach strengthens the translational validity of our models and provides a more robust foundation for preclinical research.

1.9 Aims and objectives of the present thesis

1.9.1 Overall aims

The overarching aim of this thesis is to explore what exactly the classical translational tests of anxiety are measuring, through the lens of computational neuroscience. This approach is rooted in the assumption that clinical anxiety involves specific information-processing mechanisms that can be elicited in certain task environments, and that these mechanisms are engaged not only in non-human animals but also in human participants who are not clinically anxious.

A better understanding of the information-processing mechanisms engaged in classical translational tests of anxiety have the potential to refine their validity and utility. To this end, I adopt a computational neuroscience approach as it offers a precise method for describing the putative neurocognitive processes that drive anxious behaviour, and how these processes might be affected by anxiety, or by external interventions relating to anxiety. Across several experiments, I use forward-translated versions of classical preclinical anxiety tests, both conditioned and unconditioned variants, to study their computational underpinnings in humans. By applying computational methods to

observed behaviours in these human tests, the present thesis aims to bridge the gap between animal models and human anxiety.

1.9.2 Overview of remaining chapters

In **Chapter 2**, I explore how anxiety impacts human behaviour in a forward-translated variant of conditioned tests of anxiety, in the form of an approach-avoidance conflict task that involves learning about rewards and punishments simultaneously. In rodent subjects, anxiety is typically correlated with greater punishment avoidance and reduced reward pursuit, so I investigated whether similar patterns can be found in humans. I also designed the measure in a way that facilitates its computational analysis, allowing me to investigate the precise cognitive mechanisms of behaviour and anxiety. I tested the measure in two independent, general-population samples (combined N > 900) who completed the task online, and I assessed whether and how task performance is associated with state anxiety and clinical anxiety symptoms. I predicted that, as seen in the rodent variants of this paradigm, anxiety would be correlated with a bias for punishment avoidance over reward pursuit in humans. The results demonstrate the potential of this measure as a translational and computational measure of anxiety-related avoidance.

Having found evidence suggesting that the behavioural measure from Chapter 2 engages anxiety-related computational mechanisms, I then tested its sensitivity to within-subjects anxiety manipulations in **Chapter 3**. Specifically, I conducted a double-blind, placebo-controlled, crossover experiment testing the effect of an acute dose of the anxiolytic, lorazepam, on task performance to investigate the effect of reducing anxiety on behaviour and on the putative computational mechanisms of behaviour identified in Chapter 2. In a parallel, within-subjects experiment in an independent sample, I used threat of unpredictable electric shock to induce a state of heightened anxiety during task performance, to investigate the behavioural and computational effects of increasing anxiety. Based on the findings of the previous chapter, I predicted that lorazepam, relative to placebo, would decrease avoidance behaviour and, computationally, lead individuals to place greater weight on reward pursuit relative to punishment avoidance. Conversely, I predicted that threat-of-shock, relative to safety, would increase avoidance behaviour and greater weighting of punishment avoidance over reward pursuit.

The behavioural task implemented in Chapters 2 and 3 was inspired by relatively simple behavioural tests from the preclinical literature (i.e. the conditioned tests of anxiety), which involve learning about the consequences of a limited set of behavioural responses. This simplicity is what facilitated a generative-computational analysis of behaviour. However, as noted earlier, unconditioned tests are popular in contemporary preclinical research, as they are less abstract, more naturalistic, and easier

to implement. In **Chapter 4**, I describe a new computational model of behaviour in unconditioned tests of anxiety which involve spontaneous exploration in space, and a balance between the natural drives of rodents to explore their environment, and their drives to avoid spatial locations associated with threat. I first proposed a generative, computational model of such exploration-avoidance behaviour within a reinforcement learning framework. Then, I validated the model with empirical data from a previous study implementing a virtual-reality, human version of the elevated plus maze, that also tested the effects of an anxiolytic (lorazepam) and an anxiogenic drug (yohimbine) on behaviour (Biedermann et al., 2017). As this chapter was more exploratory by nature, I tentatively predicted that parameters from the proposed computational model, when fitted to human data, would be able to explain individual differences in anxiety, and the effects of lorazepam and yohimbine on behaviour.

Finally, in **Chapter 5**, the General Discussion, I discuss the broader implications of my research and directions for future work.

2 Approach-avoidance reinforcement learning as a translational and computational model of anxiety-related avoidance

2.1 Preface

Avoidance is a core symptom across many anxiety disorders and other psychiatric conditions. However, progress in understanding the cognitive and neural mechanisms underlying avoidance behaviour has been hindered by differences in how avoidance is measured across human and non-human animal studies. This chapter describes the development and validation of a novel translational task designed to quantify anxiety-related avoidance in humans using the same computational principles as approach-avoidance conflict tests in animal research. By translating a reinforcement learning paradigm from rodents to humans, this task captures the key decision-making processes involved when individuals must weigh reward-pursuit against punishment-avoidance. I validate the task in two large online samples of participants drawn from the general population and investigate how anxiety affects their performance in the task.

Note to examiners.

The work in this chapter was published as a research article in eLife, where the examiners can also find publicly available peer review which was overall favourable. The text is taken largely verbatim from this published article with the exception that it has been put into first person and the methods and results sections have been re-ordered to make them consistent with the rest of the thesis. The reference to the article is:

Yumeya Yamamori, Oliver J Robinson, Jonathan P Roiser (2023). Approach-avoidance reinforcement learning as a translational and computational model of anxiety-related avoidance. *eLife* 12:RP87720. https://doi.org/10.7554/eLife.87720.4

2.2 Abstract

Although avoidance is a prevalent feature of anxiety-related psychopathology, differences in the measurement of avoidance between humans and non-human animals hinder our progress in its theoretical understanding and treatment. To address this, here I developed a novel translational measure of anxiety-related avoidance in the form of an approach-avoidance reinforcement learning task, by adapting a paradigm from the non-human animal literature to study the same computational processes in human participants. I used computational modelling to probe the putative cognitive mechanisms underlying approach-avoidance behaviour in this task and investigated how they relate to subjective task-induced anxiety. In a large online study (n = 372), participants who experienced greater task-induced anxiety avoided choices associated with punishment, even when this resulted in lower overall reward. Computational modelling revealed that this association was explained by greater individual sensitivities to punishment relative to rewards. I replicated these findings in an independent sample (n = 627), and I also found fair-to-excellent reliability of measures of task performance in a sub-sample retested 1 week later (n = 57). My findings demonstrate the potential of approach-avoidance reinforcement learning tasks as translational and computational models of anxiety-related avoidance.

2.3 Introduction

Avoiding harm or potential threats is essential for our well-being and survival, but it can sometimes lead to negative consequences in and of itself. This occurs when avoidance is costly, that is, when the act to avoid harm or threat requires the sacrifice of something positive. When offered the opportunity to attend a job interview, for example, the reward of landing a new job must be weighed against the risk of experiencing rejection and/or embarrassment due to fluffing the interview. On balance, declining a single interview would likely have little impact on one's life, but routinely avoiding job interviews would jeopardise one's long-term professional opportunities. More broadly, consistent avoidance in similar situations, which are referred to as involving approach-avoidance conflict, can have an increasingly negative impact as the sum of forgone potential reward accumulates, ultimately leading to missed opportunities in life. Such excessive avoidance is a hallmark symptom of pathological anxiety (Barlow, 2002) and avoidance biases during approach-avoidance conflict are thought to drive the development and maintenance of anxiety-related (Aupperle & Paulus, 2010; Hayes, 1976; Mkrtchian et al., 2017; Stein & Paulus, 2009) and other psychiatric disorders (Aldao et al., 2010; Kakoschke et al., 2019).

How is behaviour during approach-avoidance conflict measured quantitatively? In non-human animals, this can be achieved through behavioural paradigms that induce a conflict between natural approach and avoidance drives. For example, in the Geller-Seifter conflict test (Geller et al., 1962), rodents decide whether or not to press a lever that is associated with the delivery of both food pellets and shocks, thus inducing a conflict. Anxiolytic drugs typically increase lever presses during the test (Treit et al., 2010), supporting its validity as a model of anxiety-related avoidance. Consequently, this test along with others inducing approach-avoidance conflict are often used as non-human animal tests of anxiety (Griebel & Holmes, 2013). In humans, on the other hand, approach-avoidance conflict has historically been measured using questionnaires such as the Behavioural Inhibition/Activation Scale (Carver & White, 1994), or cognitive tasks that rely on motor/response time biases, for example by using joysticks to approach/move towards positive stimuli and avoid/move away from negative stimuli (Guitart-Masip et al., 2012; Kirlic et al., 2017; Mkrtchian et al., 2017; Phaf et al., 2014).

Translational approaches, specifically when equivalent tasks are used to measure the same cognitive construct in humans and non-human animals, benefit the study of avoidance and its relevance to mental ill-health for two important reasons (Bach, 2021; Pike et al., 2021). First, precise causal manipulations of neural circuitry such as chemo/optogenetics are only feasible in non-human animals, whereas only humans can verbalise their subjective experiences – it is only by using translational measures that we can integrate data and theory across species to achieve a

comprehensive mechanistic understanding. Second, translational paradigms can make the testing of novel anxiolytic drugs more efficient and cost-effective, since those that fail to reduce anxiety-related avoidance in humans could be identified earlier during the development pipeline.

Since the inception of non-human animal models of psychiatric disorders, researchers have recognised that it is infeasible to fully recapitulate a disorder like generalised anxiety disorder in non-human animals due to the complexity of human cognition and subjective experience. Instead, the main strategy has been to reproduce the physiological and behavioural responses elicited under certain scenarios across species, such as during fear conditioning (Craske et al., 2006; Milad & Quirk, 2012). Computational modelling of behaviour allows us to take a step further and probe the mechanisms underlying such physiological and behavioural responses, using a mathematically principled approach. A fundamental premise of this approach is that the brain acts as an information-processing organ that performs computations responsible for observable behaviours, including approach and avoidance (for a recent review on the application of computational methods to approach-avoidance conflict, see Letkiewicz et al., 2023). With respect to translational models, it follows that a valid translational measure of some behaviour not only matches the observable responses across species, but also their underlying computational mechanisms. This criterion, which has been termed computational validity, has been proposed as the primary basis for evaluating the validity of translational measures (Redish et al., 2022).

To exploit the advantages of translational and computational approaches, there has been a recent surge in efforts to create relevant approach-avoidance conflict tasks in humans, based on non-human animal models (referred to as back-translation; Kirlic et al., 2017). One stream of research has focused on exploration-based paradigms, which involve measuring how participants spend their time in environments that include potentially anxiogenic regions. Rodents typically avoid such regions, which can take the form of tall, exposed platforms as in the elevated plus maze (Pellow et al., 1985), or large open spaces as in the open-field test (Hall, 1934). Similar behaviours have been observed in humans when virtual reality was used to translate the elevated plus maze (Biedermann et al., 2017), and when participants were free to walk around a field or around town (Walz et al., 2016). These translational tasks have excellent face validity in that they involve almost identical behaviour compared to their non-human analogues, and their predictive validity is supported by findings that individual differences in self-reported anxiety are associated with greater avoidance of the exposed/open regions of space (Biedermann et al., 2017; Walz et al., 2016). But what of their computational validity? The computations underlying free movement are difficult to model as the action space is large and complex (consisting of two-dimensional directions, velocity, etc.), which

makes it difficult to understand the precise cognitive mechanisms underlying avoidance in these tasks.

Other studies have designed tasks that emulate operant conflict paradigms in which non-human participants learn to associate certain actions with both reward and punishment, such as the Geller-Seifter and Vogel conflict tests (Geller et al., 1962; Vogel et al., 1971). These studies used decision-making tasks in which participants choose between offers consisting of both reward and punishment outcomes (Aupperle et al., 2011; Sierra-Mercado et al., 2015). The strengths of such decision tasks lie in the simplicity of their action space (accept/reject an offer; choose offer A or B), which facilitates the computational modelling of decision-making, through for example value-based logistic regression models (Ironside et al., 2020) or drift-diffusion modelling (Pedersen et al., 2021). However, these tasks have less face validity compared to exploration-based tasks, as explicit offers ('Will you accept £1 to incur an electric shock?') are incomprehensible to the majority of non-human animals (although non-human primates can be trained to do so, e.g. Ironside et al., 2020; Sierra-Mercado et al., 2015). Instead, non-human participants such as rodents typically have to learn the possible outcomes given their actions, for example that a lever press is associated with both food pellets and electric shocks. Critically, this means that humans performing offer-based decision-making are doing something computationally dissimilar to non-human participants during operant conflict tasks, and the criterion of computational validity is unmet. Moreover, evidence from economic decision-making suggests that explicit offers of probabilistic outcomes can impact decision-making differently compared to when probabilistic contingencies need to be learned from experience (referred to as the 'description-experience gap'; Hertwig & Erev, 2009); this finding raises potential concerns regarding the use of offer-based tasks in humans as approximations of non-human tasks that do not involve explicit offers.

Therefore, in the current study, I developed a novel translational task that was specifically designed to mimic the computational processes involved in non-human operant conflict tasks, and was also amenable to computational modelling of behaviour. During the task, participants chose between options that were asymmetrically associated with rewards and punishments, such that the more rewarding options were more likely to result in punishment. However, in contrast to previous translations of these tests using purely offer-based designs, participants had to learn the likelihoods of receiving rewards and/or punishments from experience, thus ensuring computational validity.

Preserving the element of learning in the task did not necessarily risk confounding the study of approach-avoidance behaviour. Rather, a critical advantage of my approach was that I could model the cognitive processes underlying approach-avoidance behaviour using classical reinforcement

learning algorithms (Sutton & Barto, 2018), and then use these algorithms to disentangle the relative contributions of learning and value-based decision-making to anxiety-related avoidance. Specifically, using the reinforcement learning framework; I accounted for individual differences in learning by estimating learning rates, which govern the degree to which recently observed outcomes affect subsequent actions; and I explained individual differences in value-based decision-making using outcome sensitivity parameters, which model the subjective values of rewards and punishments for each participant.

I therefore aimed to develop a new translational measure of approach-avoidance conflict. Based on previous findings that anxiety is associated with avoidance biases under conflict (Biedermann et al., 2017; Walz et al., 2016), I predicted that anxiety would increase avoidance responses (and conversely decrease approach responses) and that this would be reflected in changes in computational parameters. Having found evidence for this in a large initial online study, I retested the prediction that anxiety-related avoidance is explained by greater sensitivity to punishments relative to rewards in a pre-registered independent replication sample (https://osf.io/8dm95).

2.4 Methods

2.4.1 Participants

I recruited participants from Prolific (https://www.prolific.com/). Participants had to be aged 18–60, speak English as their first language, have no language-related disorders/literacy difficulties, have no vision/hearing difficulties/have no mild cognitive impairment or dementia, and be resident in the UK. Participants were reimbursed at a rate of £7.50 per hour, and could earn a bonus of up to an additional £7.50 per hour based on their performance on the task. The study received ethical approval from the University College London Research Ethics Committee (ID 15253/001).

I first recruited a discovery sample to investigate whether anxiety was associated with avoidance in the novel task. I conducted an *a priori* power analysis to detect a minimally interesting correlation between anxiety and avoidance of r = 0.15 ($alpha_{two\text{-}tailed}$=0.05, power = 0.80), resulting in a required sample size of 346. I aimed to replicate my initial finding (that task-induced anxiety was associated with the model-derived measure of approach-avoidance bias) in a replication sample. In the initial sample, the size of this effect was Kendall's τ = 0.099, which approximates to a Pearson's r of 0.15 – this was the basis for the power analysis for the replication sample. Therefore, the power analysis was configured to detect a correlation of r = 0.15 ($alpha_{one\text{-}tailed}$=0.05, power = 0.95), resulting in a required sample size of n = 476. A total of 423 participants completed the initial study (mean age = 38, SD = 10, 59% female), and 726 participants completed the replication study (mean age = 34, SD =

10, 51% female). The predictions for the replication study were pre-registered (https://osf.io/8dm95).

2.4.2 Headphone screen and calibration

I used aversive auditory stimuli as my punishments in the reinforcement learning task, which consisted of combinations of female screams and high-frequency tones that were individually rated as highly aversive in a previous online study (Seow & Hauser, 2022). The stimuli can be found online at https://osf.io/m3zev/. To maximise the aversiveness of the stimuli, I asked participants to use headphones whilst they completed the task. All participants completed an open-source, validated auditory discrimination task (Woods et al., 2017) to screen out those who were unlikely to be using headphones. Participants could only continue to the main task if they achieved an accuracy threshold of five out of six trials in this screening task.

To further increase the aversiveness of the punishments, I wanted each participant to hear them at a loud volume. I initially used a calibration task to increase the system volume and implemented attention checks in the main task to ensure that participants had kept the volume at a sufficiently high level during the task (see below). Before beginning the calibration task, participants were advised to set their system volume to around 10% of the maximum possible. Then, I presented two alternating auditory stimuli: a burst of white noise which was set to have the same volume as the aversive sounds in the main task, and a spoken number played at a lower volume to one ear ('one', 'two', or 'three'; these stimuli are also available online at https://osf.io/m3zev/). I instructed participants to adjust their system volume so that the white noise was loud but comfortable. At the same time, they also had to press the number key on their keyboard corresponding to the number spoken between each white noise presentation. The rationale for the difference in volume was to fix a lower boundary for each participant's system volume, so that the volume of the punishments presented in the main task was locked above this boundary. I used the white noise as a placeholder for the punishments to limit desensitisation. Since the volume of the punishments, and therefore at least in part their aversiveness, was determined by the participants themselves, I conducted sensitivity analyses on all reported findings after accounting for self-reported unpleasantness ratings for the punishment – this did not materially affect my main findings (see Appendix 2.9).

2.4.3 The approach-avoidance reinforcement learning task

I developed a novel approach-avoidance conflict task, based on the 'restless bandit' design (Daw et al., 2006), in which the participants' goal was to accrue rewards whilst avoiding punishments (Figure 2.1a). On each trial, participants had 2 seconds to choose one of two options, depicted as distinct visual stimuli (Figure 2.1a), using the left/right arrow keys on their keyboard. The chosen option was

then highlighted by a white border, and participants were presented with the outcome for that trial for 2 seconds. This could be one of (1) no outcome, (2) a reward (an animated coin image), (3) a punishment (an aversive sound), or (4) both the reward and punishment (Figure 2.1b). The probability of observing the reward and the aversive sound at each option fluctuated randomly and independently over trials, except the probability of observing the aversive sound at the safe option, which was set to 0 across all trials (Figure 2.1c).

The options were asymmetric in their associated outcomes, such that one option (which I refer to as the 'conflict' option) was associated with both the reward (with high probability on average) and aversive sound, whereas the other (the 'safe' option) was only associated with the reward (with low probability on average, Figure 2.1d). In other words, choosing the conflict option could result in any of the four possible outcomes on any given trial, but the safe option could only lead to either the reward or no reward, and never the aversive sounds. Therefore, I induced an approach-avoidance conflict between the options as the conflict option was generally more rewarding but also more likely to result in punishment. Participants were not informed about this asymmetric feature of the task, but they were told that the probabilities of observing the reward or the aversive sound were independent and that they could change across trials.

I incentivised participants to collect the coins by informing them that they could earn a bonus payment of up to £2.50 based on the number of coins earned during the task. If no choice was made, the text 'Too slow!' was shown for 2 seconds and the trial was skipped.

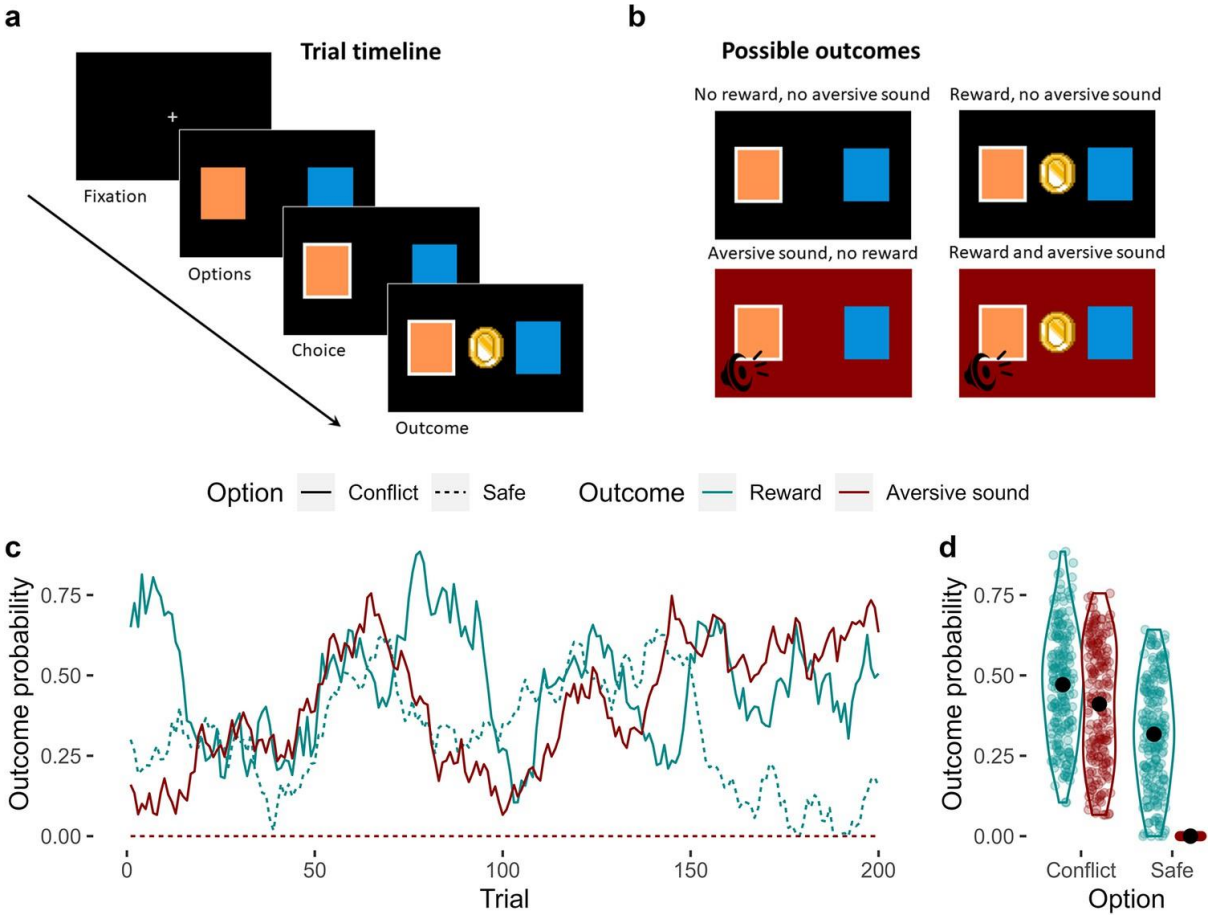


Figure 2.1, The approach-avoidance reinforcement learning task. *(a) Trial timeline. A fixation cross initiates the trial. Participants are presented with two options for up to 2 s, from which they choose one. The outcome is then presented for 1 s. (b) Possible outcomes. There were four possible outcomes: (1) no reward and no aversive sound; (2) a reward and no aversive sound; (3) an aversive sound and no reward; or (4) both the reward and the aversive sound. (c) Probabilities of observing each outcome given the choice of option. Unbeknownst to the participant, one of the options (which I refer to as the 'conflict' option – solid lines) was generally more rewarding compared to the other option (the 'safe' option – dashed line) across trials. However, the conflict option was also the only option of the two that was associated with a probability of producing an aversive sound (the probability that the safe option produced the aversive sound was 0 across all trials). The probabilities of observing each outcome given the choice of option fluctuated randomly and independently across trials. The correlations between these dynamic probabilities were negligible (mean Pearson's r = 0.06). (d) Distribution of outcome probabilities by option and outcome. On average, the conflict option was more likely to produce a reward than the safe option. The conflict option had a variable probability of producing the aversive sound across trials, but this probability was always 0 for the safe option. Black points represent the mean probability.*

The latent outcome probabilities were generated as decaying Gaussian random walks in a similar fashion to previous restless bandit tasks (Daw et al., 2006), where the probability, p, of observing outcome, o, at option, i, on trial, t, drifted according to the following:

$$p_{i,o}(t+1) = \lambda p_{i,o}(t) + (1-\lambda)\theta + e$$

$$e \sim \mathrm{N}(0, 0.04^2)$$

The decay parameter, λ, was set to 0.97, which pulled all values towards the decay centre θ, which was set to 0.5. The random noise, e, added on each trial was drawn from a zero-mean Gaussian with a standard deviation of 0.04. All probabilities of observing the reward at the safe option were scaled

down by a factor of 0.7 so that it was less rewarding than the conflict option, on average (Figure 2.1d). I initialised each walk at the following values: conflict option reward probability – 0.7, safe option reward probability – 0.35, conflict option punishment probability – 0.2. I only used sets of latent outcome probabilities where the cross-correlations across each probability walk were below 0.3. A total of four sets of outcome probabilities were used across the studies.

2.4.4 Attention checks

I implemented three auditory attention checks during the main task, to ensure that participants were using an appropriate volume to listen to the auditory stimuli. As participants could potentially mute their device during the outcome phase of the trial only (i.e. only when the aversive sounds could occur), the timing of these checks had to coincide with moment of outcome presentation. Therefore, I implemented these checks at the outcome phase of two randomly selected trials where the participants' choices led to no outcome (no reward or punishment), and also once at the very end of the task. Specifically, I played one of the spoken numbers used during the calibration task (see above) and the participant had 2.5 seconds to respond using the corresponding number key on their keyboard.

2.4.5 Subjective ratings

Immediately after the task and the final attention check, I obtained subjective ratings about the task using the following questions: 'How unpleasant did you find the sounds?' and 'How anxious did you feel during the task?'. Responses were measured on a continuous slider ranging from 'Not at all' (0) to 'Extremely' (50).

2.4.6 Data cleaning

As the task was implemented online, meaning I could not ensure the same testing standards as would be possible in person, I implemented three exclusion criteria to improve data quality. First, I excluded participants who missed a response to more than one auditory attention check (see above; 8% in both discovery and replication samples) – as these occurred infrequently and the stimuli used for the checks were played at relatively low volume, I included participants with incorrect responses so long as a response was made. Second, I excluded those who responded with the same response key on 20 or more consecutive trials (>10% of all trials; 4%/6% in discovery and replication samples, respectively) – note that as the options randomly switched sides on the screen across trials, this did not exclude participants who frequently and consecutively chose a certain option. Lastly, I excluded those who did not respond on 20 or more trials (1%/2% in discovery and replication samples, respectively). Overall, I excluded 51 out of 423 (12%) in the discovery sample, and 98 out of 725 (14%) in the replication sample. I conducted all the analyses with and without these exclusions.

These exclusions had no appreciable effects on the main findings reported in the manuscript, and differences before and after excluding data are reported in Appendix 2.1.

2.4.7 Mixed-effects regressions of trial-by-trial choice

I specified two hierarchical logistic regression models to test the effect of the dynamic outcome probabilities on choice. Both models included predictors aligned with approach and avoidance responses, based on the latent probabilities of observing an outcome given the chosen option at trial t, $\mathrm{P}(\mathrm{outcome} \mid \mathrm{option})_t$. First, if individuals were learning to choose to maximise their rewards earned, they should generally be choosing the option which was relatively more likely to produce the reward on each trial. This would result in a significant effect of the difference in reward probabilities across options, as given by:

$$\delta\mathrm{P}(\mathrm{reward})_t = \mathrm{P}(\mathrm{reward} \mid \mathrm{conflict})_t \text{ - } \mathrm{P}(\mathrm{reward} \mid \mathrm{safe})_t$$

Second, if individuals were learning to avoid punishments, they should generally be avoiding the conflict option on trials when it is likely to produce the punishment. This would result in a significant effect of the punishment probability at the conflict option:

$$\mathrm{P}(\mathrm{punishment} \mid \mathrm{conflict})_t$$

The first regression model simply included these two predictors and random intercepts varying by participant to predict choice:

$$\mathrm{Choice} \sim \delta\mathrm{P}(\mathrm{reward}) + \delta\mathrm{P}(\mathrm{punishment}) + (1 \mid \mathrm{participant})$$

The second model tested the effect of task-induced anxiety on choice, by including anxiety as a main effect and its interactions with $\delta\mathrm{P}(\mathrm{reward})_t$ and $\mathrm{P}(\mathrm{punishment} \mid \mathrm{conflict})_t$:

$$\mathrm{Choice} \sim \delta\mathrm{P}(\mathrm{reward}) + \delta\mathrm{P}(\mathrm{punishment}) + \mathrm{Anxiety} +$$
$$\mathrm{Anxiety} * \delta\mathrm{P}(\mathrm{reward}) + \mathrm{Anxiety} * \delta\mathrm{P}(\mathrm{punishment}) + (1 \mid \mathrm{participant})$$

2.4.8 Computational modelling using reinforcement learning

Model specification

I used variations of classical reinforcement learning algorithms (Sutton and Barto, 2018), specifically the 'Q-learning' algorithm (Watkins & Dayan, 1992), to model participants' choices in the main task (Table 2.1). Briefly, the models assume that participants estimate the probabilities of observing a reward and punishment separately for each punishment, and continually update these estimates after observing the outcomes of their actions. Consequently, participants can then use these estimates to choose the option which will maximise subjective value, for example by choosing the option which is more likely to produce a reward, or avoiding an option if it is very likely to produce a

punishment. By convention, I refer to the probability estimates as 'Q-values'; for example, the estimated probability of observing a reward after choosing the conflict option is written as $Q^r(\text{conflict})$, whereas the probability of observing a punishment after choosing the safe option would be written as $Q^p(\text{safe})$.

In all models, the probability estimates of observing outcome, o = [r, p], of the chosen option, a = [conflict, safe], on trial, t, were updated via: $Q^o_{t+1}(a) = Q^o_t(a) + \alpha \cdot [o^t - Q^o_t(a)]$, where the learning rate is controlled by $\alpha \in (0,1)$. In some models (Table 2.1), the learning rate was split into reward- and punishment-specific parameters, α^r and α^p.

On each trial, the probability estimates for observing rewards and aversive sounds are integrated into action weights, W, via:

$$W = \beta \cdot (Q^r - Q^p)$$

where individual differences in sensitivity to the outcomes (how much the outcomes affect value-based choice) are parameterised by the outcome sensitivity parameter, β. In models with a single sensitivity parameter, participants were assumed to value the reward and punishment equally, such that obtaining a reward had the same subjective value as avoiding a punishment. However, as individual differences in approach-avoidance conflict could manifest in asymmetries in the valuation of reward/punishment, I also specified models (Table 2.1) where β was split into reward- and punishment-specific parameters, β^r and β^p:

$$W = \beta^r \cdot Q^r - \beta^p \cdot Q^p$$

The action weights are then the basis on which a choice was made between the options, which was modelled with the softmax function:

$$P(a) = \frac{W(a)}{\sum_i W(i)}$$

I also tested all models with the inclusion of a lapse term, which allows for choice stochasticity to vary across participants, but since this did not improve the evidence model for any model, I do not discuss this further.

Table 2.1, Model specification

Model	**Parameters**	
Reward learning only	α^r	β^r
Punishment learning only	α^p	β^p
Symmetrical learning	α	β

Asymmetric learning rates	α^r	α^p	β	
Asymmetric sensitivities		α	β^r	β^p
Asymmetric learning rates and sensitivities	α^r	α^p	β^r	β^p

Model fitting and comparison

I fitted the models using a hierarchical EM algorithm (Huys et al., 2011) – the algorithm and supporting code are available online at https://www.quentinhuys.com/pub/emfit/. All prior distributions of the untransformed parameter values were set as Gaussian distributions to regularise fitting and to limit extreme values. Parameters were transformed within the models, via the sigmoid function for parameters constrained to values between 0 and 1; or exponentiated for strictly positive parameters. I compared evidence across models using integrated Bayesian information criterion (Huys et al., 2011) which integrates both model likelihood and model complexity into a single measure for model comparison.

2.4.9 Mediation analysis of anxiety-related avoidance

To investigate the cognitive processes underlying anxiety-related avoidance, I tested for potential mediators of the effect of task-induced anxiety on avoidance choices in the approach-avoidance reinforcement learning task. Mediation was assessed using structural equation modelling with the 'lavaan' package in R (Rosseel, 2012). This allowed me to test multiple potential mediators within a single model. I used maximum-likelihood estimation with and without bootstrapped standard errors (using 2000 bootstrap samples) for estimation. The potential mediators were any parameters from the computational model that significantly correlated with task-induced anxiety. Therefore, the model included the following variables: task-induced anxiety, avoidance choices (computed as the proportion of safe option choices for each participant), and model parameters which correlated with task-induced anxiety. The model was constructed such that the model parameters were parallel mediators of the effect of task-induced anxiety on avoidance choices. The model parameters were allowed to covary with one another.

2.4.10 Symptom questionnaires

Psychiatric symptoms were measured using the Generalised Anxiety Disorder scale (GAD7; Spitzer et al., 2006), Patient Health Questionnaire depression scale (PHQ8; Kroenke et al., 2009), and the Brief Experiential Avoidance Questionnaire (BEAQ; Gamez et al., 2014). The BEAQ indexes an unwillingness to experience negative emotions/experiences (Hayes et al., 1996). I included experiential avoidance as a direct measure of avoidance symptoms/beliefs, given that the task involved avoiding stimuli inducing negative affect.

2.4.11 Test-retest reliability analysis

Split-half reliability

I assessed the split-half reliability of the task by correlating the overall proportion of conflict option choices and model parameters from the winning model across the first and second half of trials. For overall choice proportion, reliability was simply calculated via Pearson's correlations. For the model parameters, I calculated model-derived estimates of Pearson's r values from the parameter covariance matrix when first- and second-half parameters were estimated within a single model, following a previous approach recently shown to accurately estimate parameter reliability (Waltmann et al., 2022). I interpreted indices of reliability based on conventional values of <0.40 as poor, 0.4–0.6 as fair, 0.6–0.75 as good, and >0.75 as excellent reliability (Fleiss, 1986).

Test-retest reliability

I retested a subset of the sample (n = 57) to assess the reliability of my task. I determined my sample size via an a priori power analysis using GPower (Faul et al., 2007), which was set to detect the smallest meaningful effect size of r = 0.4, as any lower test-retest correlations are considered poor by convention (Cicchetti, 1994; Fleiss, 1986). The required sample size was thus 50 participants, based on a one-tailed significance threshold of 0.05 and 90% power. Participants were recruited from Prolific. I invited participants from the replication study (providing they had not been excluded on the basis of data quality) to complete the study again, at least 10 days after their first session. The retest version of the approach-avoidance reinforcement learning task was identical to that in the first session, except for the use of new stimuli to represent the conflict and safe options and different latent outcome probabilities to limit practice effects.

I calculated intraclass correlation coefficient (ICCs) to assess the reliability of my model-agnostic measures. Specifically, I used consistency-based ICC(3,1) (following the notation of Shrout & Fleiss, 1979). To estimate the reliability of the model-derived measures, I estimated individual- and session-level model parameters via a joint model in an approach recently shown to improve the accuracy of estimating reliability (Waltmann et al., 2022). Briefly, this involved fitting behavioural data from both test and retest sessions per participant in a single model, but specifying unique parameters per session. Since the fitting algorithm assumes parameters are drawn from a multivariate Gaussian distribution, information about the reliability of the parameters can be derived from the covariance of each parameter across time. These covariance values can subsequently be converted to Pearson's r values, and thus indices of reliability.

2.5 Results

2.5.1 Participants

The discovery sample had a significantly greater proportion of female participants than the replication sample (59% vs 52%, $\chi^2 = 4.64$, $p = 0.031$). The average age was also significantly different across samples (discovery sample mean = 37.7, SD = 10.3, replication sample mean = 34.3, SD = 10.4; $t_{785.5} = 5.06$, $p < 0.001$). The differences in self-reported psychiatric symptoms across samples did not reach significance ($p > 0.086$). Since the behavioural findings were broadly comparable across samples, I primarily report findings from the discovery sample and discuss deviations from these findings in the replication sample. A detailed comparison of statistical findings across samples is reported in the Appendix.

2.5.2 Choices reflect both approach and avoidance

First, I verified that participants demonstrated both approach and avoidance responses during the task through a hierarchical logistic regression model, in which I used the latent outcome probabilities to predict trial-by-trial choice. Overall, participants learned to optimise their choices to accrue rewards and avoid the aversive sounds. As expected, across trials participants chose (i.e. approached) the option with relatively higher reward probability ($\beta = 0.98 \pm 0.03$, $p < 0.001$; Figure 2.2a and c). At the same time, they avoided the conflict option when it was more likely to produce the punishment ($\beta = -0.33 \pm 0.04$, $p < 0.001$; Figure 2.2a and d). These effects were also clearly evident in the replication sample (effect of relative reward probability: $\beta = 0.95 \pm 0.02$, $p < 0.001$; effect of punishment probability: $\beta = -0.52 \pm 0.03$, $p < 0.001$; Appendix 2.1—table 1). While sex was significantly associated with choice in the hierarchical logistic regression in the discovery sample ($\beta = 0.16 \pm 0.07$, $p = 0.028$) with males being more likely to choose the conflict option, this pattern was not evident in the replication sample ($\beta = 0.08 \pm 0.06$, $p = 0.173$), and age was not significantly associated with choice in either sample ($p > 0.2$). Across individuals, there was considerable variability in overall choice proportions (discovery sample: mean = 0.52, SD = 0.14, min/max = [0.03, 0.96]; replication sample: mean = 0.52, SD = 0.15, min/max = [0.01, 0.99]). When asked to rate the unpleasantness of the punishments from ‘Not at all’ to ‘Extremely’ (scored from 0 to 50, respectively), participants gave a mean rating of 31.7 (SD = 12.8).

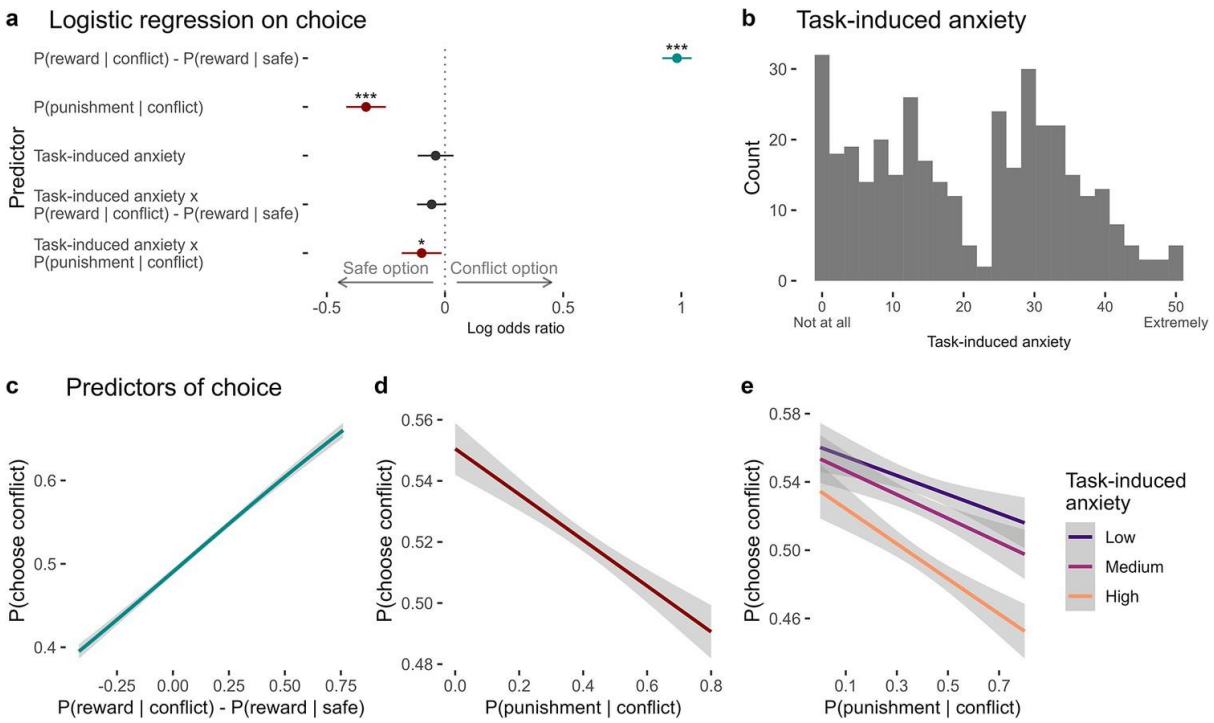


Figure 2.2, Predictors of choice in the approach-avoidance reinforcement learning task. *(a) Coefficients from the mixed-effects logistic regression of trial-by-trial choices in the task (n = 369). On any given trial, participants chose the option that was more likely to produce a reward. They also avoided choosing the conflict option when it was more likely to produce the punishment. Task-induced anxiety significantly interacted with punishment probability. Significance levels are shown according to the following: p < 0.05 – *; p < 0.01 – **; p < 0.001 – ***. Error bars represent confidence intervals. (b) Subjective ratings of task-induced anxiety, given on a scale from 'Not at all' (0) to 'Extremely' (50). (c) On each trial, participants were likely to choose the option with greater probability of producing the reward. (d) Participants tended to avoid the conflict option when it was likely to produce a punishment. (e) Compared to individuals reporting lower anxiety during the task, individuals experiencing greater anxiety showed greater avoidance of the conflict option, especially when it was more likely to produce the punishment. Note. Figures c–e show logistic curves fitted to the raw data using the 'glm' function in R. For visualisation purposes, I categorised continuous task-induced anxiety into tertiles. I show linear curves here since these effects were estimated as linear effects in the logistic regression models, however the raw data showed non-linear trends – see Appendix 2.11—figure 1.*

2.5.3 Task-induced anxiety is associated with greater avoidance

On average, participants reported a moderate level of task-induced anxiety (mean rating of 21, SD = 14; Figure 2.2b). Task-induced anxiety was significantly associated with multiple aspects of task performance. At the most basic level, task-induced anxiety was correlated with the proportion of conflict/safe option choices made by each participant across all trials, with greater anxiety corresponding to a preference for the safe option over the conflict option (permutation-based non-parametric correlation test: Kendall's $\tau = -0.074$, $p = 0.033$), at the cost of receiving fewer rewards on average.

To investigate this association further, I included task-induced anxiety into the hierarchical logistic regression model of trial-by-trial choice. Here, task-induced anxiety significantly interacted with punishment probability ($\beta = -0.10 \pm 0.04$, $p = 0.022$; Figure 2.2a), such that more anxious individuals were proportionately less willing to choose the conflict option during times it was likely to produce

the punishment, relative to less-anxious individuals (Figure 2.2e). This explains how task-induced anxiety influenced participants' choices, as the main effect of anxiety was not significant in the model, although it was when the interaction effects were excluded. The analogous interaction between task-induced anxiety and relative reward probability was non-significant (β = –0.06 ± 0.03, *p* = 0.074, Figure 2.2a).

These associations with task-induced anxiety were again clearly evident in the replication sample (correlation between task-induced anxiety and proportion of conflict/safe option choices: τ = –0.075, *p* = 0.005; interaction between task-induced anxiety and punishment probability in the hierarchical model: β = –0.23 ± 0.03, *p* < 0.001; Appendix 2.1—table 1).

2.5.4 A reinforcement learning model explains individual choices

Next, I investigated the putative cognitive mechanisms driving behaviour on the task by fitting standard reinforcement learning models to trial-by-trial choices. The winning model in both the discovery and replication samples included reward- and punishment-specific learning rates, and reward- and punishment-specific sensitivities (i.e. four parameters in total; Figure 2.3a and b).

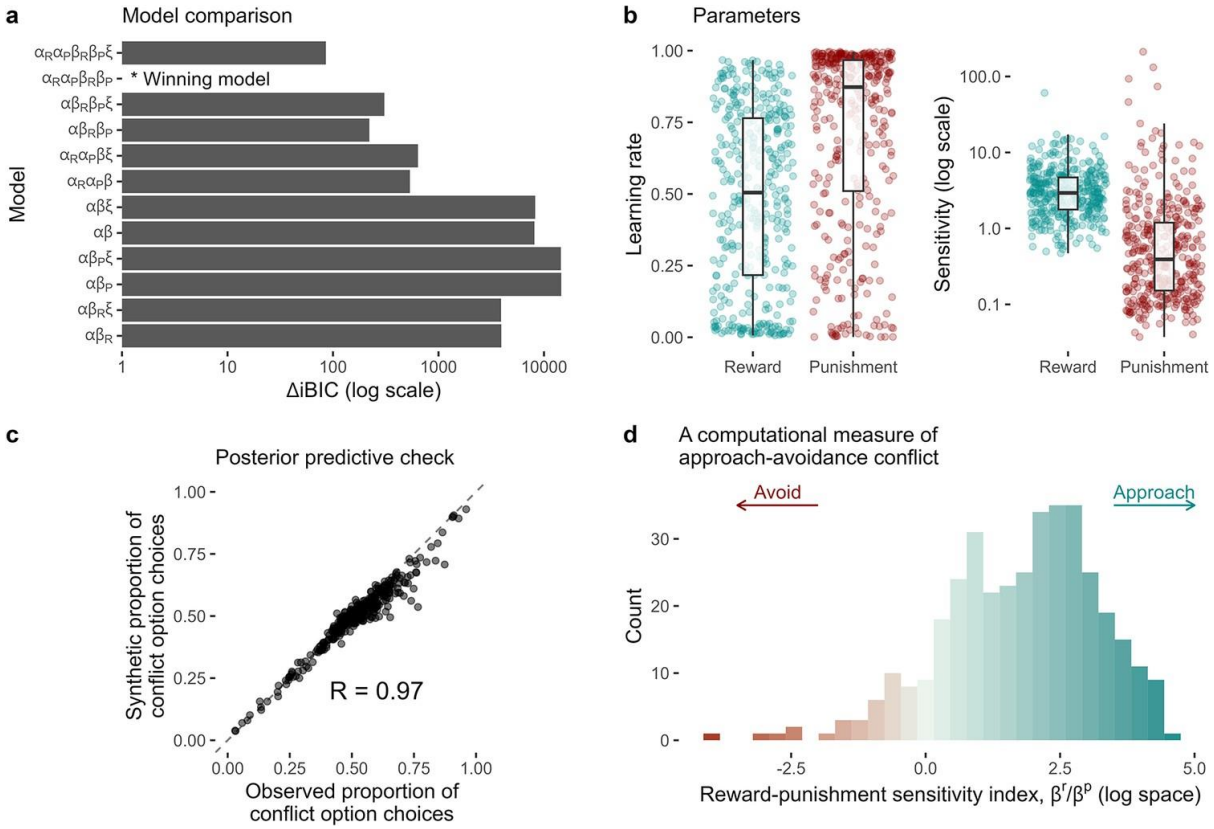


Figure 2.3, Computational modelling of approach-avoidance reinforcement learning. *(a) Model comparison results (n = 369). The difference in integrated Bayesian information criterion scores from each model relative to the winning model is indicated on the x-axis. The winning model included specific learning rates for reward (αR) and punishment learning (αp), and specific outcome sensitivity parameters for reward (βR) and punishment (βP). Some models were tested with the inclusion of a lapse term (ξ). (b) Distributions of individual parameter values from the winning model. (c) The winning model was able to reproduce the proportion of conflict option choices over all trials in the observed data with high accuracy*

(observed vs predicted data r = 0.97). (d) The distribution of the reward-punishment sensitivity index – the computational measure of approach-avoidance bias. Higher values indicate approach biases, whereas lower values indicate avoidance biases.

To assess how well this model captured the observed data, I simulated data from the model given each participant's parameter values, choices, and observed outcomes. These synthetic choice data closely resembled the actual choices made by the participants (Figure 2.3c). Parameter recovery from the winning model was excellent, with Pearson's r values between data-generating and recovered parameters ranging from 0.77 to 0.86 (Appendix 2.4—figure 1)

To quantify individual approach-avoidance bias computationally, I calculated the ratio between the reward and punishment sensitivity parameters (β^r/β^p). As the task requires simultaneously balancing reward pursuit (i.e. approach) and punishment avoidance, this composite measure provides an index of where an individual lies on the continuum of approach vs avoidance, with higher and lower values indicating approach and avoidance biases, respectively. I refer to this measure as the 'reward-punishment sensitivity index' (Figure 2.3d).

Comparing parameters across sexes via Welch's t-tests revealed significant differences in reward sensitivity (t_{289} = –2.87, p = 0.004, d = 0.34; lower in females) and consequently reward-punishment sensitivity index (t_{336} = –2.03, p = 0.043, d = 0.22; lower in females, i.e. more avoidance-driven). In the replication sample, I observed the same effect on reward-punishment sensitivity index (t_{626} = –2.79, p = 0.005, d = 0.22; lower in females). However, the sex difference in reward sensitivity did not replicate (p = 0.441), although I did observe a significant sex difference in punishment sensitivity in the replication sample (t_{626} = 2.26, p = 0.024, d = 0.18).

2.5.5 Computational mechanisms of anxiety-related avoidance

I assessed the relationship between task-induced anxiety and individual model parameters through permutation testing of non-parametric correlations, since the distribution of task-induced anxiety was non-Gaussian. The punishment learning rate and reward-punishment sensitivity index were significantly and negatively correlated with task-induced anxiety in the discovery sample (punishment learning rate: Kendall's τ = –0.088, p = 0.015; reward-punishment sensitivity index: τ = –0.099, p = 0.005; Figure 2.4a and b). These associations were also significant in the replication sample (punishment learning rate: τ = –0.064, p = 0.020; reward-punishment sensitivity index: τ = –0.096, p < 0.001), with an additional positive correlation between punishment sensitivity and anxiety (τ = 0.076, p = 0.004; Appendix 2.1—table 1) which may have been detected due to greater statistical power. The learning rate equivalent of the reward-punishment sensitivity index (i.e. α^r/α^p) did not significantly correlate with task-induced anxiety (discovery sample: τ = 0.048, p = 0.165; replication sample: τ = 0.029, p = 0.298).

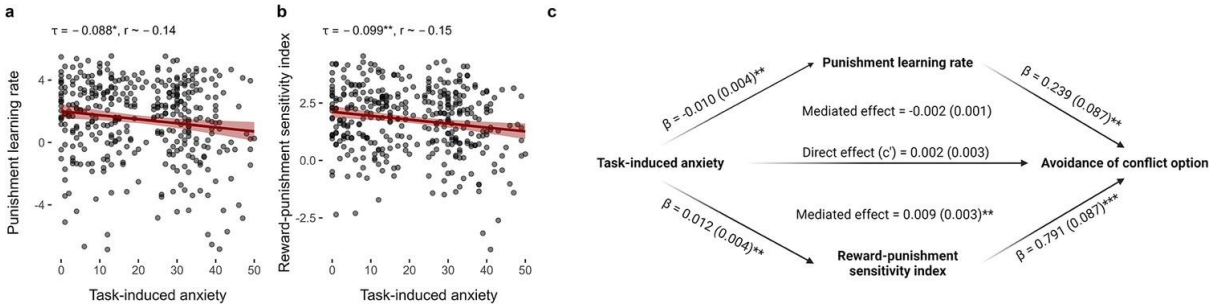


Figure 2.4, Relationships between task-induced anxiety, model parameters, and avoidance. *(a) Task-induced anxiety was negatively correlated with the punishment learning rate. (b) Task-induced anxiety was also negatively correlated with reward-punishment sensitivity index. Kendall's tau correlations and approximate Pearson's r equivalents are reported above each figure (n = 369). (c) The mediation model. Mediation effects were assessed using structural equation modelling. Bold terms represent variables and arrows depict regression paths in the model. The annotated values next to each arrow show the regression coefficient associated with that path, denoted as coefficient (standard error). Only the reward-punishment sensitivity index significantly mediated the effect of task-induced anxiety on avoidance. Significance levels in all figures are shown according to the following: p < 0.05 – *; p < 0.01 – **; p < 0.001 – ***.*

Given that both the punishment learning rate and the reward-punishment sensitivity index were significantly and reliably correlated with task-induced anxiety, I next asked whether these computational measures explained the effect of anxiety on avoidance behaviour. I tested this via structural equation modelling, including both the punishment learning rate and reward-punishment sensitivity index as parallel mediators for the effect of task-induced anxiety on proportion of choice for conflict/safe options. This suggested that the mediating effect of the reward-punishment sensitivity index (standardised β = –0.009 ± 0.003, p = 0.003) was over four times larger than the mediating effect of the punishment learning rate, which did not reach the threshold as a significant mediator (standardised β = –0.002 ± 0.001, p = 0.052; Figure 2.4c). A similar pattern was observed in the replication sample (mediating effect of reward-punishment sensitivity index: standardised β = –0.009 ± 0.002, p < 0.001; mediating effect of punishment learning rate: standardised β = –0.001 ± 0.0003, p = 0.132; Appendix 2.1—table 1, Appendix 2.3—figure 1).

This analysis reveals two effects: first, that individuals who reported feeling more anxious during the task were slower to update their estimates of punishment probability; and second, more anxious individuals placed a greater weight on the aversive sound relative to the reward when making their decisions, meaning that the potential for a reward was more strongly offset by potential for punishment. Critically, the latter but not the former effect explained the effect of task-induced anxiety on avoidance behaviour, thus providing an important insight into the computational mechanisms driving anxiety-related avoidance.

2.5.6 Psychiatric symptoms predict task-induced anxiety, but not avoidance

Given that I recruited samples without any inclusion/exclusion criteria relating to mental health, the instances of clinically relevant anxiety symptoms, defined as a score of 10 or greater in the

Generalised Anxiety Disorder 7-item scale (Spitzer et al., 2006), were low in both the discovery (16%, 60 participants) and the replication sample (20%, 127 participants).

Anxiety, depression, and experiential avoidance symptoms were all significantly and positively correlated with task-induced anxiety ($p < 0.001$; Appendix 2.1—table 1). Only experiential avoidance symptoms were significantly correlated with the proportion of conflict/safe option choices ($\tau = -0.059$, $p = 0.010$; Appendix 2.1—table 1), but this effect was not observed in the replication sample ($\tau = -0.029$, $p = 0.286$). Experiential avoidance was also significantly correlated with punishment learning rates (Kendall's $\tau = -0.09$, $p = 0.024$) and the reward-punishment sensitivity index ($\tau = -0.09$, $p = 0.034$) in the discovery sample. However, these associations were non-significant in the replication sample ($p > 0.3$), and no other significant correlations were detected between the psychiatric symptoms and model parameters (Appendix 2.1—table 1).

2.5.7 Split-half reliability

Overall choice proportion showed good reliability (discovery sample $r = 0.63$; replication sample $r = 0.63$; Appendix 2.5—figure 1). The model parameters showed good-to-excellent reliability (model-derived r values ranging from 0.61 to 0.85 [0.76–0.92 after Spearman-Brown correction]; Appendix 2.5—figure 1).

2.5.8 Test-retest reliability

Task-induced anxiety, overall proportion of conflict option choices, and the model parameters demonstrated fair to excellent reliability (reliability indices ranging from 0.4 to 0.77; Appendix 2.6—figure 1), except for the punishment learning rate which showed a model-derived reliability of $r = -0.45$. Practice effects analyses showed that task-induced anxiety and the punishment learning rate were also significantly lower in the second session (anxiety: $t_{56} = 2.21$, $p = 0.031$; punishment learning rate: $t_{56} = 2.24$, $p = 0.029$), whilst the overall probability of choosing the conflict option and the other parameters did not significantly change over sessions (Appendix 2.7—figure 1). As the dynamic (latent) outcome probabilities were different across the test vs. retest sessions, in other words participants were not performing the exact same task across sessions, arguably even 'fair' reliability (i.e. estimates in range of 0.4–0.6) is encouraging, as this suggests that model parameters generalise across different environments if the required computations are preserved. However, potential within-subjects inferences based on the punishment learning rate may not be justified in this task.

2.5.9 Sensitivity analyses

As the procedure for estimating model parameters (the expectation maximisation (EM) algorithm, see 'Methods') produced some high inter-parameter correlations in my data (Appendix 2.2—figure

2), I also re-estimated the parameters using Stan's variational Bayesian inference (VBI) algorithm (Stan Development Team, 2023) – this resulted in lower inter-parameter correlations, but my primary computational finding, that the effect of anxiety on choice is mediated by relative sensitivity to reward/punishment was consistent across algorithms (see Appendix 2.8 for details).

Since participants self-determined the volume of the punishment stimuli in the task, and therefore (at least in part) their aversiveness, I also conducted a sensitivity analysis by accounting for self-reported unpleasantness ratings of the punishment (see Appendix 2.9). My finding that anxiety impacts approach-avoidance behaviour was robust to this sensitivity analysis ($p < 0.001$), however the mediating effect of the reward-punishment sensitivity index was not ($p > 0.1$; see Appendix 2.9 for details). Although sound volume and stimulus unpleasantness correlate (Seow & Hauser, 2022), unpleasantness ratings are not a perfect indicator of volume, so this analysis should be interpreted with caution.

2.6 Discussion

To develop a translational and computational model of anxiety-related avoidance, I created a novel task that involves learning under approach-avoidance conflict and investigated the impact of anxiety on task performance. Importantly, I show that individuals who experienced more task-induced anxiety avoid actions that lead to aversive outcomes, at the cost of potential reward. Computational modelling of behaviour revealed that this was explained by relative sensitivity to punishment over reward. However, this effect was limited to task-induced anxiety, and was only evident on a self-report measure of general avoidance in my first sample. Finally, split-half and test-retest analyses of the task indicated fair-to-excellent reliability indices of task behaviour and model parameters. These results demonstrate the potential of approach-avoidance reinforcement learning tasks as translational measures of anxiety-related avoidance.

2.6.1 Task validity

I evaluated the translational validity of the task based on the criteria of face, construct, predictive, and computational validity (Redish et al., 2022; Willner, 1984). The task shows sufficient face validity when compared to the Vogel/Geller-Seifter conflict tests; the design involves a series of choices between an action associated with high reward and punishment, or low reward and no punishment, and these contingencies must be learned. One potentially important difference is that my task involved a choice between two active options, referred to as a two-alternative forced choice (2AFC), whereas in the Vogel/Geller-Seifter tests, animals are free to perform an action (e.g. press the lever) or withhold the action (don't press the lever, i.e. a go/no-go design). The disadvantage of the latter approach is that human participants can and typically do show motor response biases (i.e. for either

action or inaction), which can confound value-based response biases with approach or avoidance (Guitart-Masip et al., 2012). Fortunately, there is ample evidence that rodents can perform both 2AFC learning (Metha et al., 2020) and conflict tasks (Glover et al., 2020; Oberrauch et al., 2019; Simon et al., 2009). That avoidance responses were positively associated with task-induced anxiety also lends support for the construct validity of my measure, given the theoretical links between anxiety and avoidance under approach-avoidance conflict (Aupperle & Paulus, 2010; Hayes, 1976; Stein & Paulus, 2009). However, the empirical effect size was small (|r|~0.13) – I discuss potential reasons for this below.

The predictive and computational validity of the task need to be determined in future work. Predictive validity can be assessed by investigating the effects of anxiolytic drugs on task behaviour, as described in the subsequent chapter. Reduced avoidance after anxiolytic drug administration, or indeed any manipulation designed to reduce anxiety, would support the measure's predictive validity. Similarly, computational validity would also need to be assessed directly in non-human animals by fitting models to their behavioural data. This should be possible even in the face of different procedures across species such as number of trials or outcomes used (shock or aversive sound). I am encouraged by my finding that the winning computational model in my study relies on a relatively simple classical reinforcement learning strategy. There exist many studies showing that non-human animals rely on similar strategies during reward and punishment learning (Mobbs et al., 2020; Schultz, 2013); albeit to my knowledge this has never been modelled in non-human animals where rewards and punishment can occur simultaneously.

2.6.2 Implications

One benefit of designing a task amenable to the computational modelling of behaviour was that I could extract model parameters that represent individual differences in precise cognitive mechanisms. The model recapitulated individual- and group-level behaviour with high fidelity, and mediation analyses suggested a potential mechanism of how anxiety may induce avoidance, namely through the relative sensitivity to punishment compared to reward. This demonstrates the potential of computational methods in allowing for mechanistic insights into anxiety-related avoidance. Further, I found satisfactory test-retest reliability of the reward-punishment sensitivity index in test-retest analyses, suggesting that the task can be used for within-subject experiments (e.g. on/off medication) to assess the impact of interventions on the mechanisms underlying avoidance.

The composite nature of the reward-punishment sensitivity index may be important for contextualising prior results of anxiety and outcome sensitivity. The traditional view has long been that anxiety is associated with greater automatic sensitivity to punishments and threats (i.e. that

anxious individuals show a negative bias in information processing; Bishop, 2007), due to amygdala hypersensitivity (Hartley & Phelps, 2012; Mathews & Mackintosh, 1998). However, subsequent studies have found inconsistent findings relative to this hypothesis (Bishop & Gagne, 2018). I found that, rather than a straightforward link between anxiety and punishment sensitivity, the relative sensitivity to punishment over reward provided a better explanation of anxiety's effect on behaviour. This effect is more in keeping with cognitive accounts of anxiety's effect on information processing that involve higher-level cognitive functions such as attention and cognitive control; these accounts propose that anxiety involves a global reallocation of resources away from processing positive and towards negative information (Beck & Clark, 1988; Cisler & Koster, 2010; Robinson et al., 2013), possibly through prefrontal control mechanisms (Carlisi & Robinson, 2018). Future studies using experimental manipulations of anxiety will be needed to test the causal role of anxiety on relative sensitivities to punishments versus rewards.

Since I developed my task with the primary focus on translational validity, its design diverges from other reinforcement learning tasks that involve reward and punishment outcomes (Pike & Robinson, 2022). One important difference is that I used distinct reinforcers as my reward and punishment outcomes, compared to many studies which use monetary outcomes for both (e.g. earning and losing £1 constitute the reward and punishment, respectively; Aylward et al., 2019; Jean-Richard-Dit-Bressel et al., 2021; Pizzagalli et al., 2005; Sharp et al., 2022). Other tasks have been used that induce a conflict between value and motor biases, relying on prepotent biases to approach/move towards rewards and withdraw from punishments, which makes it difficult to approach punishments and withdraw from rewards (Guitart-Masip et al., 2012; Mkrtchian et al., 2017). However, since translational operant conflict tasks typically induce a conflict between different types of outcome (e.g. food and shocks/sugar and quinine pellets; Oberrauch et al., 2019; van den Bos et al., 2014), I felt it was important to implement this feature. One study used monetary rewards and shock-based punishments, but also included four options for participants to choose from on each trial, with rewards and punishments associated with all four options (Seymour et al., 2007). This effectively requires participants to maintain eight probability estimates (i.e. reward and punishment at each of the four options) to solve the task, which may be too difficult for non-human animals to learn efficiently.

Notably, although a recent computational meta-analysis of reinforcement learning studies showed that symptoms of anxiety and depression are associated with elevated punishment learning rates (Pike & Robinson, 2022), I did not observe this pattern in my data. Indeed, I even found the contrary effect in relation to task-induced anxiety, specifically that anxiety was associated with lower rates of learning from punishment. However, other work has suggested that the direction of this effect can

depend on the form of anxiety, where cognitive anxiety may be associated with elevated learning rates, but somatic anxiety may show the opposite pattern (Wise & Dolan, 2020) and this may explain the discrepancy in findings. Additionally, parameter values are highly dependent on task design (Eckstein et al., 2022), and study designs to date may be more optimised in detecting differences in learning rate (Pike & Robinson, 2022) – further work is needed to better understand the potentially complex association between anxiety and punishment learning rate. Lastly, as punishment learning rate was severely unreliable in the test-retest analyses, and the associations between punishment learning rate and state anxiety were not robust to an alternative method of parameter estimation (VBI), the negative correlation observed in my study should be treated with caution.

2.6.3 Limitations/opportunities for development

One potentially important limitation of my findings is the small effect size observed in the correlation between task-induced anxiety and avoidance (Kendall's tau values < 0.1, mediation coefficients < 0.01). This may be attributed to the simplicity of using overall choice proportion as a measure of approach/avoidance, as the effect of anxiety on choice was also influenced by punishment probability. This may have also been due to the aversiveness of the punishment outcomes (aversive sounds) used in my study, compared to other studies which included aversive images (Aupperle et al., 2011), electric shocks (Talmi et al., 2009), or used virtual reality (Biedermann et al., 2017) to deliver negative reinforcement. Using more salient punishments would likely produce larger effects. Further, whilst I included tests to check whether participants were wearing headphones for this online study, I could not be certain that the sounds were not played over speakers and/or played at a sufficiently high volume to be aversive. Finally, it may simply be that the relationships between psychiatric symptoms and behaviour are small due to the inherent variability of individuals, the lack of precision of self-reported symptoms, or the unreliability of mental health diagnoses (Freedman et al., 2013; Wise et al., 2022). Despite these limitations, an advantage of using web-based auditory stimuli is their efficiency in scaling with sample size, as only a web-browser is needed to complete the task. This is especially relevant for psychiatric research, given the need for studies with larger sample sizes in the field (Rutledge et al., 2019).

Relatedly, participants had some control over the intensity at which the punishments were presented, which may have driven my findings relating to anxiety and putative mechanisms of anxiety-related avoidance. Sensitivity analyses showed that my finding that anxiety is positively associated with avoidance in the task was robust to individual differences in self-reported punishment unpleasantness, whilst the mediation effects were not. Future work imposing better control over the stimuli presented, and/or using within-subjects designs will be needed to validate the role of reward/punishment sensitivities in anxiety-related avoidance.

The punishment learning rate was severely unreliable given a model-derived correlation across sessions of r = –0.45, whereas the other indices of task performance showed fair-to-excellent reliability. This might be a result of practice effects, where participants learned about the punishments differently across sessions, but not about the rewards. Alternatively, as participants self-determined the loudness of the punishments, differences in volume settings across sessions may have impacted the reliability of this parameter (and indeed other measures). Further, the asymmetric nature of the task may have impacted my ability to estimate the punishment learning rate, as there were fewer occurrences of the punishment compared to the reward. However, given that the sensitivity parameters were more important for explaining how anxiety affected avoidance, the low observed reliability of the punishment learning rate may not present a barrier for future use of this task in repeated-measures designs.

I also did not detect any significant correlations between task performance and clinical symptoms of anxiety, which raises questions around the predictive validity of the task. As participants were recruited without imposing specific inclusion/exclusion criteria related to mental health, the low level of clinically relevant symptoms in the data may have contributed to these results. Another possibility is that the participants with higher anxiety levels may have been taking medication to manage their symptoms, which could have affected their task performance and obscured potential correlations with clinical measures. Further studies in samples with high levels of anxiety or examining individuals at the extreme ends of symptom distributions may offer increased sensitivity in assessing the predictive validity of the task. Further, it is worth noting that many animal paradigms were developed and widely adopted due to their sensitivity to anxiolytic medication (Cryan & Holmes, 2005). Given the lack of associations with clinical measures in my results, it is possible that current translational models of anxiety may not fully capture behaviours that are directly relevant to pathological anxiety. To develop translational paradigms of clinical utility, future research should place a stronger emphasis on assessing their clinical validity in humans. Using multi-level or continuous outcomes would also improve the ecological validity of the present approach and interpretation of the sensitivity indices.

One limitation of the present analysis was the primary focus on computational modelling over more traditional, model-agnostic behavioural metrics. This was a deliberate choice, directly motivated by the overarching aim of this thesis which was to investigate the specific computational constructs engaged by translational tests of anxiety. By employing computational analyses, I sought to move beyond descriptive measures and study the latent cognitive mechanisms that might not be apparent in behavioural summary statistics. However, the interpretability of the model parameters would be strengthened with additional behavioural metrics, and so future work could incorporate the analysis

of other metrics such as reaction times or win-stay/lose-shift probabilities, to provide a more comprehensive understanding of task performance and to further validate the computational findings. For instance, slower reaction times for conflict choices might indicate greater deliberation or cognitive conflict experienced by more anxious individuals when faced with the decision of whether to approach potential punishment. Similarly, decreased win-stay probabilities following punishment could reflect a stronger impact of aversive outcomes on subsequent choices, suggesting heightened sensitivity to negative feedback.

Finally, whilst there is a broad literature on the roles of behavioural inhibition and avoidance tendency traits on decision-making and behaviour (Carver & White, 1994; Corr, 2004; Gray, 1982), I did not replicate the correlation of experiential avoidance and avoidance responses or the reward-punishment sensitivity index. Since there were also no significant correlations across task performance indices and clinical symptom measures, my findings suggest that the measure may be more sensitive to behaviours relating to state anxiety, rather more stable traits. Nevertheless, how performance in the present task relates to other traits such as behavioural approach/inhibition tendencies (Carver & White, 1994), as has been found in previous studies on reward/punishment learning (Sharp et al., 2022; Wise & Dolan, 2020) and approach-avoidance conflict (Aupperle et al., 2011), will be an important question for future work.

2.7 Conclusion

Avoidance is a hallmark symptom of anxiety-related disorders and there is a need for better translational models of avoidance. In a novel translational task designed to involve the same computational processes as those involved in non-human animal tests, anxiety was reliably induced, associated with avoidance behaviour, and linked to a computational model of behaviour. I hope that similar approaches will facilitate progress in the theoretical understanding of avoidance and ultimately aid in the development of novel anxiolytic treatments for anxiety-related disorders.

2.8 Appendices

2.8.1 Appendix 2.1: Comparison of results in the discovery and replication samples, with and without data cleaning exclusions

Appendix 2.1—table 1. Statistical results across the discovery and replication samples, and the effect of data cleaning exclusions

Statistical test/model	Sub-test	Sample (*without data cleaning*)	
		Discovery	**Discovery**
Hierarchical logistic regression; no task-induced anxiety	Reward coefficient	β = 0.98 ± 0.03, p < 0.001 *β = 0.95 ± 0.03, p < 0.001*	β = 0.95 ± 0.02, p < 0.001 *β = 0.90 ± 0.02, p < 0.001*
	Punishment coefficient	β = -0.33 ± 0.04, p < 0.001 *β = -0.29 ± 0.04, p < 0.001*	β = -0.52 ± 0.03, p < 0.001 *β = -0.50 ± 0.03, p < 0.001*
Distribution; task-induced anxiety		Mean = 21, SD = 14 *Mean = 21, SD = 14*	Mean = 22, SD = 13 *mean = 22, SD = 14*
Correlation; task-induced anxiety and choice		τ = -0.074, p = 0.033 *τ = -0.068, p = 0.036*	τ = -0.075, p = 0.005 *τ = -0.081, p = 0.002*
Hierarchical logistic regression; with task-induced anxiety	Reward coefficient	β = 0.98 ± 0.03, p < 0.001 *β = 0.95 ± 0.03, p < 0.001*	β = 0.95 ± 0.02, p < 0.001 *β = 0.90 ± 0.02, p < 0.001*
	Punishment coefficient	β = -0.33 ± 0.04, p < 0.001 *β = -0.29 ± 0.04, p < 0.001*	β = -0.52 ± 0.03, p < 0.001 *β = -0.50 ± 0.03, p < 0.001*
	Task-induced anxiety coefficient	β = -0.04 ± 0.04, p = 0.304 *β = -0.03 ± 0.04, p = 0.377*	β = 0.02 ± 0.03, p = 0.637 *β = 0.003 ± 0.03, p = 0.925*
	Interaction of task-induced anxiety and reward	β = -0.06 ± 0.03, p = 0.074 *β = -0.05 ± 0.03, p = 0.076*	β = -0.003 ± 0.02, p = 0.901 *β = -0.006 ± 0.02, p = 0.774*
	Interaction of task-induced anxiety and punishment	β = -0.10 ± 0.04, p = 0.022 *β = -0.10 ± 0.04, p = 0.011*	β = -0.23 ± 0.03, p < 0.001 *β = -0.23 ± 0.03, p < 0.001*
Hierarchical logistic regression; with task-induced anxiety, no interaction	Reward coefficient	β = 0.99 ± 0.03, p < 0.001 *β = 0.95 ± 0.03, p < 0.001*	β = 0.95 ± 0.02, p < 0.001 *β = 0.89 ± 0.02, p < 0.001*
	Punishment coefficient	β = -0.33 ± 0.04, p < 0.001 *β = -0.29 ± 0.04, p < 0.001*	β = -0.52 ± 0.03, p < 0.001 *β = -0.50 ± 0.03, p < 0.001*
	Task-induced anxiety coefficient	β = -0.09 ± 0.03, p = 0.012 *β = -0.08 ± 0.03, p = 0.016*	β = -0.08 ± 0.03, p = 0.005 *β = -0.10 ± 0.03, p < 0.001*
Computational model comparison	Winning model	2 learning rate, 2 sensitivity *2 learning rate, 2 sensitivity*	2 learning rate, 2 sensitivity *2 learning rate, 2 sensitivity*
Correlation; task-induced anxiety and model parameters	Reward learning rate	τ = -0.019, p = 0.596 *τ = -0.006, p = 0.847*	τ = -0.010, p = 0.749 *τ = -0.012, p = 0.637*
	Punishment learning rate	τ = -0.088, p = 0.015	τ = -0.064, p = 0.019

		τ = -0.073, p = 0.026	*τ = -0.074, p = 0.003*
	Learning rate ratio	τ = 0.048, p = 0.175 *τ = -0.037, p = 0.276*	τ = 0.029, p = 0.282 *τ = 0.045, p = 0.072*
	Reward sensitivity	τ = -0.038, p = 0.286 *τ = -0.047, p = 0.149*	τ = -0.028, p = 0.285 *τ = -0.023, p = 0.345*
	Punishment sensitivity	τ = 0.068, p = 0.051 *τ = 0.047, p = 0.153*	τ = 0.076, p = 0.004* *τ = 0.094, p < 0.001*
	Reward-punishment sensitivity index	τ = -0.099, p = 0.005 *τ = -0.084, p = 0.011*	τ = -0.096, p < 0.001 *τ = -0.103, p < 0.001*
Mediation	Punishment learning rate	β = -0.002 ± 0.001, p = 0.052 *β = -0.001 ± 0.001, p = 0.222*	β = -0.001 ± 0.003, p = 0.132 *β = -0.001 ± 0.003, p = 0.031*†
	Reward-punishment sensitivity index	β = 0.009 ± 0.003, p = 0.003 *β = 0.006 ± 0.002, p = 0.011*	β = 0.009 ± 0.002, p < 0.001 *β = 0.009 ± 0.002, p < 0.001*
Correlation; psychiatric symptoms and choice (overall proportion of conflict option choices)	GAD7	τ = -0.026, p = 0.458 *τ = 0.005, p = 0.894*	τ = -0.001, p = 0.988 *τ = -0.002, p = 0.919*
	PHQ8	τ = -0.02, p = 0.579 *τ = 0.014, p = 0.684*	τ = -0.013, p = 0.639 *τ = -0.012, p = 0.646*
	BEAQ	τ = -0.059, p = 0.010 *τ = -0.048, p = 0.151*†	τ = -0.029, p = 0.286* *τ = -0.020, p = 0.423*
Correlation; psychiatric symptoms and task-induced anxiety	GAD7	τ = 0.256, p < 0.001 *τ = 0.267, p < 0.001*	τ = 0.222, p < 0.001 *τ = 0.231, p < 0.001*
	PHQ8	τ = 0.233, p < 0.001 *τ = 0.244, p < 0.001*	τ = 0.184, p < 0.001 *τ = 0.194, p < 0.001*
	BEAQ	τ = 0.15, p < 0.001 *τ = 0.172, p < 0.001*	τ = 0.176, p < 0.001 *τ = 0.17, p < 0.001*
Correlation; GAD7 and model parameters	Reward learning rate	τ = -0.06, p = 0.076 *τ = -0.061, p = 0.074*	τ = -0.03, p = 0.304 *τ = -0.028, p = 0.264*
	Punishment learning rate	τ = -0.07, p = 0.077 *τ = -0.037, p = 0.265*	τ = -0.01, p = 0.621 *τ = -0.017, p = 0.534*
	Learning rate ratio	τ = -0.001, p = 0.999 *τ = -0.032, p = 0.35*	τ = -0.034, p = 0.229 *τ = -0.026, p = 0.301*
	Reward sensitivity	τ = -0.01, p = 0.802 *τ = -0.005, p = 0.877*	τ = -0.01, p = 0.745 *τ = -0.009, p = 0.718*
	Punishment sensitivity	τ = 0.05, p = 0.154 *τ = 0.022, p = 0.513*	τ = 0.003, p = 0.907 *τ = 0.012, p = 0.633*
	Reward-punishment sensitivity index	τ = -0.05, p = 0.149 *τ = -0.023, p = 0.496*	τ = 0.01, p = 0.746 *τ = 0.007, p = 0.797*

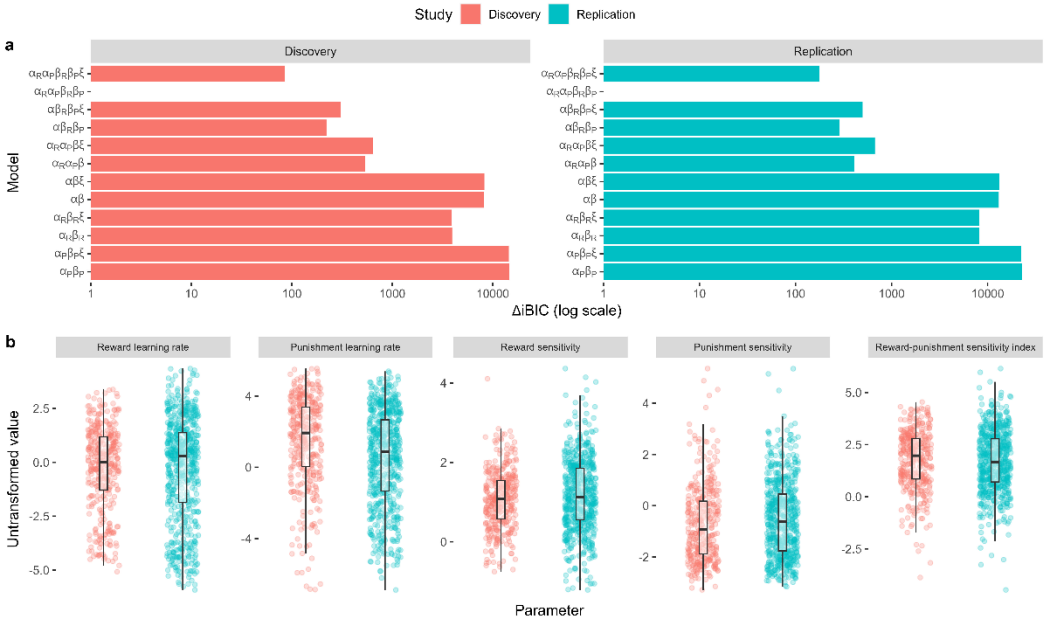
Correlation; PHQ8 and model parameters	Reward learning rate	τ = -0.03, p = 0.396 *τ = -0.04, p = 0.239*	τ = -0.03, p = 0.219 *τ = -0.035, p = 0.17*
	Punishment learning rate	τ = -0.06, p = 0.100 *τ = -0.022, p = 0.504*	τ = -0.01, p = 0.630 *τ = -0.025, p = 0.348*
	Learning rate ratio	τ = 0.008, p = 0.809 *τ = -0.038, p = 0.259*	τ = -0.033, p = 0.231 *τ = -0.016, p = 0.521*
	Reward sensitivity	τ = -0.012, p = 0.610 *τ = -0.009, p = 0.801*	τ = 0.01, p = 0.729 *τ = 0.004, p = 0.901*
	Punishment sensitivity	τ = 0.05, p = 0.179 *τ = 0.008, p = 0.802*	τ = -0.002, p = 0.948 *τ = 0.012, p = 0.619*
	Reward-punishment sensitivity index	τ = -0.06, p = 0.123 *τ = -0.01, p = 0.773*	τ = 0.02, p = 0.557 *τ = 0.005, p = 0.867*
Correlation; BEAQ and model parameters	Reward learning rate	τ = -0.06, p = 0.085 *τ = -0.047, p = 0.147*	τ = -0.02, p = 0.394 *τ = -0.017, p = 0.503*
	Punishment learning rate	τ = -0.08, p = 0.024 *τ = -0.071, p = 0.032*	τ = -0.03, p = 0.337* *τ = -0.031, p = 0.228*
	Learning rate ratio	τ = 0.018, p = 0.618 *τ = 0.002, p = 0.938*	τ = -0.005, p = 0.883 *τ = 0.007, p = 0.759*
	Reward sensitivity	τ = -0.01, p = 0.739 *τ = -0.036, p = 0.29*	τ = 0.01, p = 0.753 *τ = 0.002, p = 0.927*
	Punishment sensitivity	τ = 0.07, p = 0.061 *τ = 0.05, p = 0.127*	τ = 0.02, p = 0.477 *τ = 0.025, p = 0.308*
	Reward-punishment sensitivity index	τ = 0.02, p = 0.034 *τ = -0.073, p = 0.028*	τ = -0.01, p = 0.745* *τ = -0.015, p = 0.548*

Note. All correlations tests were conducted using permutation-based Kendall's tau correlations, using 10,000 permutations per test. Symbols represent the following: * - Inconsistent findings across discovery replication samples. † - Inconsistent findings before and after data cleaning exclusions. Abbreviations: GAD7 – Generalised Anxiety Disorder 7-item scale; PHQ8 – Patient Health Questionnaire 8-item depression scale; BEAQ – Brief Experiential Avoidance Questionnaire.

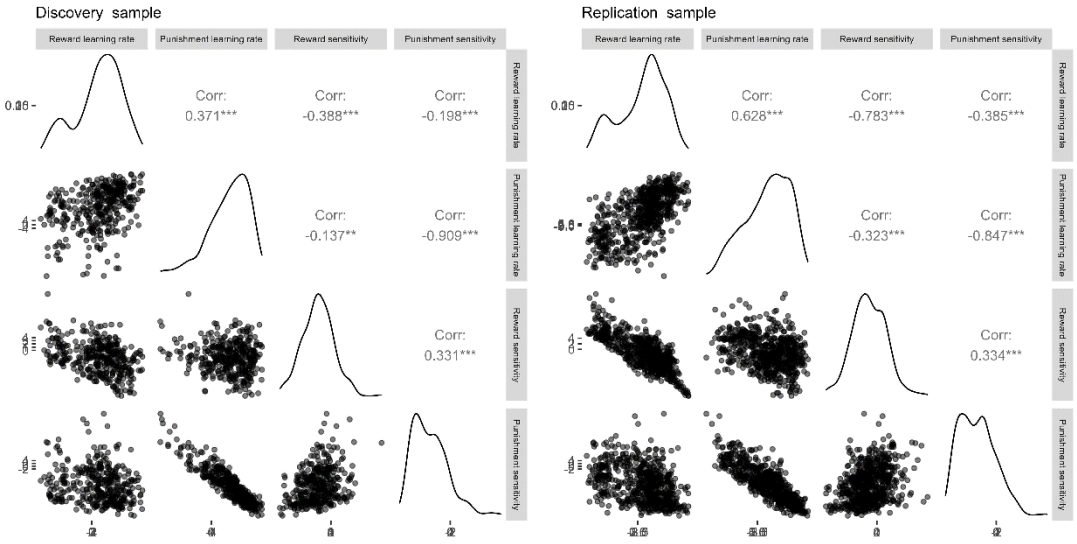
A sensitivity analyses examining the impact of implementing my data cleaning exclusions showed that the significant and negative correlation between experiential avoidance symptoms and overall proportion of conflict option choices (τ = -0.059, *p* = 0.010) was not significant when the data cleaning exclusions were omitted (τ = -0.029, *p* = 0.286); however since this effect was not replicated in the independent sample (with and without data exclusions), I do not discuss this further. The mediating effect of the punishment learning rate was also significant in the replication sample only

before excluding data, however the effect sizes were small and similar before and after excluding data, and so I also do not discuss this further.

2.8.2 Appendix 2.2: Computational modelling

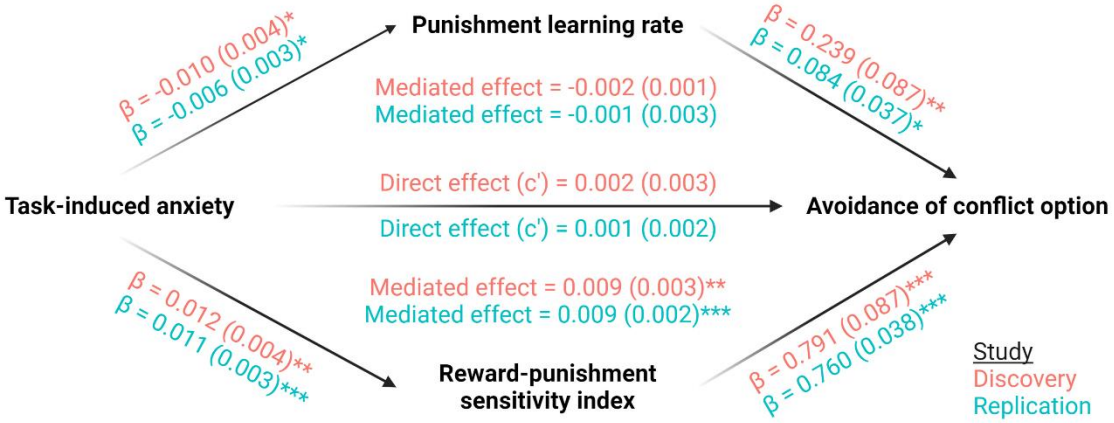


Appendix 2—figure 1. Model comparison and parameter distributions across studies. *a, Model comparison results. The difference in integrated Bayesian Information Criterion scores from each model relative to the winning model is indicated on the x-axis. The winning model in both studies included specific learning rates for reward (α^R) and punishment learning (α^p), and specific outcome sensitivity parameters for reward (β^R) and punishment (β^P). Some models were tested with the inclusion of a lapse term (ξ). b, Distributions of individual parameter values from the winning model across studies. The reward-punishment sensitivity index constituted my computational measure of approach-avoidance bias, calculated by taking the ratio between the reward and punishment sensitivity parameters.*



Appendix 2—figure 2. Correlation matrices for the estimated parameters across studies. *Lower-right diagonal of each matrix shows a scatterplot of cross-parameter correlations. Upper-right diagonal denotes the Pearson's r correlation coefficients for each pair of parameters, based on the untransformed parameter values.*

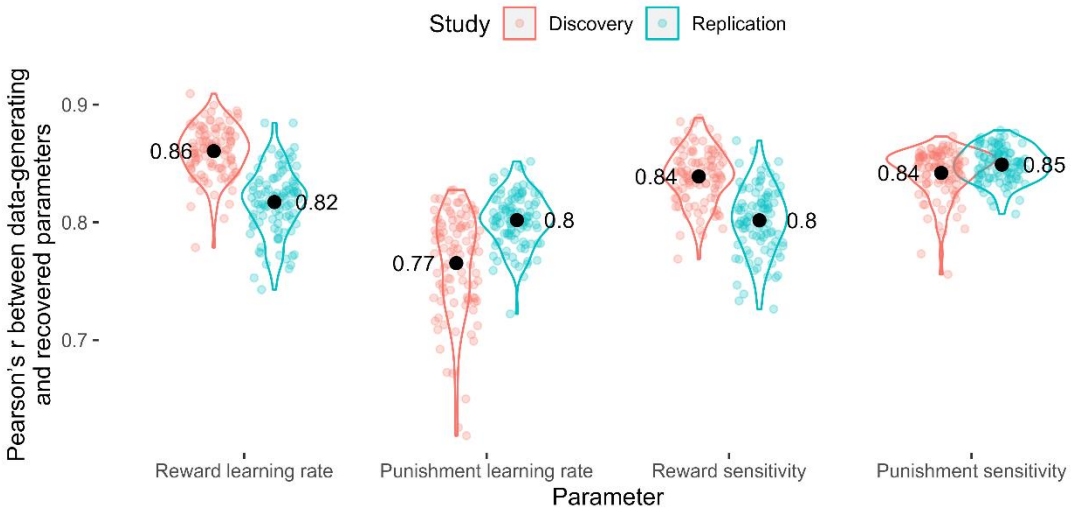
2.8.3 Appendix 2.3: Mediation analyses



Appendix 3—figure 1. Mediation analyses across studies. *Mediation effects were assessed using structural equation modelling. Bold terms represent variables and arrows depict regression paths in the model. The annotated values next to each arrow show the regression coefficient associated with that path, denoted as coefficient (standard error). Only the reward-punishment sensitivity index significantly mediated the effect of task-induced anxiety on avoidance. Significance levels in all figures are shown according to the following: p < 0.05 - *; p < 0.05 - **; p < 0.001 - ***.*

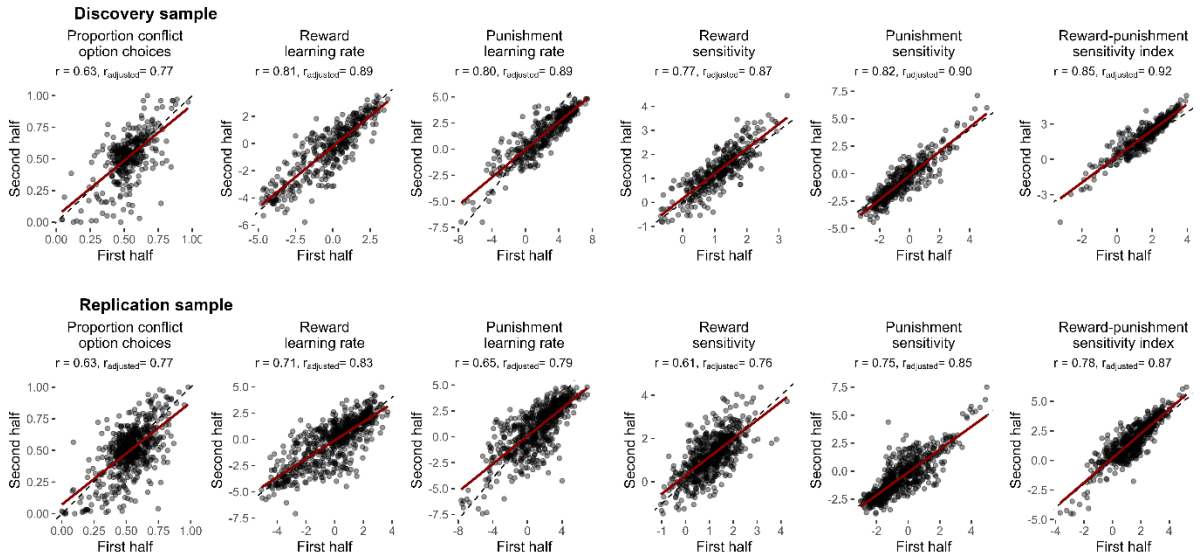
2.8.4 Appendix 2.4: Parameter recovery

To establish how well I could recover the parameters from the best fitting model, I simulated data from each study (discovery, replication) and refitted these new datasets with the identical model fitting procedure used for the empirical data. I matched the number of participants, number of trials per participant, and latent outcome probabilities in the simulated datasets to the empirical data, and used the best-fitting parameters for each participant to generate their choices. Trial-by-trial outcomes were generated stochastically according to the latent outcome probabilities, which meant that the simulated participants did not necessarily receive the same series of outcomes to the real participants. To average over this stochasticity in the outcomes, we repeated the simulations 100 times for each study. After fitting the best-fitting model to the simulated data, I computed Pearson's correlation coefficients across the data-generating parameters and the recovered parameters. Across all parameters, the average Pearson's r values between data-generating and recovered parameters ranged from 0.77 to 0.86, indicating excellent recovery (Appendix 2.4—figure 1).



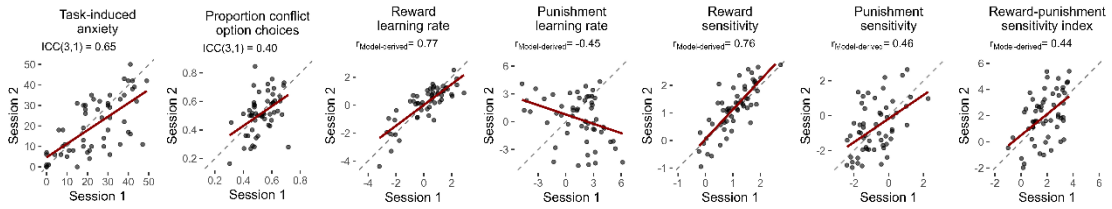
Appendix 4—figure 1. Parameter recovery. *Pearson's r values across the data-generating and recovered parameters by parameter. Coloured points represent Pearson's r values for each of 100 simulation iterations, and black points represent the mean value across simulations.*

2.8.5 Appendix 2.5: Split-half reliability



Appendix 2.5—figure 1. Split-half reliability of the task. *Reliability via Pearson's correlations of measures calculated from the first and second halves of the task are shown with their estimates of reliability. Reliability estimates for the computational measures from the winning computational model were computed by fitting split-half parameters within a single model, then using the parameter covariance matrix to derive Pearson's correlation coefficients for each parameter across halves. Reliability estimates are reported as unadjusted values (r) and after adjusting for reduced number of trials via Spearman-Brown correction ($r_{adjusted}$). Dotted lines represent the reference line, indicating perfect correlation. Red lines show lines-of-best-fit.*

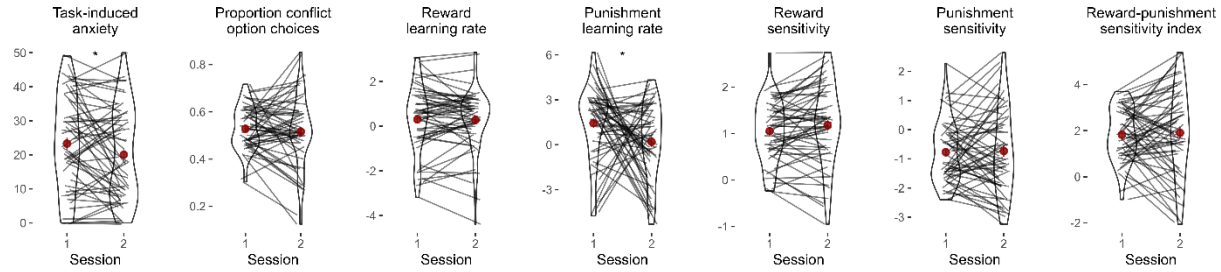
2.8.6 Appendix 2.6: Test-retest reliability



Appendix 6—figure 1. Test-retest reliability of the task. *Correlations of measures across the test and retest sessions are shown with their estimates of reliability. Reliability estimates for the model-agnostic measures (task-induced anxiety, proportion of conflict option choices) were estimated using intra-class correlation coefficients. Reliability estimates for the computational measures from the winning computational model were computed by first fitting both sessions' parameters within a single model, then using the parameter covariance matrix to derive a Pearson's correlation coefficient ($r_{Model-derived}$) for each parameter across sessions to be calculated from their covariance. Dotted lines represent the reference line, indicating perfect correlation. Red lines show lines-of-best-fit.*

2.8.7 Appendix 2.7: Practice effects

I checked for practise effects in the behavioural data by comparing the behavioural measures and computational model parameters across sessions via paired t-tests. These tests were all non-significant ($p > 0.07$) except for task-induced anxiety ($t_{56} = 2.21$, $p = 0.031$) and the punishment learning rate ($t_{56} = 2.24$, $p = 0.029$; Appendix 2.7—figure 1).



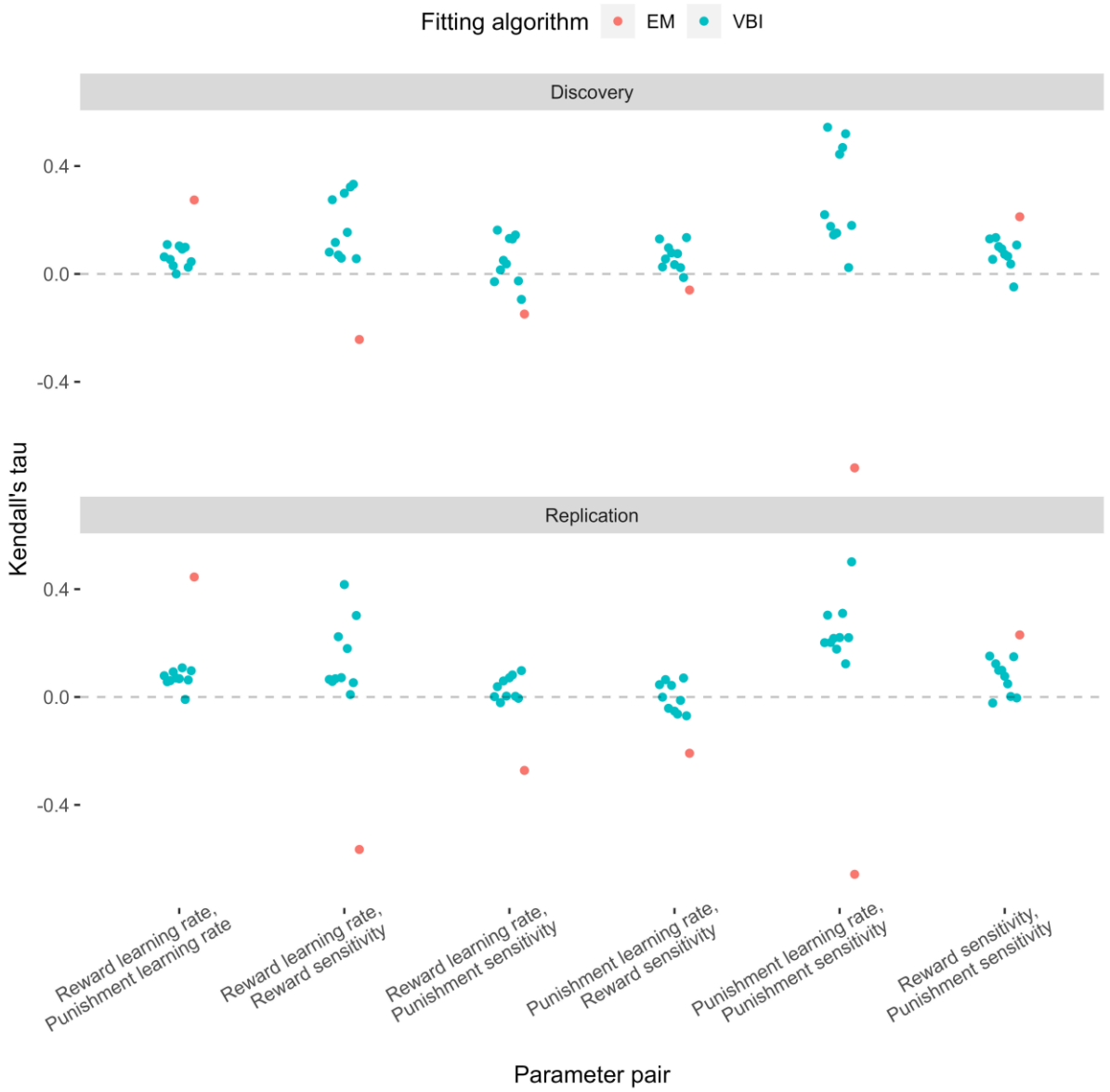
Appendix 7—figure 1. Task practice effects. *Comparison of behavioural measures and model parameters across time. Lines represent individual data, red points represent mean values, and red lines represent standard error bars. P-values of paired t-tests are annotated above each plot. Task-induced anxiety and the punishment learning rate was significantly lower in the second session, whilst the other measures did not change significantly across sessions. Significance levels are shown according to the following: p < 0.05 - *.*

2.8.8 Appendix 2.8: Sensitivity analysis: estimating parameters via expectation maximisation and variational Bayesian inference algorithms

Given that the expectation maximisation (EM) algorithm produced high inter-parameter correlations, I ran a sensitivity analysis by assessing the robustness of my computational findings to an alternative method of parameter estimation – (mean-field) variational Bayesian inference (VBI) via Stan (Stan Development Team, 2023). Since, unlike EM, the results of VBI are very sensitive to initial values, I fitted the data 10 times with different initial values.

Inter-parameter correlations

The VBI produced lower inter-parameter correlations than the EM algorithm (Appendix 2.8—figure 1).



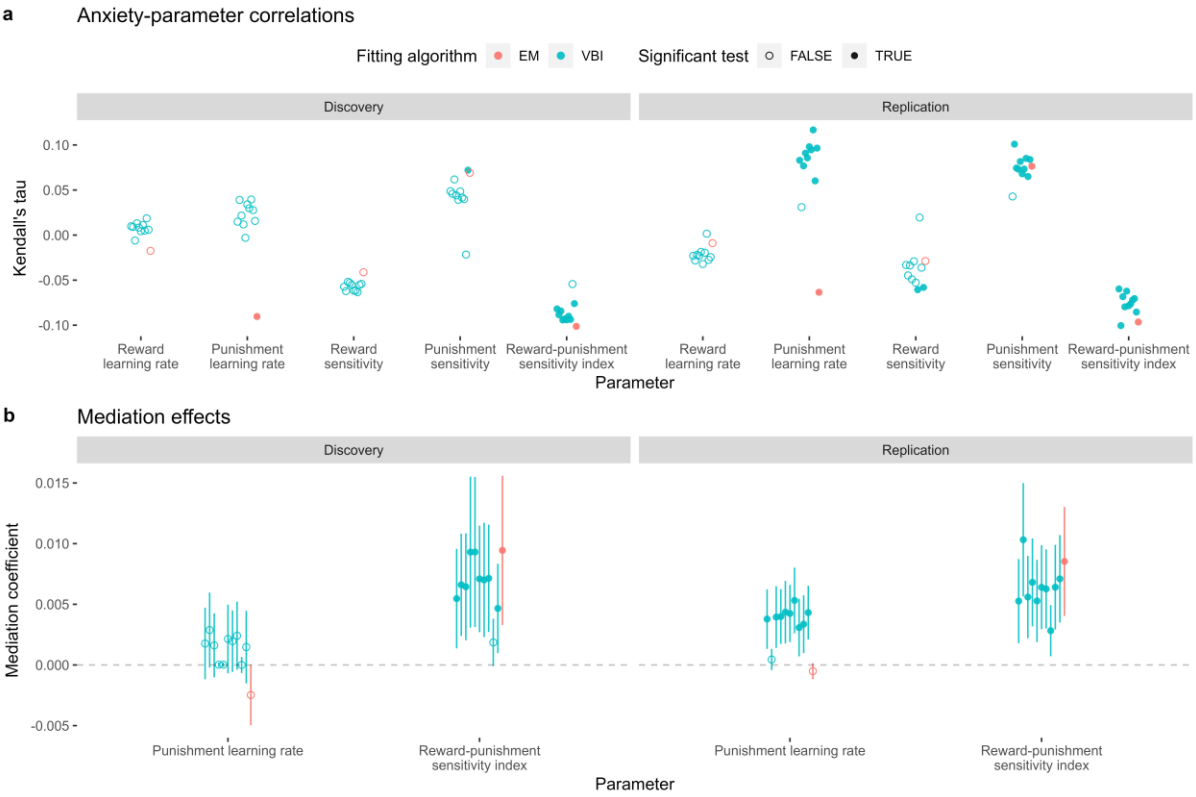
Appendix 2.8—figure 1. Inter-parameter correlations across the expectation-maximisation (EM, red) and variational Bayesian inference (VBI, blue) algorithms. *Overall, the VBI algorithm produced lower correlations compared to EM.*

Sensitivity analysis

Since multicollinearity in the VBI-estimated parameters was lower than for EM, indicating less trade-off in the estimation, I re-tested my computational findings as part of a sensitivity analysis. I first assessed whether I observed the same correlations between task-induced anxiety and punishment learning, and reward-punishment sensitivity index (Appendix 2.8—figure 2a). Punishment learning rate was not significantly associated with task-induced anxiety in any of the 10 VBI iterations in the discovery sample, although it was in 9/10 in the replication sample. On the other hand, the reward-punishment sensitivity index was significantly associated with task-induced anxiety in 9/10 VBI iterations in the discovery sample and all iterations in the replication sample. This suggests that the correlation of anxiety and sensitivity index is robust to these two fitting approaches.

I also re-estimated the mediation models, where in the EM-estimated parameters, I found that the reward-punishment sensitivity index mediated the relationship between task-induced anxiety and task choice proportions (Appendix 2.8—figure 2b). Again, I found that the reward-punishment sensitivity index was a significant mediator in 9/10 VBI iterations in the discovery sample and all iterations in the replication sample. Punishment learning rate was also a significant mediator in 9/10 iterations in the replication sample, although it was not in the discovery sample for all iterations, and this was not observed for the EM-estimated parameters.

Overall, I found that my key results, that anxiety is associated with greater sensitivity to punishment over reward, and this mediates the relationship between anxiety and approach-avoidance behaviour, were robust across both fitting methods.

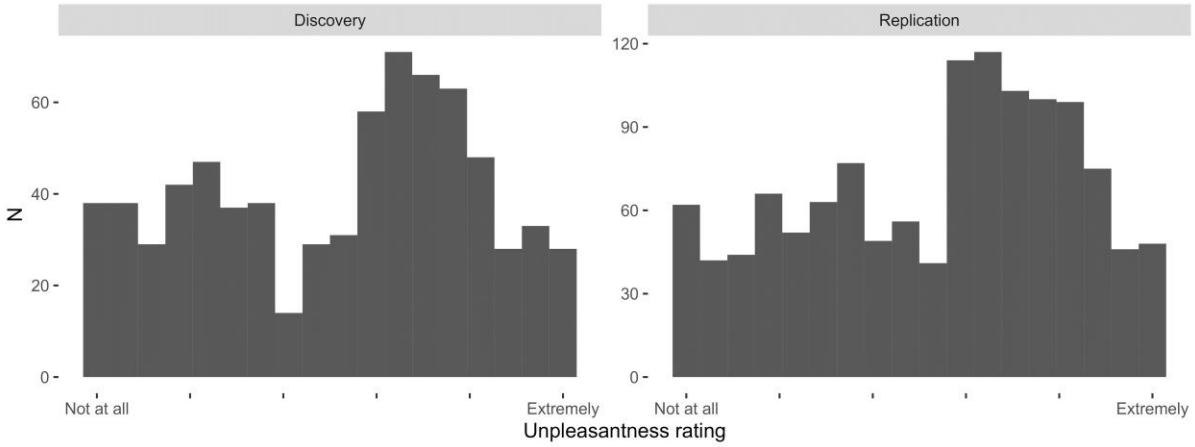


Appendix 2.8—figure 2. Sensitivity analysis of the computational findings relating to task-induced anxiety; comparing results when using parameters estimated via expectation maximisation (EM, red) and variational Bayesian inference (VBI, blue). *a, Kendall's tau correlations across each parameter and task-induced anxiety. b, Mediating effects of the punishment learning rate and reward-punishment sensitivity index.*

2.8.9 Appendix 2.9: Sensitivity analysis: punishment unpleasantness

Distribution of unpleasantness

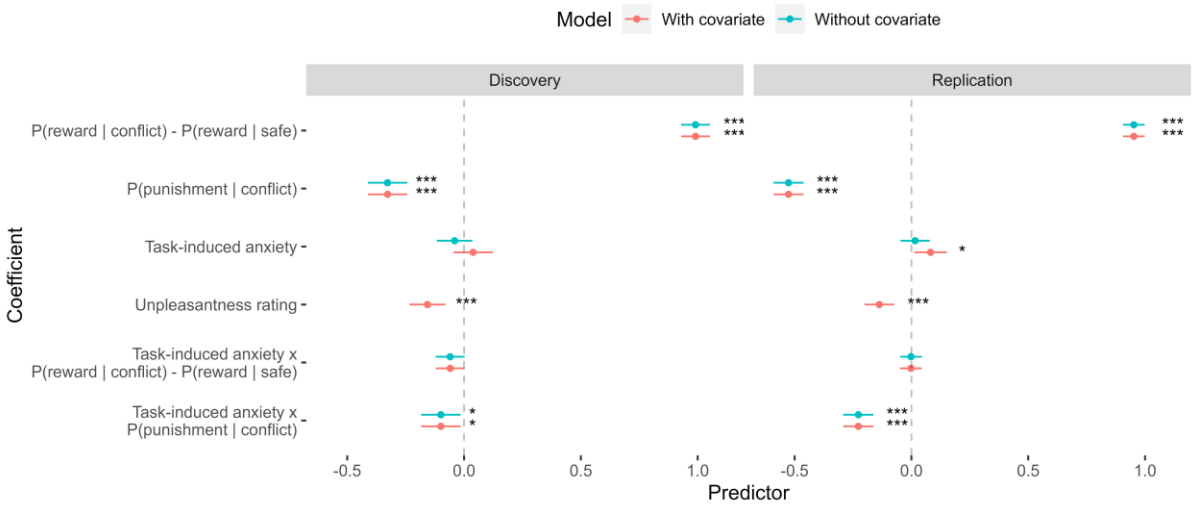
The punishments were rated as unpleasant by the participants, on average (discovery sample: mean rating = 31.1 [scored between 0 and 50], SD = 13.1; replication sample: mean rating = 32.1, SD = 12.7; Appendix 2.9—figure 1).



Appendix 2.9—figure 1. Distribution of self-reported punishment unpleasantness ratings. *Ratings were scored from 'Not at all' to 'Extremely' (encoded as 0 and 50, respectively). Distributions are shown across the discovery and replication samples.*

Approach-avoidance hierarchical logistic regression model

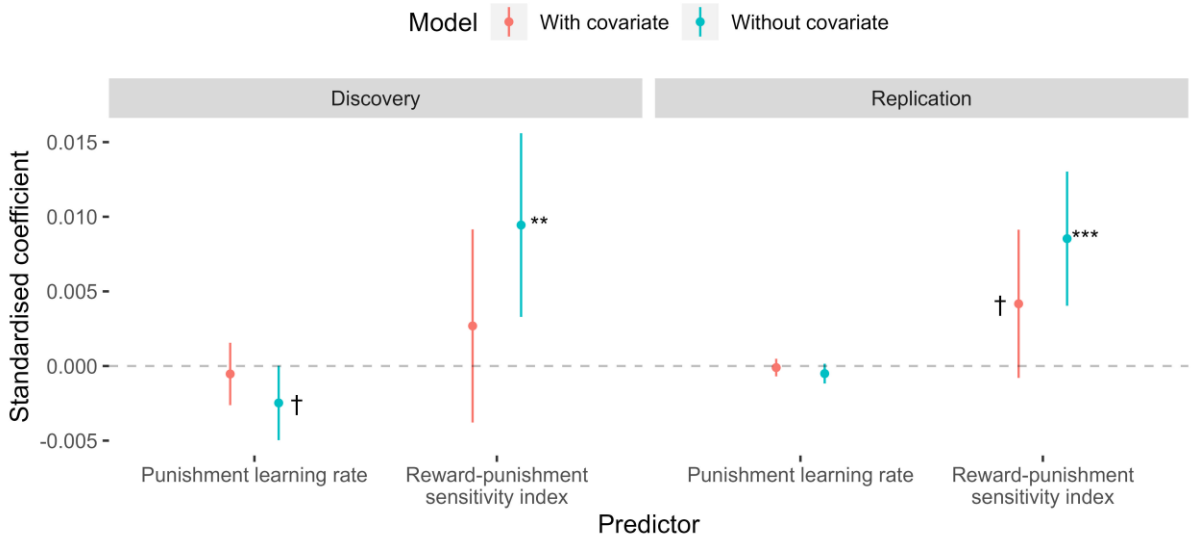
I assessed whether approach and avoidance responses, and their relationships with state anxiety, were impacted by punishment unpleasantness, by including unpleasantness ratings as a covariate into the hierarchical logistic regression model. Whilst unpleasantness was a significant predictor of choice (positively predicting safe option choices), all significant predictors and interaction effects from the model without unpleasantness ratings survived (Appendix 2.9—figure 2). Critically, this suggests that punishment unpleasantness does not account for all of the variance in the relationship between anxiety and avoidance.



Appendix 2.9—figure 2. The effect of including unpleasantness ratings as a covariate in the hierarchical logistic regression models of task choices. *Dots represent coefficient estimates from the model, with confidence intervals. Models are shown for both the discovery and replication samples. Significance levels are shown according to the following: p < 0.05 - *; p < 0.05 - **; p < 0.001 - ***.*

Mediation model

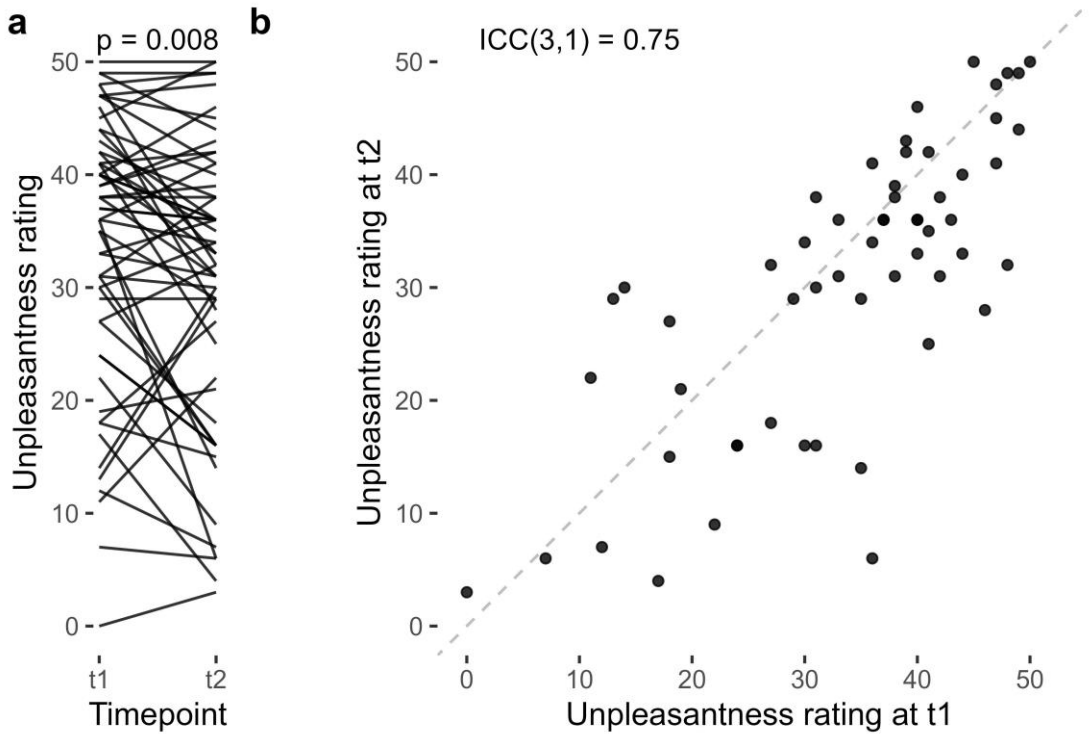
When unpleasantness ratings were included in the mediation models, the mediating effect of the reward-punishment sensitivity index did not survive (discovery sample: standardised β = 0.003 ± 0.003, *p* = 0.416; replication sample: standardised β = 0.004 ± 0.003, *p* = 0.100; Appendix 2.9—figure 3). Pooling the samples resulted in an effect that narrowly missed the significance threshold (standardised β = 0.004 ± 0.002, *p* = 0.068).



Appendix 2.9—figure 3. The effect of including unpleasantness ratings as a covariate in the mediation models. *Dots represent coefficient estimates from the model, with confidence intervals. Models are shown for both the discovery and replication samples. Significance levels are shown according to the following: p < 0.1 - †; p < 0.05 - *; p < 0.05 - **; p < 0.001 - ***.*

Test-retest reliability of unpleasantness

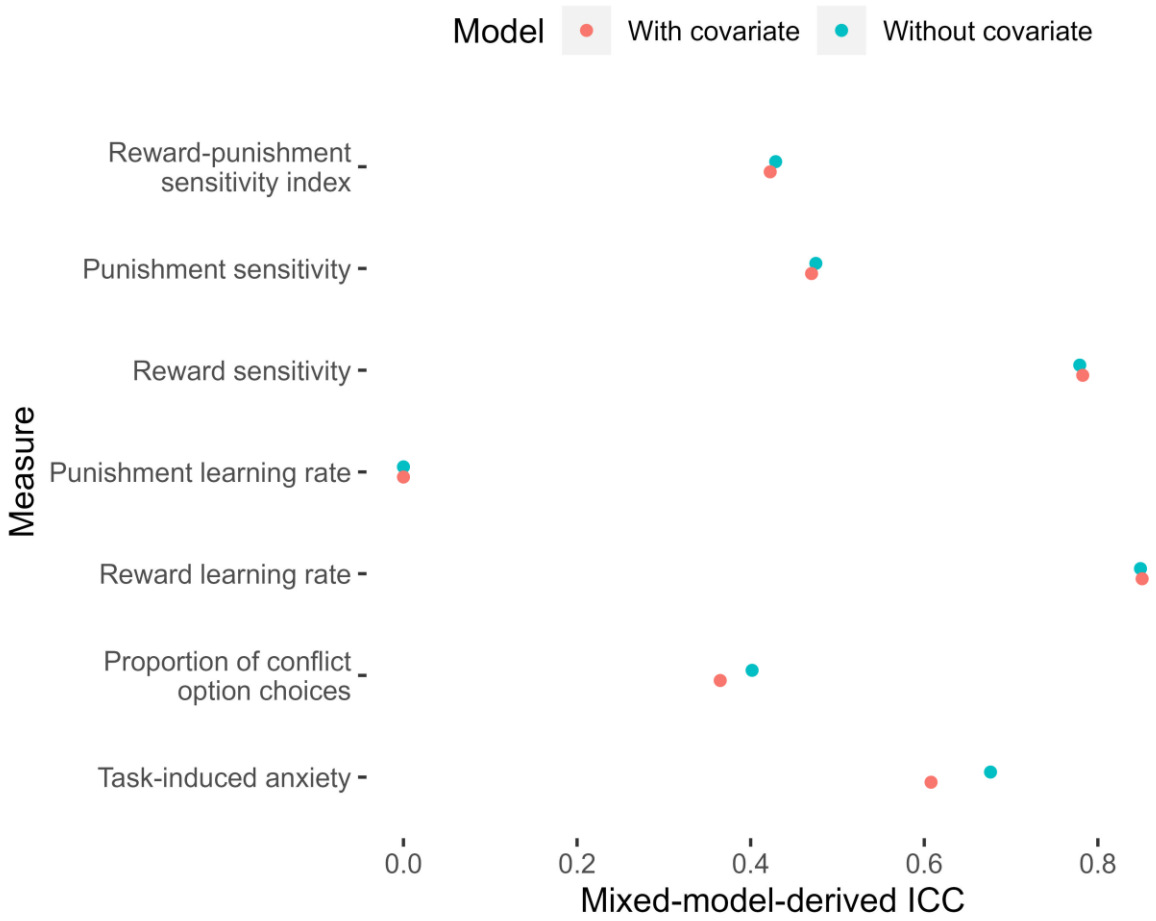
The test-retest reliability of unpleasantness ratings was excellent (ICC(3,1) = 0.75), although participants gave significantly lower ratings in the second session (t_{56} = 2.7, p = 0.008, d = 0.37; mean difference of 3.12, SD = 8.63).



***Appendix 2.9—figure 4. Test-retest reliability of unpleasantness ratings. a**, Comparing unpleasantness ratings across timepoints, participants rated the punishments as significantly less unpleasant in the second session. **b**, Correlation of ratings across timepoints.*

Reliability of other measures with/out unpleasantness

To assess the effect of accounting for unpleasantness ratings on reliability estimates of task performance, I extracted variance components from linear mixed models, following a standard approach (Nakagawa et al., 2017) – note that this was not the method used to estimate reliability values in the main analyses, but I used this specific approach to compare the reliability values with and without the covariate of unpleasantness ratings. The results indicated that unpleasantness ratings did not have a material effect on reliability (Appendix 2.9—figure 5).

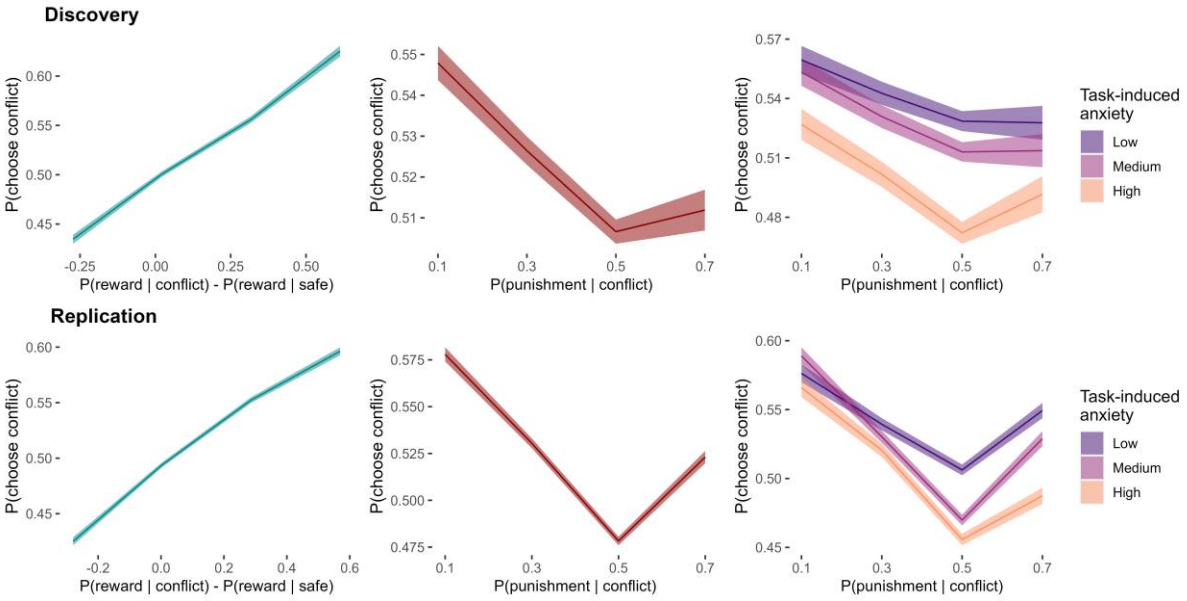


Appendix 2.9—figure 5. Mixed-model-derived intraclass correlation coefficients (ICCs) for measures of task performance, with and without accounting for unpleasantness. *Dots represent model-derived ICCs.*

2.8.10 Appendix 2.10: Sensitivity analysis: anxiety and inflexibility

As anxious participants were slower to update their estimates of punishment probability, I determined whether this was due to greater general inflexibility by examining the model including two sensitivity parameters, but one general learning rate (i.e. not split by outcome). The correlation between this general learning rate and task-induced anxiety was not significant in either sample (discovery: tau = -0.02, p = 0.504; replication: tau = -0.01, p = 0.625), suggesting that the effect is specific to punishment.

2.8.11 Appendix 2.11: Effects of outcome probabilities on choice in the task: non-linear effects



Appendix 2.11—figure 1. Effects of outcome probabilities on proportion of conflict option choices. *Mean probabilities of choosing the conflict arm across the sample are plotted with standard errors. The relationships between the drifting outcome probabilities in the task and group choice proportions showed non-linear trends in both the discovery and replication samples, especially for the effect of punishment probability on choice (both main effect and interaction effect with anxiety). Note. For visualisation purposes, the continuous predictors (based on the latent outcome probabilities or task-induced anxiety) were categorised into discrete bins.*

3 Computational effects of anxiolytic and anxiogenic manipulations on approach-avoidance reinforcement learning

3.1 Preface

In the previous chapter, I found that individuals who reported feeling more anxious during an approach-avoidance reinforcement learning task avoided punishments at the cost of potential rewards to a greater extent than less anxious individuals. Computational analyses indicated that this was potentially mediated by a heightened sensitivity to punishments, relative to rewards. However, owing to the cross-sectional nature of the study design, I could not draw strong conclusions about any causal relationship between anxiety and avoidance. Therefore, in the present chapter, I test this relationship directly through two parallel experiments; each experiment involved an experimental manipulation designed to either dampen or heighten anxiety during task performance.

Special acknowledgements.

The work from this chapter would not have been possible without a team of collaborators. I would like to thank Dr Harry Costello for providing clinical cover. I would also like to thank my students, Alan Liao, Udhay Cheema and Martina Tacconis for their assistance in data collection. Finally, I would like to thank Drs Harry Costello and Paulina Lukow for acting as the unblinded members of the research team for the lorazepam experiment.

3.2 Abstract

This chapter builds on the findings from Chapter 2, by examining the causal role of anxiety in avoidance behaviours using an approach-avoidance reinforcement learning task in which participants learned to choose between high-reward/high-punishment and low-reward/no-punishment options. To this end, I conducted two parallel within-subjects experiments: one using the fast-acting anxiolytic lorazepam to decrease anxiety, and the other using threat-of-shock to increase anxiety, during task performance. Previously, I found that individuals who experienced greater state anxiety during the task tended to avoid the high-reward/high-punishment options in favour of the low-reward/no-punishment options, and this effect was mediated by heightened sensitivity to punishments relative to the rewards. In this chapter, I predicted that lowering anxiety experimentally would reduce avoidance behaviour (where avoidance behaviour was operationalised as a preference for the low-reward/no-punishment option) and shift sensitivities towards rewards over punishments, whereas increasing anxiety would enhance avoidance behaviour and shift sensitivities towards punishments over rewards. The results showed that lorazepam did not significantly reduce anxiety experienced during the task, and it also did not appear to affect avoidance behaviour, as reflected in choices on the task. On the other hand, threat-of-shock elicited state anxiety, relative to safety, and it also induced a small increase in avoidance behaviour. The pre-registered computational analyses showed that lorazepam increased individual sensitivities to rewards, relative to punishments, whilst threat-of-shock had the opposite effect, increasing sensitivities to punishments, relative to rewards. Mediation analyses indicated that these shifts in the reward-punishment sensitivity index mediated the effects of anxiety manipulations on behaviour, supporting the role of this computational measure as a mechanism of anxiety-induced avoidance. These findings validate the task's computational validity as a model of anxiety-induced avoidance, and in the case of the lorazepam experiment, even in the absence of changes in self-reported anxiety. The research also offers novel insights into the mechanisms of how anxiety impacts avoidance behaviour.

3.3 Introduction

Anxiety-related disorders are among the leading causes of illness worldwide (Ferrari et al., 2022; Vos, 2020), and one of their core features is the tendency to avoid situations, activities or stimuli that provoke distress despite the potential for negative long-term consequences (American Psychiatric Association, 2013a; Barlow, 2002). While avoidance can provide temporary relief from anxiety, it is considered a key maintaining factor that undermines therapeutic efficacy and perpetuates anxiety by preventing individuals from learning that avoided stimuli do not need to be avoided (Salters-Pedneault et al., 2004). A better understanding of the underlying mechanisms of anxiety-related avoidance is needed to improve treatments for anxiety.

In the previous chapter, I developed a novel approach-avoidance reinforcement learning task to investigate learning and decision-making under conditions of competing rewards and punishments Participants learned to choose between an option associated with an affective punishment (an aversive sound) but also a relatively high (monetary) reward probability, versus another option with relatively lower reward probability but no punishment. I found that individuals who experienced greater state anxiety during the task tended to avoid the high reward/high-punishment options in favour of the low-reward/no-punishment options. Computational analyses suggested that this avoidance effect was mediated by a heightened sensitivity to punishments relative to rewards, thus identifying a potential computational mechanism underlying the relationship between anxiety and avoidance.

However, the cross-sectional design of the previous study meant I could not make strong claims about the causal role of anxiety on avoidance and the computational mechanism of the reward-punishment sensitivity index. To do this, I need to independently manipulate anxiety and demonstrate that this effects the identified mechanism. Therefore, in this chapter, I used established methods to experimentally modulate state anxiety as a more robust test of the causal role of anxiety in the task. Specifically, I utilised within-subjects manipulations to bidirectionally modulate state anxiety. In one experiment, I used lorazepam as an anxiolytic manipulation. Lorazepam is a benzodiazepine that acts as a reliable and fast-acting anxiolytic drug (Kaplan & DuPont, 2005) which is thought to exert its therapeutic effects via the $GABA_A$ (gamma-aminobutyric acid) receptor (Chebib & Johnston, 1999), and by inhibiting neuronal activity in key affective regions like the amygdala (Menard & Treit, 1999; Nutt & Malizia, 2001; Paulus et al., 2005). To test the opposite causal direction, in a separate experiment, I used threat-of-shock as an anxiogenic manipulation. Threat-of-shock is a well-validated and robust method for inducing anxiety in within-subjects experimental designs (Robinson et al., 2013; Schmitz & Grillon, 2012).

Through these two parallel within-subjects experiments using anxiolytic (lorazepam) and anxiogenic (threat-of-shock) manipulations, I probed three key research questions. First, does experimentally modulating state anxiety affect avoidance behaviour in the task? Finding that decreased anxiety leads to reduced avoidance, while increased anxiety has the opposite effect, would support the task's predictive validity with respect to state anxiety. Second, would modulating state anxiety correspondingly shift the hypothesised computational mechanism of relative reward/punishment sensitivity? Observing that decreased anxiety promotes greater relative weighting of rewards over punishments, while increased anxiety shows the opposite bias, would support the computational validity of this proposed mechanism. Finally, I tested whether the reward/punishment sensitivity index could mediate the effects of the anxiety manipulations on avoidance behaviour. This pattern of results would support the notion that the mechanistic computational account underlies the predictive relationship between anxiety and avoidance, providing an important demonstration of the task's integrated computational and predictive validity as a model of anxiety-related avoidance.

In summary, I aimed to rigorously test the predictive and computational validity of the approach-avoidance task as a model of anxiety-related avoidance by independently manipulating anxiety and testing its effect on task performance. This was accomplished through two separate within-subjects experiments: In one experiment, healthy participants completed the approach-avoidance reinforcement learning task after taking either lorazepam or placebo. In a separate experiment, another group of participants completed the task under threat-of-shock and safe-from-shock conditions. Using computational modelling, I examined whether anxiety causally impacts 1) avoidance behaviour, 2) the hypothesised mechanism of the reward-punishment sensitivity index, and 3) whether such computational changes mediate the behavioural effects. Following previous research showing that decreasing anxiety elicits reduced avoidance (Biedermann et al., 2017), whereas inducing anxiety increases avoidance (Clark et al., 2012; Pittig et al., 2015; Vogel & Schwabe, 2019), I predicted that the anxiolytic manipulation would *decrease* avoidance (as operationalised by increased likelihoods of choosing the high-reward/high-punishment option), and bias individual sensitivities towards *rewards* over punishments. Conversely, I predicted that the anxiogenic manipulation would *increase* avoidance (i.e. increase the likelihood of choosing the low-reward/no-punishment option), and bias individual sensitivities towards *punishments* over rewards. These hypotheses and predictions were pre-registered for both experiments (https://osf.io/7h9w5, https://osf.io/38krn).

3.4 Methods

3.4.1 Deviations from pre-registered plan

All methods and statistical analyses were consistent with the pre-registered plan, except for the estimation and analysis of the computational model parameters (see below for details). The pre-registered plan was to estimate the parameters using an expectation-maximisation algorithm, but I instead used Markov-Chain Monte Carlo sampling as it produces more accurate parameter estimates.

3.4.2 Participants

I previously found that, across participants, task-induced anxiety was associated with the computational measure of approach-avoidance conflict (the relative sensitivity to punishments versus rewards) at Pearson's $r = 0.15$, and that this computational measure had a test-retest reliability value of 0.41 (Chapter 2; Yamamori et al., 2023). To detect similar effects in within-subjects intervention designs, I converted this r value to a Cohen's d of 0.303. Therefore, I calibrated the *a priori* power analysis using G*Power (Faul et al., 2007), to detect a within-subjects effect of experimental condition on relative sensitivity to threat vs reward, via a paired t-test sensitive to an effect size $d = 0.303$, at 80% power and $\alpha = 0.05$ (one-tailed), with a correlation value of 0.41 of the dependent variable within-subjects. This resulted in a required sample size of 81 participants per experiment, leading to a total sample size of n = 162 across both studies.

I recruited participants aged 18-64 years from online advertising and institutional recruitment platforms. The exclusion criteria for both experiments were: current diagnosis of a psychiatric condition; currently on a course of psychiatric medication or therapy; history of neurological disorders/injuries; language-related disorders/literacy difficulties; and vision/hearing difficulties. Additional exclusion criteria for the lorazepam experiment were: a body mass index outside the range of 18.5 to 30; (potential) pregnancy or breastfeeding; allergies to the lorazepam or its excipients; alcohol or substance abuse/dependency; current use of any medications contraindicated for use with lorazepam, previous benzodiazepine use; and use of recreational drugs in the past month. Participants were reimbursed at a rate of £10 per hour in the lorazepam experiment and £8 per hour in the threat-of-shock experiment, and could earn a bonus of up to the respective hourly rates based on their performance in the task. Both experiments received ethical approval from the University College London Research Ethics Committee (refs. lorazepam experiment: 14557/001; threat-of-shock experiment: 6198/002). Encouragingly, despite the pre-registered exclusion criteria in place to ensure data quality (https://osf.io/38krn), no participants met these criteria.

3.4.3 The approach-avoidance reinforcement learning task

I used the same approach-avoidance reinforcement learning task as described in Chapter 2, in which participants learned to simultaneously pursue rewards and avoid punishments (Yamamori et al., 2023). On each trial, participants had 2 seconds to choose one of two options, depicted as distinct visual stimuli (Figure 3.1A), using the left/right arrow keys on their keyboard. The chosen option was then highlighted by a white border, and participants were presented with the outcome for that trial for 2 seconds. This could be one of 1) no outcome, 2) a reward (depicted with an image of a coin), 3) a punishment (an aversive sound), or 4) both the reward and punishment (Figure 3.1B). I incentivised participants to collect the coins by informing them that they could earn a bonus payment based on the number of coins earned during the task (1p for every two coins). If no choice was made, the text 'Too slow!' was shown for 2 seconds and the trial was skipped. Each option was associated with separate probabilities of producing the reward/punishment outcomes, and these probabilities fluctuated randomly over trials. Both options were associated with the reward, but only one was associated with the punishment (i.e., the 'conflict' option could produce both rewards and punishments, but the 'non-conflict' option could produce only rewards and never punishments). The participants could take a break every 100 trials during the task. The experimental task was programmed in MATLAB R2015b (The MathWorks Inc., 2015) using the Psychophysics Toolbox extension ('Psychtoolbox-3'; Brainard, 1997).

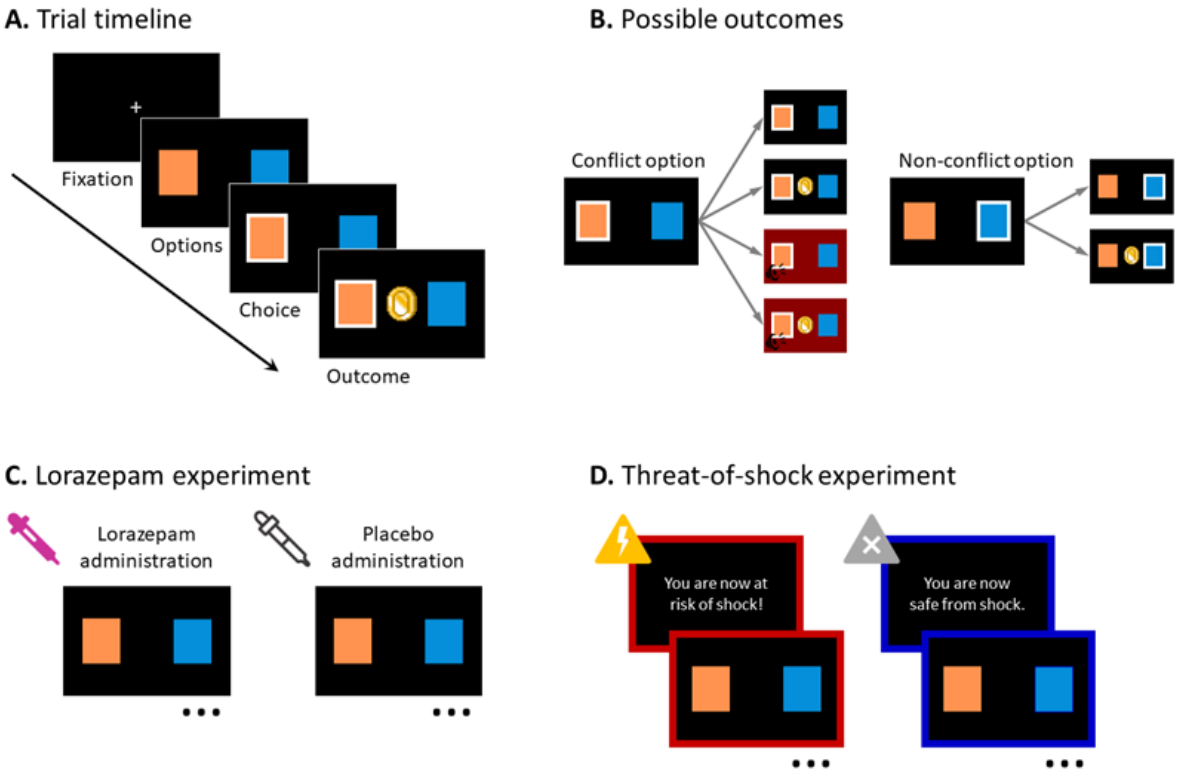


Figure 3.1, The approach-avoidance reinforcement learning task. A, *Trial timeline. A fixation cross initiates the trial. Participants are presented with two options from which they choose one, which then produces a choice-dependent outcome.* ***B,*** *Possible outcomes. There were four possible outcomes: (1) no reward and no aversive sound; (2) a reward and no aversive sound; (3) an aversive sound and no reward; or (4) both the reward and the aversive sound. Choosing the conflict option could result in any of these four outcomes, whilst the non-conflict could never lead to an aversive sound.* ***C,*** *The lorazepam experiment. In this within-subjects experiment, all participants completed the task under two conditions: after lorazepam administration and after placebo administration (note, the participants and experimenter were blind to these conditions). These were completed over separate days.* ***D,*** *The threat-of-shock experiment. In this within-subjects experiment, all participants completed the task under two conditions: during threat-of-shock where they could receive unpredictable electric shocks, and during safe blocks where there were no shocks.*

I used aversive auditory stimuli as the punishments in the task, consisting of combinations of female screams and high-frequency tones which were individually rated as aversive in a previous online study (Seow & Hauser, 2022) – these stimuli were identical to those used in Chapter 2. To maximise their aversiveness, participants listened to the sounds using headphones during the task, and I asked participants to increase the volume setting to the loudest possible whilst still being comfortable. This was achieved in a calibration task prior to the main task, in which participants were presented with short bursts of white noise, matched in loudness to the auditory punishment stimuli in the main task, to which they adjusted the volume setting. I used the white noise as a placeholder for the punishments to limit desensitisation.

3.4.4 Self-report questionnaires

Psychiatric symptoms were measured using the Generalised Anxiety Disorder scale (GAD-7; Spitzer et al., 2006), Patient Health Questionnaire depression scale (PHQ-8; Kroenke et al., 2009), the Brief

Experiential Avoidance Questionnaire (BEAQ; Gamez et al., 2014), and the Behavioural Inhibition/Activation Scales (Carver & White, 1994).

After completing the task on each experimental session, participants were also asked to rate the unpleasantness of the auditory punishments on a continuous slider ranging from 'Not at all' (0) to 'Extremely' (50).

In the lorazepam experiment, I also collected ratings of subjective drug effects for feelings of anxiety, drowsiness, nausea, and ability to concentrate. These were measured on visual analogue scales ranging from 0 ('Not at all') to 10 ('Extremely') at two timepoints: 1) baseline, just before drug administration, and 2) 90 minutes after drug administration (i.e. just before participants completed the task).

3.4.5 Procedures

The two experiments were conducted separately and consist of separate groups of participants. While these experiments could be presented as separate chapters, I report them together as a combined chapter given their complementary nature; both manipulate anxiety levels to observe their putatively opposing effects on approach-avoidance behaviour. Specifically, studying the effects of the anxiolytic and anxiogenic manipulations together provides clearer insights into which computational mechanisms align across both interventions.

Experiment I: lorazepam administration (anxiolytic manipulation)

This experiment consisted of a randomised, placebo-controlled, double-blind, within-subjects design. Participants were first administered either 1 mg of lorazepam as an oral solution suspended in medium-chain triglyceride oil, or matched placebo which was simply the medium-chain triglyceride oil without lorazepam. Both the active drug and placebo were mixed into 100ml of orange juice to mask differences in flavour.

Participants attended two testing sessions on separate days, with an interval of at least week in between sessions. Each session involved the administration of either the active drug or placebo, and after a 90-minute drug absorption waiting period, participants completed 200 trials of the approach-avoidance reinforcement learning task (Figure 3.1C). After the task, participants rated how anxious they felt during the task, on a visual analogue scale from 'Not at all' to 'Extremely' (scored from 0 and 50).

The dose and drug absorption duration were chosen based on a previous study which found anxiolytic effects of lorazepam (identical dose) on an approach-avoidance test, 60 min after drug administration (Biedermann et al., 2017).

Experiment II: threat-of-shock (anxiogenic manipulation)

In this experiment, I induced a heightened state of anxiety in participants using threat of unpredictable shocks. The shocks were delivered through electrodes attached to the participants' left wrist, using a Digitimer DS5 Constant Current Stimulator (Digitimer Ltd., Welwyn Garden City, United Kingdom). Each shock consisted of a 200 ms, 40 Hz train of stimulation. Shock intensities ranged from 0 to 10 V and were calibrated to each individual according to a work-up procedure. Here, participants received shocks of increasing intensity and rated them on a scale from 0 (not felt) to 10 (unbearable), until they reported a subjective rating of 8.

Participants attended a single testing session in which they completed alternating 'threat-of-shock' and 'safe-from-shock' blocks of the approach-avoidance reinforcement learning task (Figure 3.1D), with each block lasting for 100 trials. The identity of each block was instructed to participants at the beginning of each block, and signalled during blocks by a coloured border of the screen (red for threat-of-shock, blue for safe-from-shock). Participants were told that they could receive up to three shocks during each threat block, but all participants were given two shocks in the first threat-of-shock block, and one shock in the second threat-of-shock block. The shocks were delivered during the inter-trial intervals to avoid interactions with task events. After each block, participants rated how anxious they felt during the block, on a scale visual analogue scale from 'Not at all' to 'Extremely' (scored from 0 and 50).

3.4.6 Analysis

Model-agnostic analyses

Anxiety

In line with my pre-registered analyses, I tested the effects of experimental condition (lorazepam vs placebo; threat-of-shock vs safe-from-shock) on self-reported anxiety ratings via paired t-tests across conditions for each experiment. To account for temporal effects, namely session for the lorazepam experiment and block number for the threat-of-shock experiment, I also specified a hierarchical linear regression model in each experiment. For the lorazepam experiment, this was tested as:

$$\text{Anxiety} \sim \text{Condition} + \text{Session} + \text{Condition} \cdot \text{Session} + (1 \mid \text{participant})$$

For the threat-of-shock experiment, this was tested as:

$$\text{Anxiety} \sim \text{Condition} + \text{Block} + \text{Condition} \cdot \text{Block} + (1 \mid \text{participant})$$

Task performance

To assess for effects of each experimental condition on avoidance behaviour, I first used the pre-registered paired t-tests to test whether the proportion of conflict option choices differed across conditions for each experiment.

As a more granular assessment of task performance, I also specified hierarchical logistic regression models of trial-level choices to test the effect of the dynamic outcome probabilities and how these interacted with the experimental conditions. These models included predictors aligned with approach and avoidance responses, based on the latent probabilities of observing an outcome given the chosen option at trial t, $\mathrm{P}(\mathrm{outcome} \mid \mathrm{option})_t$. Firstly, if individuals were learning to choose to maximise their rewards earned, they should generally be choosing the option which was relatively more likely to produce the reward on each trial. This would result in a significant effect of the difference in reward probabilities across options, as given by:

$$\delta\mathrm{P}(\mathrm{reward})_t = \mathrm{P}(\mathrm{reward} \mid \mathrm{conflict})_t - \mathrm{P}(\mathrm{reward} \mid \mathrm{safe})_t$$

Secondly, if individuals were learning to avoid punishments, they should generally be avoiding the conflict option on trials when it is likely to produce the punishment. This would result in a significant effect of the difference in punishment probabilities across options, as given by:

$$\delta\mathrm{P}(\mathrm{punishment})_t = \mathrm{P}(\mathrm{punishment} \mid \mathrm{conflict})_t$$

where $\mathrm{P}(\mathrm{punishment} \mid \mathrm{safe})$ is omitted as it was always 0 across trials.

The models included these predictors based on the outcome probabilities, and their interactions with the experimental condition in each experiment. The general model specification was therefore:

$$\mathrm{Choice} \sim \delta\mathrm{P}(\mathrm{reward}) + \delta\mathrm{P}(\mathrm{punishment}) + [\delta\mathrm{P}(\mathrm{reward}) + \delta\mathrm{P}(\mathrm{punishment})] \cdot \mathrm{Condition} + (1 \mid \mathrm{participant})$$

Computational model specification

For the computational analyses, I used a similar strategy to that in Chapter 2, by fitting reinforcement learning models to the data. I first assessed whether the winning model from Chapter 2, which included reward- and punishment-specific learning rates and outcome sensitivities, was also the most parsimonious model in the present data when it was collapsed across conditions in each experiment. I compared this model against models in which the learning rates and/or sensitivity parameters were fixed across reward/punishment outcomes (Table 3.1).

Table 3.1, Condition-agnostic model specification

Model	**Parameters**		
Symmetrical learning	α		β
Asymmetric learning rates	α_r	α_p	β

Asymmetric sensitivities		α	β_r	β_p
Asymmetric learning rates and sensitivities	α_r	α_p	β_r	β_p

Briefly, all models implemented a Rescorla-Wagner mechanism of learning for updating the expected values of performing certain actions, $Q(a)$ (i.e. choosing certain options). Each model learned the expected value for obtaining the reward, Q_r, and the punishment, Q_p, separately, according to:

$$Q_{r,t+1}(a) = Q_{r,t}(a) + \alpha \cdot \left[r_t - Q_{r,t}(a)\right]$$

$$Q_{p,t+1}(a) = Q_{p,t}(a) + \alpha \cdot \left[p_t - Q_{p,t}(a)\right]$$

where t is the trial number; r and p are the reward and punishment outcomes observed on a given trial, respectively; and α is a learning rate parameter that could take on different values across reward/punishment outcomes, depending on the model.

These action values (Q_r and Q_p) are then integrated into a single decision weight:

$$W = \beta \cdot (Q_r - Q_p)$$

where β is an outcome sensitivity parameter that scales the degree to which these action values impact choices. Again, some models implemented separate values of the sensitivity parameter across reward/punishment outcomes. via:

$$W = \beta_r \cdot Q_r - \beta_p \cdot Q_p$$

As in Chapter 2, the reward-punishment sensitivity index was calculated as the ratio between the reward and punishment sensitivity parameters, β_r/β_p.

Choice was implemented via the softmax function based on these action weights:

$$P(a) = \frac{W(a)}{\sum_i W(i)}$$

I then used the most parsimonious model from the collapsed data in each experiment to test for any effects of the anxiety manipulations on the model parameters. To achieve this, I fitted models which estimated both treatment- and control-specific values of a certain parameter (e.g. a reward learning rate under lorazepam and a reward learning rate under placebo). The first model fitted the model parameters separately, more specifically as drawn from independent distributions. The second model specified correlations between treatment and control parameters (i.e. allowed the reward learning under lorazepam to correlate with that under placebo) using Cholesky factors implemented in Stan. This model specification allows the model to pool data across conditions to estimate the parameter

values, which has recently been shown to reduce estimation noise in repeated-measures modelling (Waltmann et al., 2022).

Model fitting and comparison

The models were fitted using an MCMC algorithm in Stan (Stan Development Team, 2023). Each model was sampled for 1000 iterations over 15 chains (for a total of 15000 draws per model), with 500 warm-up samples per chain. Model fit was assessed with standard diagnostic measures: visual inspection of parameter samples; no divergences; Gelman-Rubin R-hat statistics below 1.1; and energy-Bayes fractions of missing information above 0.2. I compared across models using the leave-one-out information criterion (LOOIC; Vehtari et al., 2017) where the lowest LOOIC value indicated the winning model.

Parameter analyses

To study how the model parameters are affected by the experimental conditions, I extracted the mean parameter estimates from the individual-level parameter distributions of the winning model. I then used paired t-tests to assess the effects of the experimental conditions on the parameter values. In line with the pre-registered analysis plan, I used one-tailed tests for the reward-punishment sensitivity index, testing the predictions that lorazepam, compared to placebo, would increase individual sensitivity to *rewards* relative to punishments and that threat-of-shock, compared to safety, would increase individual sensitivity to *punishments* relative to rewards. The other comparisons of parameters between conditions were tested using two-tailed tests.

Given that the MCMC algorithm provided distributions over parameter values, I employed a Bayesian approach to complement the frequentist analyses. For each parameter, I calculated the posterior distributions of the differences in its values across conditions, first as t-statistics, taking into account the covariance of parameter values across conditions, and then converted these to effect size estimates (d). From these distributions, I extracted the *maximum a posteriori* (MAP) estimate, which represents the most probable effect size, and the 95% highest density interval (HDI), which provides a range of credible values for the effect size. To assess practical significance, I used a region of practical equivalence (ROPE) analysis, with a threshold of ± 0.2, corresponding to the commonly used cut-off for a small effect size. The percentage of the posterior distribution falling within this ROPE indicates the probability that the effect is practically equivalent to zero. Additionally, I calculated the probability of direction (pd), which represents the probability that the effect is positive or negative. In interpreting these results, one can consider an effect to be meaningful (akin to 'significant' in frequentist analysis) if its 95% HDI excludes zero and if less than 5% of the posterior distribution falls within the ROPE. The probability of direction provides further evidence for the consistency of the

effect's sign. This approach allows for an interpretation of the parameter effects which takes into account their uncertainty.

Exploratory mediation analyses

In a similar vein to the mediation analyses from the previous chapter (Chapter 2), I also conducted mediation analyses here to test whether the model parameters mediated the effect of the interventions on behaviour. More precisely, I tested for mediation effects using a bootstrapped product-of-coefficients approach (Baron & Kenny, 1986; Preacher & Hayes, 2008). I refer to these analyses as exploratory as they were not pre-registered.

For each experiment and parameter, I ran a mediation analysis involving 4000 bootstrapped samples at the participant level. Sampling at the participant level means that for each bootstrap sample, the entire dataset of individual choices for each participant was resampled, rather than drawing individual choices randomly. This preserves the within-subject correlation structure of the data. To compute the coefficients necessary for the product-of-coefficients method, I specified two hierarchical regressions per test: first, a hierarchical linear regression model to assess the relationship between the experimental condition and the mediator parameter, with random intercepts by participant; and second, a hierarchical logistic regression model to examine trial-level choices, with choices as the dependent variable and both the experimental condition and model parameter as predictors, also with random intercepts by participant. The model specifications are presented in Appendix 3.1.

To estimate the mediation effect, the product of the fixed effect of the treatment condition on the mediator parameter (from the linear model) and the fixed effect of the mediator parameter on the outcome variable (from the logistic model) was calculated. Bias-corrected accelerated confidence intervals were computed from the bootstrapped samples for inference. A mediation effect was deemed significant if the confidence interval did not span a value of zero. This approach allowed for robust testing of the indirect effects of the experimental conditions through various mediator parameters, providing insights into the mechanisms influencing behavioural outcomes in both the lorazepam and threat-of-shock intervention studies.

3.5 Results

3.5.1 Participants

Demographic data are presented in Table 3.2. Across samples, there was no significant difference in age or proportion of female/male participants (Table 3.2). Overall, the combined samples self-reported low scores on the measure of clinical anxiety (via Generalised Anxiety Disorder scale, mean score = 4.1, SD = 3.8), and there was no evidence for a difference across samples (Table 3.2).

Table 3.2, Demographic data across experiments.

	Experiment		Comparison
	Lorazepam	**Threat-of-shock**	
N	81	81	
Age	Mean = 24 years (SD = 8)	Mean = 27 years (SD = 10)	$t_{153.7} = 1.93$, $p = 0.055$
Sex	$N_{female} = 49$; $N_{male} = 32$	$N_{female} = 54$; $N_{male} = 27$	$\chi_1 = 0.43$, $p = 0.514$
Clinical anxiety symptoms (via Generalised Anxiety Disorder, 7-item scale)	Mean = 4.1 (SD = 3.3)	Mean = 4.1 (SD = 4.2)	$t_{151.0} = 0.002$, $p = 0.998$

3.5.2 Basic task effects

I first verified that participants approached rewards and avoided punishments irrespective of the experimental manipulations. A hierarchical logistic regression on trial-level choice showed that, consistent with the findings reported in Chapter 2 (Yamamori et al., 2023), participants chose (i.e. approached) the option which was more likely to produce the reward on each trial (lorazepam experiment: β = 0.33 [0.30, 0.36], $p < 0.001$; threat-of-shock experiment: β = 0.27 [0.24, 0.30], $p < 0.001$; Figure 3.2A), and they also avoided the conflict option when it was likely to result in punishment (lorazepam experiment: β = -0.09 [-0.12, -0.05], $p < 0.001$; threat-of-shock experiment: β = -0.09 [-0.12, -0.06], $p < 0.001$; Figure 3.2A).

3.5.3 The effects of the interventions on anxiety

I then assessed whether the anxiety manipulations had the expected effects on self-reported anxiety. The pre-registered t-tests on anxiety ratings during the task showed no evidence for an effect of lorazepam on anxiety ($t_{80} = -0.41$, $p_{one\text{-}tailed} = 0.342$) whilst threat-of-shock elicited a large increase in anxiety ($t_{80} = 13.39$, $p_{one\text{-}tailed} < 0.001$, $d = 1.50$). Further tests via linear mixed models showed that, in the lorazepam experiment, participants were not significantly less anxious after taking lorazepam compared to placebo (β = -4.6 [-18.77, 9.52], $p = 0.520$; Figure 3.2B), but they were significantly less anxious in their second session compared to their first, irrespective of which drug they took (β = -5.7 [-10.97, -0.46], $p = 0.033$). The interaction between drug condition and order was not significant (β = 2.5 [-6.75, 11.83], $p = 0.590$). In the threat-of-shock experiment, participants were significantly more anxious under threat compared to safety (β = 13.7 [9.31, 17.98], $p < 0.001$; Figure 3.2B), and the effect of block number (i.e. time) on anxiety was not significant (β = -1.1 [-2.15, 0.03], $p = 0.056$).

The interaction between threat condition and block number was not significant (β = 0.5 [-1.10, 2.11], p = 0.539).

3.5.4 The effects of the interventions on avoidance

To test my first set of hypotheses, that decreasing anxiety would dampen avoidance behaviour whereas increasing anxiety would elicit avoidance, I conducted pre-registered paired t-tests on the proportions of conflict option choices across conditions in each experiment. Although the directions of effects were consistent with my predictions, neither effect achieved significance (effect of lorazepam: mean difference = 0.2% [-2.8%, 3.1%], t_{80} = 0.12, $p_{\text{one-tailed}}$ = 0.451; effect of threat-of-shock: mean difference = -1.6% [-4.5%, 1.3%], t_{80} = -1.12, $p_{\text{one-tailed}}$ = 0.133; Figure 3.2C).

In the hierarchical regression models, there was no significant effect of drug (β = -0.002 [-0.05, 0.04], p = 0.927; Figure 3.2A) or any significant interactions with drug (reward: β = -0.01 [-0.06, 0.04], p = 0.647; punishment: β = -0.03 [-0.07, 0.02], p = 0.248); Figure 3.2A) in experiment I. However, in experiment II, threat-of-shock was a significant predictor of choice, albeit with a small effect size (β = -0.02 [-0.03, 0.00], p = 0.004; Figure 3.2A). This significant result was likely due to increased power from utilising the full trial-level data. Threat condition did not interact with the outcome probability predictors (reward: β = -0.02 [-0.07, 0.02], p = 0.331; punishment: β = -0.01 [-0.06, 0.04], p = 0.649; Figure 3.2A).

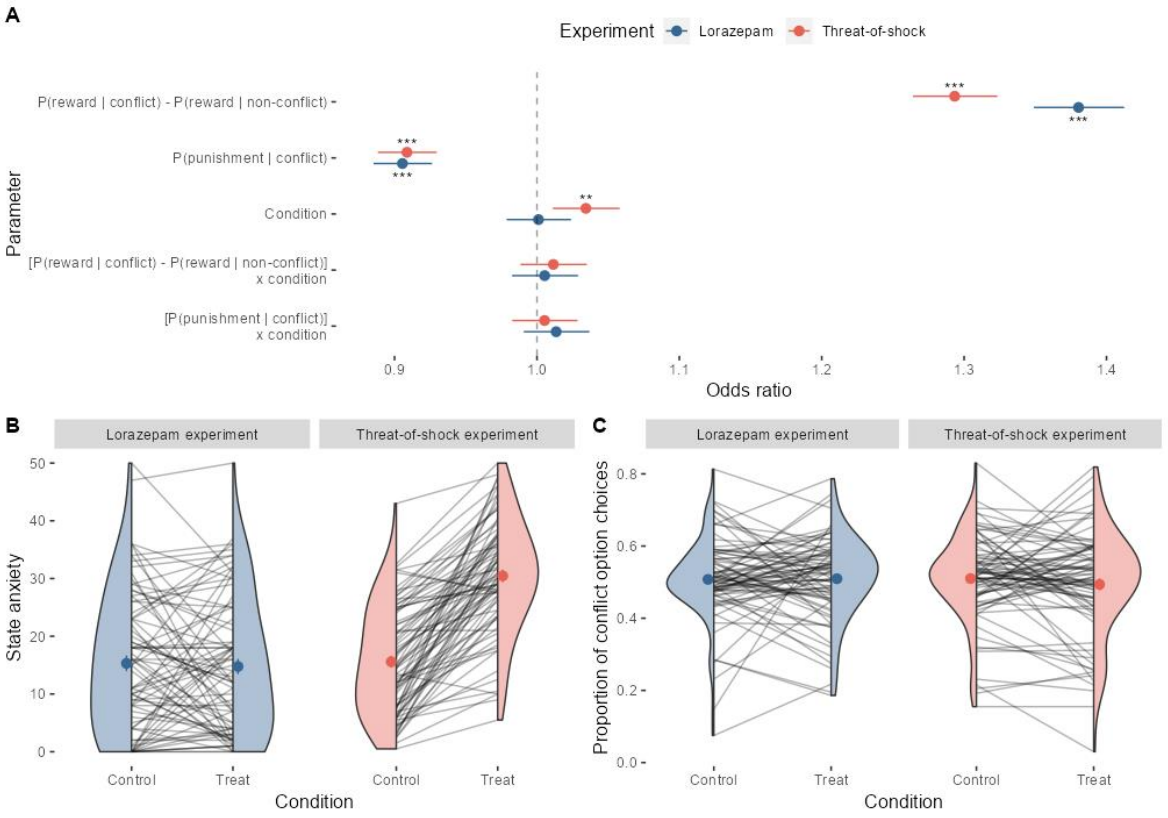


Figure 3.2, Effects of lorazepam and threat-of-shock on task performance and state anxiety. A, *Coefficients from the hierarchical logistic regression models of trial-by-trial choices in the task for each experiment. Significance levels are shown according to the following: p < 0.05 – *; p < 0.01 – **; p < 0.001 – ***. Error bars represent confidence intervals.* ***B,*** *Violin plots showing the distribution of state anxiety ratings across control and treatment conditions for both experiments. Individual participant data are connected by lines. The threat-of-shock condition significantly increased state anxiety compared to the control, while lorazepam did not significantly decrease state anxiety during the task.* ***C,*** *Violin plots displaying the proportion of conflict option choices (high-reward/high-punishment) across control and treatment conditions for both experiments. Individual participant data are connected by lines. Paired t-tests did not show significant differences in the proportion of conflict option choices across conditions for either experiment.*

3.5.5 The effects of the interventions on the computational model parameters

To test my third set of hypotheses, that the anxiety manipulations would modulate how sensitive individuals were to rewards relative to punishment, I analysed the reinforcement learning model fitted to the trial-level choices. I first verified that the winning model from Chapter 2 was also the most parsimonious model in both experiments when the data were collapsed across conditions – this was confirmed (Figure 3.3A). I then used this model to estimate condition-specific parameters for each experiment. Model comparison showed that the model specification in which the condition-specific parameters were drawn from a common, multivariate prior was more parsimonious compared to the model specification with independent priors (difference in LOOIC for lorazepam experiment: 30.6; threat-of-shock experiment: 38) – the latter also lead to poorer parameter reliability across conditions (range of Pearson's *r* values across conditions for independent priors: 0.30 – 0.79; for multivariate priors: 0.63 – 0.93; Appendix 3.2). Therefore, I tested for effects of experimental condition on parameter values from the common prior models.

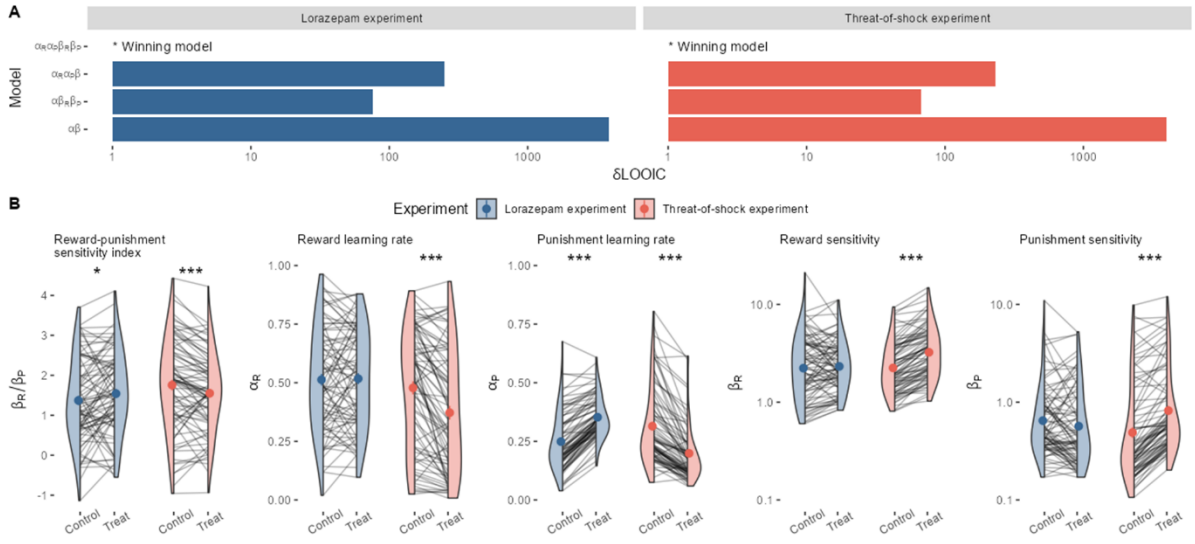


Figure 3.3, Computational modelling of task behaviour. A, *Model comparison results in the data collapsed across conditions. The difference in the leave-one-out information criteria (δLOOIC) from each model relative to the winning model is indicated on the x-axis. The winning model included specific learning rates for reward (α_R) and punishment learning (α_P), and specific outcome sensitivity parameters for reward (β_R) and punishment (β_P).* ***B,*** *Distributions of individual parameter values from the winning model for each experiment.*

My pre-registered predictions that lorazepam administration would decrease participants' relative sensitivities to punishment vs reward, and that threat-of-shock would have the opposite effect, were supported by the data, as shown by one-tailed paired t-tests on the reward-punishment sensitivity indices (lorazepam experiment: t_{80} = 1.96, $p_{\text{one-tailed}}$ = 0.027, d = 0.22, mean under lorazepam = 1.54 (SD = 1.05), mean under placebo = 1.37 (SD = 1.06); threat-of-shock experiment: t_{80} = -3.58, $p_{\text{one-tailed}}$ < 0.001, d = 0.40, mean under threat = 1.55 (SD = 1.00), mean under safe = 1.76 (SD = 1.17); Figure 3.3B). Lorazepam and threat-of-shock also had large and diverging effects on the punishment learning rate, where lorazepam increased this parameter (t_{80} = 17.84, p < 0.001, d = 1.99, mean under lorazepam = 0.35 (SD = 0.09), mean under placebo = 0.25 (SD = 0.12)), whereas threat-of-shock had the opposite effect (t_{80} = -11.08, p < 0.001, d = 1.24, mean under threat = 0.20 (SD = 0.12), mean under safe = 0.32 (SD = 0.16); Figure 3.3B). Lorazepam had no other significant effects on the parameter values (reward learning rate: t_{80} = 0.29, p = 0.774; reward sensitivity: t_{80} = -0.63, p = 0.530; punishment sensitivity: t_{80} = -0.96, p = 0.340). However, there was evidence for additional effects of threat-of-shock on reducing reward learning rate (t_{80} = -6.41, p < 0.001, d = 0.72, mean under threat = 0.37 (SD = 0.26), mean under safe = 0.48 (SD = 0.25); Figure 3.3B), increasing reward sensitivity (t_{80} = 8.70, p < 0.001, d = 0.97, mean under threat = 3.94 (SD = 2.72), mean under safe = 2.71 (SD = 1.87); Figure 3.3B) and increasing punishment sensitivity (t_{80} = 4.19, p < 0.001, d = 0.47, mean under threat = 1.47 (SD = 2.07), mean under safe = 1.07 (SD = 1.78); Figure 3.3B).

Bayesian analyses of the posterior distributions largely corroborated the above frequentist results for the relative sensitivity indices, providing further support for the pre-registered predictions, albeit with more uncertainty (Figure 3.4). The analyses showed a 78.4% probability of lorazepam biasing sensitivities towards rewards over punishments (MAP estimate = 0.17, 95% HDI [-0.24, 0.56], 5.4% in ROPE; Figure 3.4). Similarly, there was a 78.8% probability that threat-of-shock had the opposite effect, i.e. biased sensitivities towards punishments over rewards (MAP = -0.23, 95% HDI [-0.88, 0.35], 3.4% in ROPE; Figure 3.4). While similar to the frequentist results, these Bayesian metrics, particularly the wide confidence intervals and non-negligible proportions failing within the ROPE for lorazepam, suggest some uncertainty around these effects.

With respect to the other model parameters, the Bayesian analyses were also consistent with the frequentist results. There was strong evidence for diverging effects on punishment learning rates, with lorazepam showing a 97.7% probability of increasing them (MAP = 0.83, 95% HDI [-0.09, 1.95], 0.5% in ROPE; Figure 3.4), while threat-of-shock had a 94.7% probability of decreasing them (MAP = -0.58, 95% HDI [-1.69, 0.22], 1.2% in ROPE; Figure 3.4). The effects of lorazepam on the remaining model parameters (reward learning rate, reward sensitivity and punishment sensitivity) were weak or inconclusive, with substantial probabilities of the drug effect in both directions, and considerable proportions of the posterior distributions in the ROPE. In contrast, there was strong evidence for the effects of threat-of-shock on these parameters, again aligning with the frequentist results. There was a 99.9% probability of threat-of-shock reducing reward learning rates (MAP = -0.55, 95% HDI [-0.93, -0.19], 0.03% in ROPE; Figure 3.4), and significant positive effects on both reward and punishment sensitivities (reward: >99.9% probability, MAP = 1.01, 95% HDI [0.55, 1.64], 0% in ROPE; punishment: 98.7% probability, MAP = 0.81, 95% HDI [0.12, 1.6], 0.3% in ROPE; Figure 3.4).

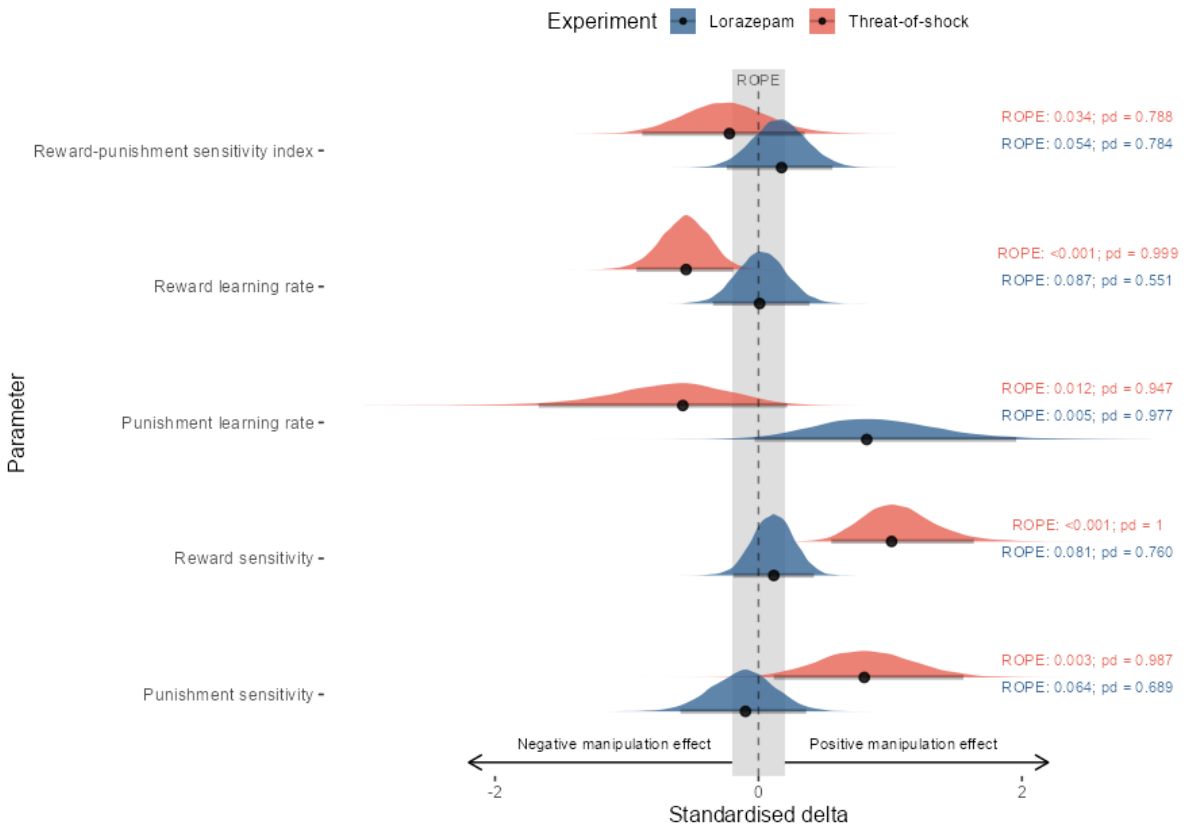


Figure 3.4, Bayesian analyses of the effect of the anxiety manipulations on the group-level computational parameters. *This forest plot shows the standardised effect of each manipulation (lorazepam and threat-of-shock) on each model parameter. Positive values indicate an increase in the parameter due to the manipulation, while negative values indicate a decrease. The densities represent the posterior distributions for each effect, with black points representing the maximum a posteriori (MAP) estimate of the effect and horizontal lines indicating the 95% Highest Density Interval (HDI). The grey shaded area represents the Region of Practical Equivalence (ROPE), centred around zero. Effects within this region are considered practically equivalent to no effect. The ROPE and probability of direction (pd) values are provided for each parameter and condition.*

The mediation analyses revealed that the reward-punishment sensitivity index was a significant mediator of the effect of intervention on choices in both experiments (lorazepam experiment: indirect effect = 0.07 [bootstrapped 95% bias-corrected accelerated confidence interval: 0.02, 0.13]; threat-of-shock experiment: indirect effect = -0.10 [-0.14, -0.05]; Figure 3.5). Punishment sensitivity significantly mediated the effects of the interventions on behaviour in both experiments, and had the largest effect sizes of the indirect effects across each experiment (lorazepam experiment: indirect effect = 0.23 [0.14, 0.31]; threat-of-shock experiment: indirect effect = -0.39 [-0.52, -0.23]; Figure 3.5). Reward sensitivity was a significant mediator only in the threat-of-shock experiment (indirect effect = 0.07 [0.02, 0.13]; Figure 3.5). These findings indicate that changes in sensitivity to rewards and punishments, especially the latter, are key mechanisms through which anxiety manipulations modulate avoidance behaviour.

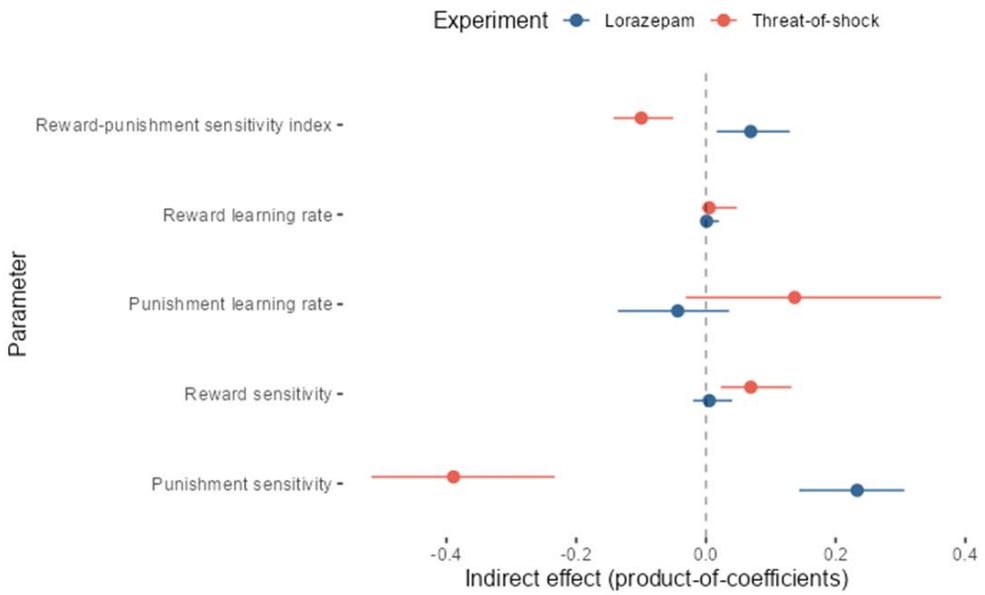


Figure 3.5, Do the computational model parameters mediate the effects of the anxiety interventions on behaviour? *Forest plot of the coefficients from the mediation analyses. The coefficients represent products of the fixed effect of the experimental conditions (treatment vs control) on the model parameter (via a linear model) and the fixed effect of the model parameter on the trial-by-trial choices (via a logistic model). Error bars represent bias-corrected accelerated confidence intervals estimated from 4000 bootstrap samples. Mediation effects were deemed as 'significant' if the confidence intervals did not contain zero.*

3.5.6 Effects of lorazepam on other subjective measures

The main effect of drug condition on subjective anxiety ratings measured prior to task completion was not significant (β = -0.1 [95% CI: -0.21, 0.02], $p = 0.099$); and the interaction between drug condition and timepoint (baseline vs 90 minutes after drug administration) was also not significant (β = 0.01 [-0.11, 0.12], $p = 0.888$). However, there was a significant interaction between drug condition and timepoint on drowsiness ratings (β = -0.2 [-0.26, 0.06], $p = 0.003$; Appendix 3.3). Post-hoc analyses examining the effect of timepoint under each drug separately showed that lorazepam administration significantly increased drowsiness ratings (β = 0.4 [0.27, 0.50], $p < 0.001$), while placebo administration did not (β = 0.1 [-0.03, 0.21], $p = 0.132$). There were no other effects relating to the drug condition on subjective drug effect ratings (Appendix 3.3).

With respect to punishment unpleasantness, the main effects of drug condition and timepoint on unpleasantness ratings were not significant (drug condition: β = -3.4 [-9.49, 2.77], $p = 0.281$, timepoint: β = -4.2 [-10.33, 1.94], $p = 0.179$); and their interaction was also not significant (β = 1.5 [-9.16, 12.16], $p = 0.781$). A similar analysis was not possible in the threat-of-shock data since unpleasantness ratings were only recorded once. Ratings were moderately high in both experiments (lorazepam experiment: mean = 27.0 (on a scale from 0 – not at all aversive, to 50 – extremely aversive), SD = 13.8, threat-of-shock experiment: mean = 26.7, SD = 11.8), suggesting that the punishments were indeed perceived as aversive.

3.5.7 Exploratory analyses: symptom effects

To explore whether there any associations with psychiatric symptoms, I tested whether symptom levels modulated the effect of the two interventions on behaviour. Specifically, I included the scores for the Generalised Anxiety Disorder (GAD-7), Brief Experiential Avoidance Questionnaire (BEAQ) and inhibition sub-scale of the Behavioural Activation/Inhibition Scales (BIS) as moderators of the effect on condition in the hierarchical logistic regression models predicting trial-level choices. Each scale was included as a main effect and also tested for its interaction with experimental condition. The p-values for the coefficients relating to the symptoms were then adjusted for multiple comparisons via the Bonferroni method.

No coefficients relating to psychiatric symptoms achieved significance in the threat-of-shock experiment, but in the lorazepam experiment, individuals with higher self-reported behavioural inhibition were significantly less likely to choose the conflict option overall (β = -0.18 [95% CI: -0.28, -0.08], $p_{\text{Bonferonni}} = 0.002$). There was also a significant interaction between the effects of lorazepam and behavioural inhibition on choice (β = 0.15 [0.10, 0.20], $p_{\text{Bonferonni}} < 0.001$). I explored this interaction effect using post-hoc tests in which I specified hierarchical logistic regression models predicting choice, separately under the lorazepam and placebo conditions, and assessing the main effect of behavioural inhibition in each model. This showed that behavioural inhibition was a negative predictor of conflict option choices in the placebo condition (β = -0.15 [-0.26, -0.04], $p = 0.009$), but not in the lorazepam condition (β = -0.02 [-0.13, 0.09], $p = 0.729$). This indicates that behavioural inhibition is associated with a preference for avoidance over approach in the placebo condition, which is attenuated following lorazepam administration (Figure 3.6A). There were no significant effects relating to generalised anxiety symptoms and experiential avoidance - the full hierarchical model results are presented in Appendix 3.4.

I followed up on these findings by investigating whether behavioural inhibition also modulated the effects of the interventions on model parameter values. Here, behavioural inhibition score was examined both as a main effect as an interaction with experimental condition to predict parameter values in hierarchical linear regression models, with p-values again adjusted via the Bonferroni method. There was a significant interaction between the effect of lorazepam and behavioural inhibition score on punishment learning rates (β = -0.02 [-0.03, -0.01], $p_{\text{Bonferonni}} = 0.009$). Using a similar approach to explore this interaction effect to that taken above, I conducted post-hoc tests by specifying linear models predicting the punishment learning rate from behavioural inhibition scores separately under each drug condition. This showed that behavioural inhibition positively associated with punishment learning rate in the placebo condition, although this did not achieve significance (β = 0.02 [-0.01, 0.05], $p = 0.162$), whereas the analogous effect was much smaller in the lorazepam

condition (β = 0.002 [-0.02, 0.02], p = 0.853). These results suggest that lorazepam could influence the relationship between behavioural inhibition and punishment learning rate, though this interpretation should be considered cautiously given the non-significant effects (Figure 3.6B). Behavioural inhibition did not significantly predict the variance of any other parameter values - the full hierarchical model results are presented in Appendix 3.4.

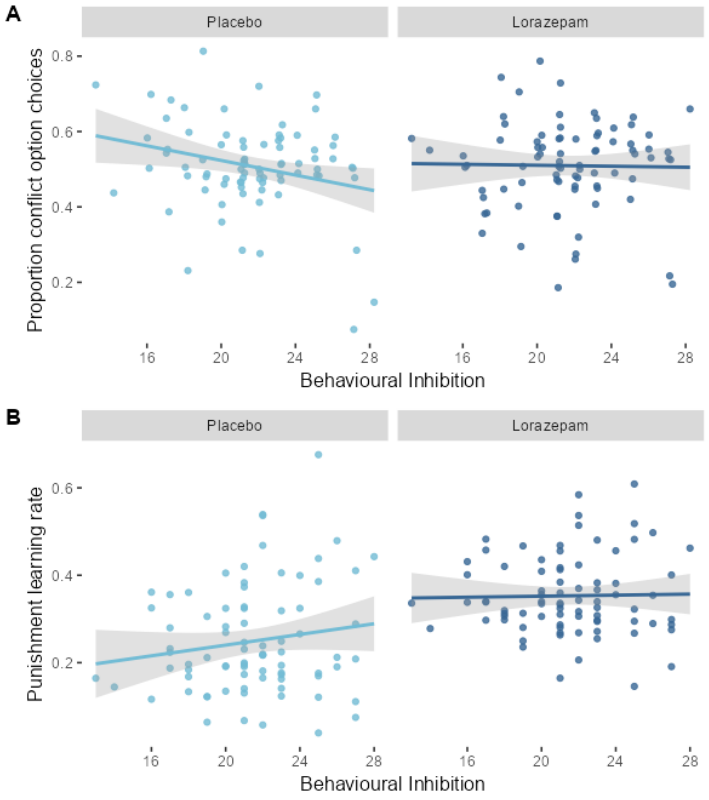


Figure 3.6, Behavioural inhibition modulates the effect of lorazepam on task performance. A, *The relationship between behavioural inhibition and the proportion of conflict option choices under placebo (left) and lorazepam (right) conditions. Higher behavioural inhibition was associated with a decreased proportion of conflict option choices under placebo, indicating a preference for avoidance over approach. This effect is normalized under lorazepam administration.* ***B,*** *The relationship between behavioural inhibition and punishment learning rate under placebo (left) and lorazepam (right) conditions. Under placebo, there is a positive relationship between behavioural inhibition and punishment learning rate. Under lorazepam, this relationship is greatly diminished, suggesting that lorazepam normalizes the association between behavioural inhibition and punishment learning rate. Note. Data points represent individual participants, and the lines represent the regression slopes with 95% confidence intervals shaded.*

3.6 Discussion

By experimentally modulating state anxiety through lorazepam administration and threat-of-shock, this chapter aimed to establish causal links between anxiety, avoidance, and the underlying computational processes. The findings from these two parallel within-subjects experiments provide novel insights into the mechanisms underlying anxiety-related avoidance. Specifically, lorazepam administration, despite not significantly reducing self-reported anxiety or avoidance, shifted sensitivities towards rewards relative to punishments. Conversely, threat-of-shock increased state anxiety and avoidance behaviour, while biasing computational sensitivities towards punishments

relative to rewards. The complementary and opposing effects of these anxiety manipulations on computational mechanisms provide consistent evidence that anxiety causally impacts reward-punishment sensitivities, which in turn mediate avoidance behaviour. This bidirectional modulation of the putative mechanism strengthens my theory of the causal pathway from anxiety through altered reward-punishment processing to avoidance, even when overt behavioural changes are subtle.

3.6.1 Implications

The results from the threat-of-shock experiment confirmed my hypotheses about how anxiety impacts approach-avoidance reinforcement learning. Participants were significantly more anxious under threat-of-shock compared to safety, and this heightened anxiety increased avoidance of the high-reward/high-punishment option – this provides causal evidence of the task's predictive validity with respect to the relationship between anxiety and avoidance. This finding is consistent with previous research suggesting that anxiety induces the avoidance of potential threats, even at the cost of forgoing substantial rewards (Clark et al., 2012; Pittig et al., 2015; Vogel & Schwabe, 2019). Further, the computational analyses provided a mechanistic interpretation of this effect, namely that heightened anxiety increases sensitivity to punishments relative to rewards, which thus leads to avoidance behaviour.

In addition to the significant pre-registered effect, there was evidence for an effect of threat-of-shock on all of the model parameters, indicating that the anxiety induced by threat had a broad effect on cognition during approach-avoidance learning. Specifically, I found threat to decrease both reward- and punishment learning rates, whereas it increased both sensitivity parameters. Despite this broad range of effects, the mediation analyses suggested that it was primarily the sensitivity parameters, especially the punishment sensitivity parameter, that explained the effect of threat on avoidance behaviour. Alongside the explanatory power of the reward-punishment sensitivity index, these results suggest that the impact of anxiety on task performance are primarily related to how individual trade-off rewards and punishments during decision-making, rather than how they learn about rewarding/punishing events.

Surprisingly, there was no effect of lorazepam on self-reported anxiety symptoms, either measured outside of the context of the approach-avoidance reinforcement learning task or during it. Yet, reassuringly, there was an effect on drowsiness ratings, indicating that the drug indeed had some effect in the present data consistent with its known psychological effects. Possibly owing to this lack of effect on anxiety, there was also no evidence for lorazepam's effect on the frequency of high-reward/high-punishment choices in the task. However, I did find evidence in favour of my pre-

registered prediction that lorazepam, compared to placebo, would bias individual sensitivities towards rewards over punishments.

That lorazepam affected the theoretical computational mechanisms of behaviour, but neither anxiety nor choices as measured via model-agnostic analyses, is interesting. One interpretation of this pattern is that the effects of the drug on anxiety and behaviour were small – the directions of effect were consistent with my predictions, but these effects did not reach significance, which may simply reflect low statistical power. Additionally, the participants in the experiment were not clinically anxious, so the weak effect on anxiety could have been due to a floor effect.

An alternative interpretation is to take these findings at face value: that lorazepam affects cognitive processes (i.e. how much weight to place on rewards vs punishment during value-based choice) but not directly anxiety. A commonly-cited theory of the cognitive mechanisms of antidepressants is that they take time to alleviate symptoms because the antidepressants primarily impact cognitive function, which takes time to translate to better mood (Harmer et al., 2009). An analogous interpretation could be taken for my results – that lorazepam affected cognitive function, but this did not translate to lower anxiety given the context of the non-anxious sample.

This interpretation may also account for why I found an effect of lorazepam on the computational parameters but not overall choices. This discrepancy is perhaps unexpected, as the assumption of these parameters as mechanisms of learning behaviour implies that an effect on the mechanisms (i.e. model parameters) should correspond with a change in the outcomes (i.e. choices). However, this may be a case of lorazepam acting on an underlying computational process related to anxiety-related avoidance, without having observable effects on avoidance behaviour. Indeed, the mediation analyses suggested that the reward-punishment sensitivity index mediates lorazepam's effect on task performance. In the context of clinical trials, findings that a clinical intervention affects underlying processes but not immediate clinical endpoints are still seen as useful outcomes (Strimbu & Tavel, 2010), and perhaps the same can be said in the field of computational psychiatry (Nair et al., 2020). These computational measures can provide early indicators of treatment effects, reflecting underlying biological or cognitive processes that precede observable behavioural changes. However, further validation will be needed before identifying relative sensitivities to rewards/punishments as reliable indicators of anxiety-induced avoidance.

In support of the idea that the (lack of) effects of lorazepam on anxiety and behaviour in my experiment was dependent somewhat on a non-clinical sample, an interesting finding from the exploratory analyses was that the effect of lorazepam on behaviour was moderated by behavioural inhibition traits. Specifically, individuals reporting greater behavioural inhibition were more likely to

avoid the option associated with punishment under placebo, but this effect was abolished after lorazepam administration. A similar effect was also observed for learning from punishment, where individuals with greater behavioural inhibition showed greater punishment learning rates after lorazepam administration, but again not after placebo. These results provide initial evidence that lorazepam affects how individuals learn from punishment, depending on their (sub-clinical) behavioural inhibition traits, and this in turn has effects on avoidance behaviour.

3.6.2 Limitations/opportunities for development

As noted above, the lack of a significant reduction in self-reported anxiety under lorazepam suggests that the drug's anxiolytic effects in this non-clinical sample were weak. Additionally, the effects of the interventions on task performance were also subtle and not observable via simple t-tests, indicating that the task might lack sensitivity in its current form. A potential avenue for future work could be to optimise the task, for example by increasing the task length or refining the reward and punishment probabilities, to improve its sensitivity in detecting behavioural changes induced by anxiety-related interventions.

As alluded to above, the relatively subtle effects of the experimental manipulations on model-agnostic behavioural metrics, i.e. overall choice proportions, raises the question of how to interpret the observed changes in computational parameters, particularly the reward-punishment sensitivity index, in the absence of robust corresponding shifts in overt behaviour. One possibility is that the computational modelling approach is more sensitive to subtle, anxiety-related changes in decision-making processes than traditional behavioural measures. The model parameters may capture latent cognitive shifts that are not fully reflected in overall behavioural metrics like choice proportions. However, the lack of clear behavioural effects also suggests that the experimental manipulations, particularly the lorazepam administration, may have been too subtle to induce substantial changes in behaviour in this non-clinical sample. Future research could address this limitation by exploring other model-agnostic metrics, such as reaction times, or win-stay lose-shift probabilities, which could provide additional insights into how the experimental manipulations affected decision-making processes. Examining these metrics in conjunction with computational parameters could help to paint a more complete picture of the effects of anxiety manipulations on behaviour. Nevertheless, the computational modelling results presented here provide valuable insights into the potential mechanisms by which anxiety and anxiolytic interventions modulate decision-making, highlighting the utility of computational approaches in extracting latent cognitive processes, such as the distinct contributions of reward and punishment sensitivity, that may be obscured when relying solely on traditional behavioural measures.

However, the Bayesian analyses of the model parameters indicate additional limitations. While these analyses supported the same directional trends as our frequentist results for the relative sensitivity indices, aligning with our pre-registered predictions, they also indicated some uncertainty around these effects. The probabilities that both lorazepam and threat-of-shock affected the reward-punishment sensitivity indices in their predicted directions were around 78%, and both highest density intervals spanned zero, suggesting less certainty in the magnitude of these effects than indicated by the t-tests alone. At the same time, it should be noted that this uncertainty does not necessarily indicate null results for my pre-registered hypotheses, but rather that the observed evidence was relatively weak. Nevertheless, taken together these results indicate consistency in that manipulations aimed at decreasing anxiety bias sensitivities towards reward over punishments, whilst anxiogenic manipulations have the opposite effect.

Another limitation is my use of a non-clinically anxious sample, which may preclude the study of clinically relevant levels of anxiety and avoidance. However, there is growing interest in research involving 'experimental models of anxiety' in healthy control participants to study anxiety (Grillon et al., 2019). This approach can mitigate the limitations of using non-clinical samples by experimentally inducing anxiety in healthy controls and examining its impact on cognitive function. This indeed showed that such transient states of heightened anxiety, in the threat-of-shock experiment, biased behaviour towards avoidance and sensitivities towards punishments, relative to rewards – these findings could be used as predictions as to how clinically anxious patients may behave in the task. Future work with clinical samples will be needed to validate the clinical relevance of the task.

Indeed, the low level of baseline anxiety in the lorazepam study may be a key factor in the absence of observable anxiety effects. This highlights another potential avenue for future research: combining the anxiolytic manipulation (lorazepam) with the anxiogenic manipulation (threat-of-shock) in healthy participants. Such a design could create a more clinically relevant model of anxiety in non-clinical samples. For instance, one could examine whether lorazepam administration mitigates or cancels out the anxiety induced by threat-of-shock. Crucially, this approach would enable a test of whether lorazepam rescues threat-induced avoidance behaviour, and whether this effect, if identified, was mediated by changes in the relative sensitivity to rewards versus punishments. This would provide another robust test of this computational mechanism, demonstrating its ability to capture both increases and decreases in anxiety-related avoidance within the same individuals.

3.6.3 Computational validity

More broadly, the findings from these experiments underscore the computational validity of the approach-avoidance reinforcement learning task. Both the anxiolytic and anxiogenic interventions

affected the putative key computational mechanism driving avoidance behaviour (i.e. the relative sensitivities to rewards versus punishments) in theoretically consistent directions. This indicates that the task engages a mechanism that explains how anxiety leads to avoidance, thus providing a robust model for studying anxiety-related behaviours.

The development of the approach-avoidance reinforcement learning task as a cross-species translational analogue of preclinical measures of anxiety in Chapter 2 aimed to preserve the computational strategies across species. The results from this chapter provide causal evidence that the task's computational foundation is relevant to human anxiety, implying that the model parameters could serve as surrogate measures of anxiety in non-human subjects where verbal reports are not feasible. For instance, if a certain intervention affects the reward-punishment sensitivity index in rodents in a matched task, it implies anxiety-relevant consequences in humans. Moreover, the finding that lorazepam had observable effects on the computational parameters but not the model-agnostic summary statistics of behaviour suggests that preclinical studies could use this type of task and computational analyses to gain a more comprehensive understanding of an intervention's potential relevance in humans. This approach goes beyond traditional behaviour-based assessments and may help to better extrapolate findings from preclinical data to humans.

3.7 Conclusion

In this chapter, I showed that anxiolytic and anxiogenic interventions (a single dose of lorazepam and threat-of-shock, respectively) elicit opposing effects on a computational measure of approach-avoidance conflict - in the latter case, even in the absence of explicit changes in anxiety. Specifically, lorazepam administration did not significantly affect self-reported anxiety, but it shifted sensitivities towards rewards over punishments, although this did not translate to observable changes in avoidance behaviour in this non-clinical sample. On the other hand, threat-of-shock increased anxiety and biased individual sensitivities towards punishments over rewards, and did result in increased avoidance behaviour. These findings support the computational validity of the task as a model of anxiety-induced avoidance, with broader implications for its utility in cross-species translational research.

3.8 Appendices

3.8.1 Appendix 3.1. Mediation model specification

For the mediation analyses, I used hierarchical regression models to assess whether the model parameters mediated the effect of the interventions on behaviour. The model specifications differed based on the parameter being tested as the mediator.

To estimate the mediation effects of the base model parameters (i.e. the reward/punishment learning rates and sensitivities), I specified models in which each parameter was tested as a mediator, with the remaining three parameters included as covariates. Specifically, the mediator model was a hierarchical linear regression model, of the form:

$$\mathrm{Mediator} \sim \mathrm{Condition} + \mathrm{Covariate}_1 + \mathrm{Covariate}_2 + \mathrm{Covariate}_3 + (1 \mid \mathrm{participant})$$

where $\mathrm{Mediator}$ was a certain model parameter, $\mathrm{Condition}$ was the experimental condition, and $\mathrm{Covariate}_i$ were the remaining model parameters as covariates. The model parameter values were standardised to facilitate model estimation.

The outcome model was a hierarchical logistic regression model, of the form:

$$\mathrm{Choice} \sim \mathrm{Condition} + \mathrm{Mediator} + \mathrm{Covariate}_1 + \mathrm{Covariate}_2 + \mathrm{Covariate}_3 + (1 \mid \mathrm{participant})$$

where Choice was the trial-level choice data from the task.

Since the reward-punishment sensitivity index was a direct function of the reward and punishment sensitivity parameters, these parameters could not be estimated in the same model. Therefore, the reward-punishment sensitivity index was tested without covarying for other model parameters.

3.8.2 Appendix 3.2. Parameter reliability for condition-specific parameters estimated under independent vs multivariate group-level prior distributions

I used two alternative approaches to estimate the condition-specific parameters from the computational model of learning behaviour. One approach was to specify the hierarchical structure of the model such that each condition-specific parameter was drawn from an independent prior, with no relationships defined across the treatment- and control-specific parameters. In other words, the model was not given the information of the natural pairings of the parameters across conditions.

The second approach was to draw the parameter pairs across conditions (e.g. the reward learning rate under treatment and under placebo) from a common, multivariate Gaussian prior in untransformed parameter space. This effectively served to pool the data across conditions, allowing the model to use more data to infer the parameter values. As Figure 3.6 shows, this led to greater correlations of parameters across conditions, compared to the independent-priors specification, indicating that the reliability of the parameters is captured more faithfully in this model.

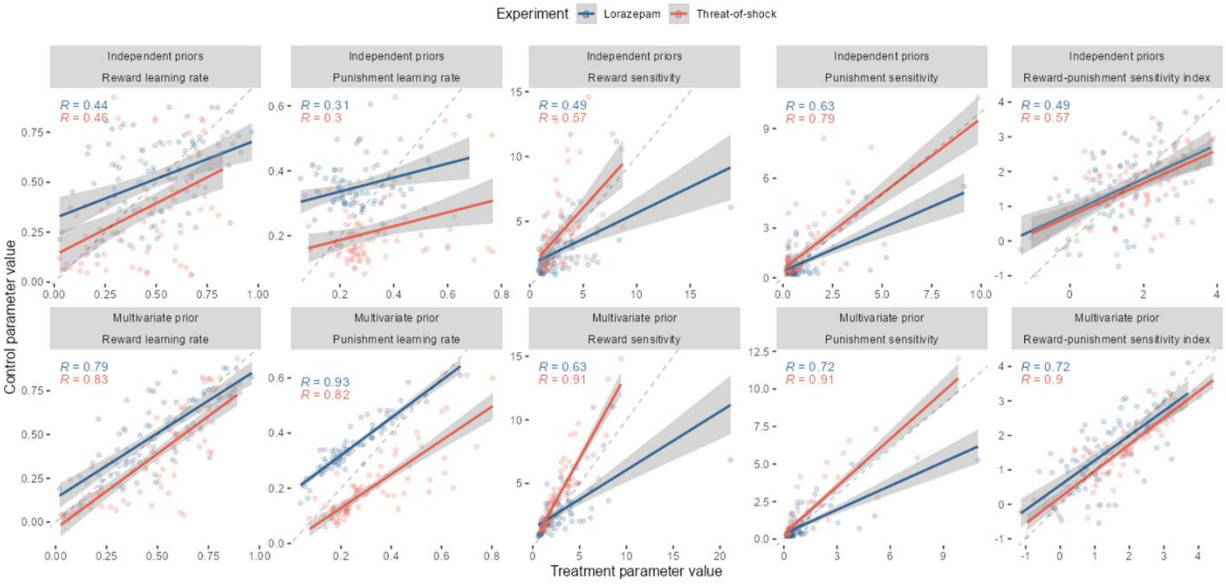


Figure 3.7, Scatter plots of estimated parameter values across conditions, by the prior specification method. *Points represent individual participants. Lines represent fitted linear relationships, with 95% errors. In general, the correlations of parameter values across conditions were lower when the data were estimated with independent priors (top row) compared to a multivariate prior (bottom row).*

3.8.3 Appendix 3.3. Subjective drug effect rating analyses

Table 3. Subjective drug effects in the lorazepam experiment			
Coefficient	**Dependent variable**	**Value [95% CI]**	**p value**
Drug condition	Anxiety	-0.10 [-0.21, 0.02]	0.099
	Drowsiness	0.08 [-0.02, 0.18]	0.132
	Nausea	0.12 [-0.10, 0.33]	0.284
	Inability to concentrate	0.03 [-0.04, 0.10]	0.346
Timepoint	Anxiety	-0.15 [-0.26, -0.04]	0.010*
	Drowsiness	0.28 [0.18, 0.39]	< 0.001*
	Nausea	0.23 [0.02, 0.44]	0.029*
	Inability to concentrate	0.10 [0.03, 0.17]	0.005
Drug condition x timepoint	Anxiety	-0.01 [-0.12, 0.11]	0.888
	Drowsiness	0.16 [0.06, 0.26]	0.003*
	Nausea	0.15 [-0.06, 0.37]	0.149
	Inability to concentrate	0.06 [-0.01, 0.13]	0.097
Session	Anxiety	-0.06 [-0.17, 0.06]	0.320
	Drowsiness	0.10 [-0.004, 0.20]	0.060
	Nausea	0.08 [-0.13, 0.29]	0.432
	Inability to concentrate	-0.004 [-0.07, 0.07]	0.918
Note. The values in the "Value" column represent the estimated regression coefficients (log-transformed) for the negative binomial regression models. These coefficients indicate that for a one-unit change in the predictor variable, the log of expected counts of the response variable changes by the respective regression coefficient.			

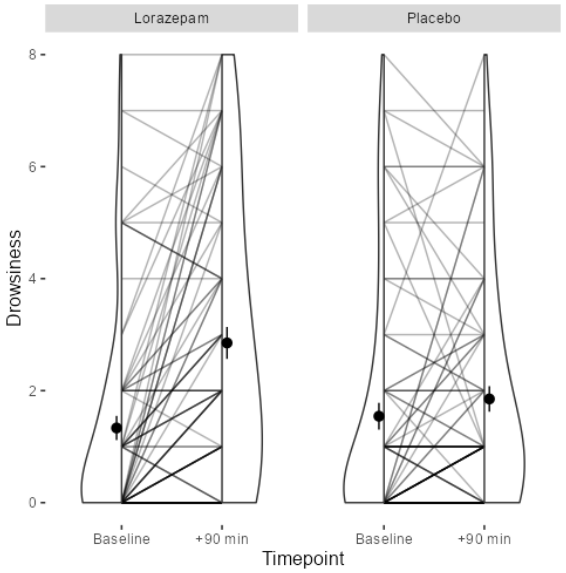


Figure 3.8, Drowsiness ratings by timepoint across drug conditions. *Lines represent individual participants. Points represent mean values, with error bars showing ±1 standard error. Lorazepam led to greater increases in drowsiness compared to placebo.*

3.8.4 Appendix 3.4. Exploratory analyses: regression tables

To explore whether symptoms modulated the effect of the anxiety manipulations on behaviour, I used hierarchical logistic regression models to predict trial-level choices, given the predictors of symptoms scores and experimental conditions (including their interaction). The general model specification was:

$$\text{Choice} \sim (\text{GAD} + \text{BEAQ} + \text{BIS}) \cdot \text{Condition} + (1 \mid \text{participant})$$

where GAD was the score from the Generalised Anxiety Disorder 7-item scale, BEAQ was the score from the Brief Experiential Avoidance Questionnaire, and BIS was the score from the Behavioural Inhibition Scale. Separate models were estimated for each experiment.

Table 4. Coefficients from the hierarchical logistic regressions of trial-level choices, including predictors of psychiatric symptoms.

Study	Coefficient	Value [95% CI]	Adjusted p-value	Original p-value
Threat-of-shock	GAD	0.07 [-0.11, 0.26]	1	0.423
	BEAQ	0.08 [-0.11, 0.26]	1	0.409
	BIS	-0.22 [-0.42, -0.02]	0.180	0.030
	GAD x condition	-0.08 [-0.14, -0.01]	0.141	0.023
	BEAQ x condition	0.07 [0.001, 0.14]	0.230	0.038
	BIS x condition	0.07 [-0002., 0.15]	0.333	0.056
Lorazepam	GAD	0.02 [-0.09, 0.12]	1	0.751
	BEAQ	0.10 [0.001, 0.20]	0.288	0.048
	BIS	-0.18 [-0.28, -0.08]	0.002	< 0.001
	GAD x condition	-0.003 [-0.05, 0.05]	1	0.911
	BEAQ x condition	-0.07 [-0.12, -0.02]	0.055	0.009
	BIS x condition	0.15 [0.10, 0.20]	< 0.001	< 0.001

Note. The adjusted p-values are values after applying Bonferroni correction.

These analyses showed that behavioural inhibition scores modulated the effect of lorazepam on task behaviour (Table 4). I followed up on this finding via hierarchical linear regression models to predict parameter values given the predictors of symptoms scores and drug condition (including their interaction). The general model specification was:

$$\text{Parameter} \sim \text{BIS} \cdot \text{Condition} + (1 \mid \text{participant})$$

where Parameter was the model parameter value across the lorazepam and placebo conditions. Separate models were estimated for each parameter. Since the key effect from the choice model was an interaction between behavioural inhibition and drug condition, I restrict the analyses here to this

analogous interaction (Table 5). This showed that behavioural inhibition significantly modulated the effect of lorazepam on punishment learning rates.

Table 5. Interaction coefficients (drug condition x behavioural inhibition) from the hierarchical linear regressions of parameter values in the lorazepam experiment.

Parameter (dependent variable)	Value [95% CI]	Adjusted p-value	Original p-value
Reward learning rate	0.01 [-0.02, 0.05]	1	0.391
Punishment learning rate	-0.02 [-0.03, -0.001]	0.009	0.002
Reward sensitivity	0.33 [-0.14, 0.80]	0.855	0.171
Punishment sensitivity	-0.02 [-0.24, 0.21]	1	0.883
Reward punishment sensitivity index	-0.02 [-0.19, 0.16]	1	0.851

Note. The adjusted p-values are values after applying Bonferroni correction.

4 A computational model of spatial exploration in aversive environments

4.1 Preface

In the previous empirical chapters, I have relied on a relatively simple behavioural task to probe anxiety-related cognitive processes. The simplicity of the task allows for a high degree of control over behaviour, which in turn limits the space of explanations to a narrow range – this is a useful feature with respect to the interpretability of observed data. As discussed previously, another benefit of a simple structure is that this simplicity also makes the task amenable to generative, theory-driven computational modelling. However, the real world is much more complex than binary decisions, and to understand anxiety in real-world situations, we need to embrace this complexity, which essentially requires us to work with more naturalistic behaviours (Mobbs et al., 2021; T. Wise et al., 2024; Yamamori & Robinson, 2024). Therefore, in the present chapter, I develop and validate a computational model of a task that involves richer behavioural complexity and is more ecologically valid, specifically how humans explore aversive spatial environments.

Note to examiners.

An initial conceptualisation of this chapter was published as part of a commentary on a wider review of computational approaches to translational psychiatry. The reference to this paper is:

Yamamori, Y., Robinson, O.J. Thinking computationally in translational psychiatry. A commentary on Neville et al. (2024). Cogn Affect Behav Neurosci (2024). https://doi.org/10.3758/s13415-024-01172-1

Special acknowledgements.

I would like to extend a special thanks to Professor Johannes Fuss and Dr Sarah Biedermann, who kindly agreed to share the raw movement data from their study (Biedermann et al., 2017) so that I could validate the computational model with empirical data.

4.2 Abstract

Behavioural paradigms that probe spontaneous spatial exploration in aversive environments are common in preclinical tests of anxiety, and recent human translation of these tests supports their validity as potential indices of some processes in humans driving psychiatric symptoms. However, such behaviours that involve free movement in space have yet to be understood from a computational perspective. Therefore, I propose a computational model of how agents explore novel, potentially aversive spatial environments, within a model-based reinforcement learning framework. The algorithm is relatively simple and able to account for a range of classical findings relating to exploration in aversive spatial environments. When I fitted this model to existing data from a mixed-reality version of the elevated plus maze with human participants, I found that the model was able to account for traditional model-agnostic metrics of anxiety-related behaviour, and the model-derived parameters provided new insights into the mechanisms of the effects of lorazepam and yohimbine on the task. The proposed model offers a simple but effective account of the computational mechanisms of spatial exploratory behaviour, in a way that is species-agnostic and generalisable to real-world cognition, therefore providing opportunities for more naturalistic measures in computational neuroscience and psychiatry.

4.3 Introduction

Computational psychiatry has produced important theoretical insights into how specific computational/cognitive processes are disrupted in psychiatric conditions such as anxiety and depression (Bishop & Gagne, 2018; Yamamori & Robinson, 2023), but the impact of these insights in clinical practice has been limited. One of the reasons for this may be field's reliance on simple and abstract lab-based tasks which, despite allowing for a high degree of interpretability, might be too constrained to generalise to behaviour in the real world. The increasing complexity and explanatory power of computational neuroscience models presents opportunities for investigating naturalistic measures - for example where individuals can freely move around space rather than choosing from abstract 'bandits' - that are more complex but are better reflections of everyday decision-making (Mobbs et al., 2021; T. Wise et al., 2024). Such measures promise greater ecological validity and potentially, by extension, clinical applicability of research findings. This chapter aims to explore this potential by applying the generative, computational approach to a common, naturalistic paradigm of anxiety-like behaviour: exploration of aversive spatial environments.

One of the most common ways to measure anxiety-like behaviour in rodents is to study their behaviour in novel spatial environments in which some regions of space are associated with aversive features (Griebel & Holmes, 2013). For example, rodents will typically avoid the brightly lit compartment of the light-dark box in favour of the darkened compartment (Crawley & Goodwin, 1980). These features-of-interest are avoided as they entail greater risk of predation in the wild; in other words, these are unconditioned survival-oriented responses. Since the behavioural tests rarely involve explicit or actual threats, avoidance of the environmental features is interpreted as 'anxiety-like' behaviour – this has been supported empirically by findings that benzodiazepines, which are anxiolytic in humans, reduce avoidance and instead promote exploration in rodents (Griebel & Holmes, 2013). Studies using human variants of these tests have observed similar behavioural patterns to that in rodents, and the individual differences in human behaviour also appear to correlate with psychiatric symptoms (Biedermann et al., 2017; Walz et al., 2016). These findings highlight the potential of this paradigm as a way to measure psychiatric processes in a naturalistic manner.

In the present chapter, I use model-based reinforcement learning algorithms (Sutton & Barto, 2018) to model how organisms navigate environments in a goal-directed, exploratory manner, specifically focusing on how they balance exploration with the avoidance of positions associated with potential threat. There are already demonstrations that model-based algorithms that encode a 'model' of the environment, such as a spatial map, can approximate how humans avoid future threat in abstract state spaces (Wang et al., 2018; Wise et al., 2021) as well as in spatial environments (Huys et al.,

2012; Wise et al., 2023). The term 'model-based' refers to the feature of these models which incorporate an understanding of the structure of the environment, which is sometimes referred to as a 'cognitive map' (Behrens et al., 2018; Tolman, 1948) - this enables agents to plan and choose future trajectories through space based on their anticipated outcomes. However, the studies on model-based avoidance in humans to date (Huys et al., 2012; Wang et al., 2018; Wise et al., 2021) have involved explicit rewards (e.g. money/points) and/or punishments (e.g. electric shocks), and to my knowledge, no study has yet modelled spontaneous exploration in environments without explicit reinforcements. This leaves open two research questions. First, how can we model avoidance when the aversive entities are not explicit events, but instead stable features of the environment? And secondly, how can we model exploratory behaviour if it is not explicitly rewarded (i.e. reinforced)?

I address both questions within the framework of a Markov decision process (MDP). MDPs offer a framework for modelling sequential decision-making problems, in which an agent chooses actions that lead to different states, each with associated rewards or punishments. In the context of spatial navigation, MDPs can conceptualise movement as a sequence of interdependent choices, where each choice affects potential future locations and must account for the potential rewards or threats associated with those locations. While MDPs describe the environment and its dynamics, reinforcement learning algorithms, particularly model-based reinforcement learning approaches, provide computational strategies for how agents can make optimal decisions within these environments. In this work, I use model-based reinforcement learning principles to approximate how agents navigate and make decisions in the MDP-structured spatial environments.

I attempt to answer the first question of how to model avoidance behaviour when aversive features are not explicit events but rather inherent in certain aspect of the environment, such as illumination level or temperature. These features can be encoded by assigning numerical values to locations based on their environmental characteristics, which then form the basis for hypotheses about which environmental features influence an organism's movement. Fitting models to empirical data can then allow one to test *whether* and *by how much* an organism is sensitive to the aspect-of-interest.

To address the second question of how to model spontaneous, uninstructed exploration, I draw on the concept of 'directed' exploration: a behavioural strategy that seeks to reduce uncertainty about the world by exploring uncertain or unfamiliar states in the environment (Auer et al., 2002). Computationally, this can be instantiated by augmenting the expected value of visiting a certain state with a bonus according to its uncertainty or informative value (Wilson et al., 2014). This bonus for seeking out uncertain or novel information serves to drive exploration of the environment to find potential rewards. In essence, exploration can be operationalised as the motivation to visit novel or

unfamiliar locations, for which an agent can experience an intrinsic reward. However, the challenge with this approach in the context of spontaneous exploration tasks is the absence of explicit rewards, which impedes the quantification of uncertainty of the value of a certain state.

Therefore, I turn to a related metric of uncertainty that is able to deal with sparse reward information – how often a state has been encountered previously (Bellemare et al., 2016). Specifically, counting the number of times a state has been visited in the past can act as a proxy for its uncertainty, which allows for a basis on which to explore without explicit reinforcements (Tang et al., 2017). There is already evidence that humans perform this strategy in exploring abstract decision-making tasks (Gershman, 2018), but not in the context of spatial environments.

In this chapter, I first explore a computational model of how agents explore spatial environments with potentially dangerous regions. Building on the framework introduced above, I implement a computational framework where exploratory behaviour involves an interaction between intrinsic rewards for novelty and extrinsic environmental factors. This approach allows me to frame navigation in potentially threatening environments as a decision-making problem, balancing the drive for exploration against the need for safety.

To model exploration behaviour that is not explicitly instructed or reinforced, I implement a dual-reward system comprising both intrinsic and extrinsic components. The intrinsic reward represents an agent's inherent motivation to explore novel or unfamiliar states, operationalised as a drive to visit locations based on how infrequently they have been encountered. This intrinsic reward decreases as a state becomes more familiar, modelling the declining novelty of repeatedly visited locations. In contrast, the extrinsic reward captures the influence of environmental features, such as potential threats or safety cues.

After the model exposition, I present simulations of how the model could be applied to explain a range of classical findings relating to naturalistic approach-avoidance behaviours. Then, I test the model against empirical data from a previous study on human exploration in a mixed-reality version of the elevated plus maze (Biedermann et al., 2017).

4.4 Methods

4.4.1 The computational model

I frame exploratory behaviours in two-dimensional space as an MDP, which describe how agents act in discrete-time environments.

Modelling movement

First, I model movement using a framework that discretises both space and time, an approach fundamental to MDPs and reinforcement learning. This method, often referred to as a 'grid-world' representation, approximates continuous motion within a simplified, computationally tractable structure.

In this model, the continuous two-dimensional space of the environment is divided into a finite grid of discrete locations, henceforth referred to as states, s, in a two-dimensional grid (Figure 4.1A, B). Time is similarly discretised into steps, t, with each step representing a moment when the agent makes a decision. This discretisation allows me to represent complex, continuous environments and behaviours in a manageable form for modelling.

At each time step, agents can choose from a set of discrete actions, a. These actions correspond to movements in the four cardinal directions (north, east, south, west) or the option to remain stationary. While this simplifies the continuous range of movements possible in reality, a sequence of these discrete actions can effectively approximate smooth, continuous motion over time. The finer the grid and the smaller the time steps, the closer this approximation comes to continuous movement.

It is worth noting that while MDPs often incorporate stochastic state transitions, I implement deterministic transitions in this model. This choice reflects the highly controlled nature of exploration-avoidance tests and my focus on decision-making rather than movement execution. In these experimental settings, subjects (whether human or animal) typically have full control over their movements within a stable, predictable environment. I assume that when an agent decides to move in a particular direction, they do so without error – i.e. without any chance of 'slipping' or inadvertently moving in an unintended direction. This deterministic approach allows me to model volitional movement, without the added complexity of accounting for motor noise or environmental unpredictability.

Modelling exploration via state visit counts

To model exploration behaviour that is not explicitly instructed or reinforced, I utilise a concept of exploration from the artificial intelligence literature: count-based exploration (Strehl & Littman, 2008). The intuition of this approach is that exploration can be operationalised as a drive to visit novel or unfamiliar states. To motivate agents to explore, I enable them to experience an intrinsic reward r_{int} upon entering a state s, which becomes discounted by the number of times s has previously been visited ($n(s)$) as:

$$r_{int}(s) = \frac{1}{(n(s)+1)^{\tau}}$$

As $n(s)$ increases (i.e. as s becomes more familiar), the intrinsic reward experienced at s decreases, providing less motivation to move there (Figure 4.1C, D). τ is a free parameter that controls the rate of decay of intrinsic reward by n, with higher values leading to more rapid reward decay. Put in cognitive terms, this parameter could be interpreted as how quickly an agent becomes familiar with a novel state, and therefore less motivated to return, after repeated visits. r_{int} is referred to as *intrinsic* as the environment does not itself provide anything of value for agent, but rather it is an internal property of the agent.

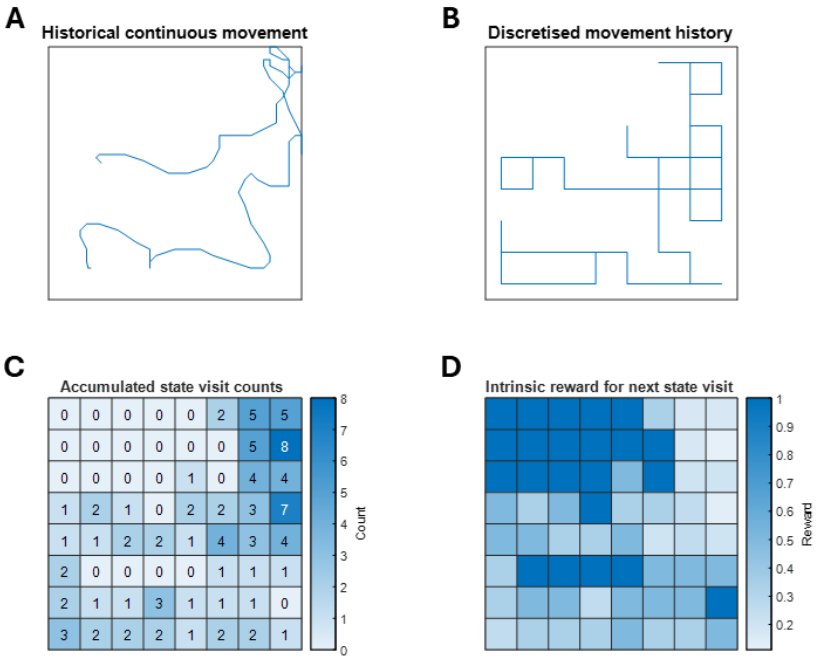


Figure 4.1, The computational model of exploration in aversive spatial environments. A, *Example history of continuous movement in a two-dimensional environment.* ***B,*** *Discretisation of the movement history into a discrete state space. The new position data takes the form of occupancy of states in Cartesian space.* ***C,*** *Accumulated state visit counts. Given the discretised movement data, the number of times each state has been visited is counted.* ***D,*** *Intrinsic reward distribution for the next state visit, given the history of state visits. The intrinsic reward available at each state given the history of movement is proportional to the inverse of the state visit counts. States which have not been visited frequently provide higher values of intrinsic reward. High state visit counts lead to lower intrinsic reward at the next visit to the state.*

Modelling the effect of environmental features

Aside from intrinsic exploration rewards, I enable agents to experience certain rewarding and/or aversive outcomes upon entering a state, depending on (predefined) environmental features, $E(s)$. For example, an agent entering a warm region of space (represented by a positive value of $E(s)$) might experience a small reward, whereas entry into a cold region (with low $E(s)$) might elicit a negative experience. Formally, these experiences can be modelled as extrinsic rewards, r_{ext} (which can be positive or negative), experienced after entering a certain state: $r_{ext}(s)$, according to:

$$r_{ext}(s) = E(s)$$

As I will discuss later, the values and distribution of $E(s)$ will depend on the specific environment.

Modelling goal-directed behaviour

I model goal-directed agents as those motivated to maximise overall reward, R, which is composed of both intrinsic and extrinsic reward:

$$R = \beta_{int} \cdot r_{int} + \beta_{ext} \cdot r_{ext}$$

where β_{int} is a free parameter which controls the degree to which an agent is motivated to explore and β_{ext} is a free parameter which controls the sensitivity of an agent to the environmental feature. Higher values of β_{int} indicate greater motivation to explore and in turn reduce avoidance behaviour, whilst higher values of β_{ext} indicate greater sensitivity (i.e. aversion) to certain environmental features, which increase avoidance.

When deciding on which direction to travel from a certain state, agents must consider not just the immediate consequences of their actions but also those in the future. This idea is captured by representing state values, $V(s)$, which depend on two components: the immediate reward at that state, $R(s)$, and potential future rewards discounted by time, given by the Bellman equation as:

$$V(s) = R(s) + \gamma \sum_{s'} P(s'|s)V(s')$$

where $P(s'|s)$ denotes the probability of transitioning to subsequent state, s', from state s, and γ is the discount parameter that controls the degree to which an agent discounts future reward. Only the immediately successive state, s', is necessary in the equation since the Markov property of MDP states ensures that s' is a sufficient statistic of the future (Sutton & Barto, 2018).

Across the breadth of model-based reinforcement learning algorithms, the transition probability matrix, P (i.e. all state-to-state values of $P(s'|s)$), is commonly conditional on a policy in model-based algorithms and can also change over time as agents learn more about the environment. However, since the behavioural tests I model here involve deterministic spatial environments, P was fixed such that there were equal probabilities of transitioning from a certain state to any of its adjacent states. This formulation effectively models agents as remaining unbiased towards particular directions of travel (after accounting for state rewards), with P serving as a static map by which to navigate and plan around the environment. This also worked to simplify the model, as the state value matrix V can be solved analytically via the matrix form of the Bellman equation:

$$V = R + \gamma P\mathrm{V} = (I - \gamma P)^{-1}R$$

where I is the identity matrix. This effectively served as a mechanism akin to the successor representation algorithm (Dayan, 1993), which efficiently manages changes in reward information if the transition dynamics of the environment remain stable. It allows agents to flexibly adapt to reward (such as decreases in intrinsic rewards associated with certain states) without having to perform full model-based planning that is computationally costly at each time step. As the successor representation depends on an action policy (i.e. the most likely action given a certain state), my implementation of an unbiased transition probability matrix equates to a successor representation under a random-action policy.

In line with convention, I modelled choice of action with a softmax function:

$$P(a|s) = \frac{e^{V(s'|s,a)}}{\sum_a e^{V(s'|s,a)}}$$

As we will see, combining extrinsic (i.e. environmental) rewards with intrinsic exploration motivation can explain a range of interesting behaviour in certain environments, specifically when external and internal rewards compete – i.e. when there is approach-avoidance conflict, or in this case, *exploration*-avoidance conflict.

4.4.2 Environmental feature specification

To explore the model's prediction in realistic experimental settings, I simulate the behaviour of agents in three classical preclinical tests of anxiety: the light-dark box, the elevated plus maze and the open field test. In all tests, organisms (usually rodents) are free to explore an environment in which some regions have features that are aversive to the organism.

Aversion to open spaces: the open field test (Figure 4.2)

The open field test is typically a square- or circle-shaped enclosed arena with no other environmental features. Here, a common behaviour-of-interest is *thigmotaxis*, which refers to an organism's preference to move along walls/edges rather than traversing open spaces. This behaviour is often interpreted as deriving from a fear of open spaces due to increased risk of predation. To computationally model this fear, I operationalise it as a 'distance from safety' signal, where safety is the proximity to a nearby wall/edge.

Let s denote the current state and z denote an edge state (i.e. one at the perimeter of the environment), with positions represented as Cartesian co-ordinates (e.g. position of s: (x_s, y_s)). Since movement in this grid-world environment is limited to north, east, south, west directions, Manhattan distance gives a measure of distance between two states, computed as $|x_i - x_j| + |y_i - y_j|$ between

states i and j. Therefore, the distance from safety of a certain state, $E(s)$, can be modelled as the Manhattan distance to the nearest wall/edge, according to:

$$E(s) = -1 \cdot \min_{z \in Z}(|x_s - x_z| + |y_s - y_z|)$$

where Z denotes the set of all states adjacent to a wall/edge in the environment. $E(s)$ was normalised to range $[-1, 0]$.

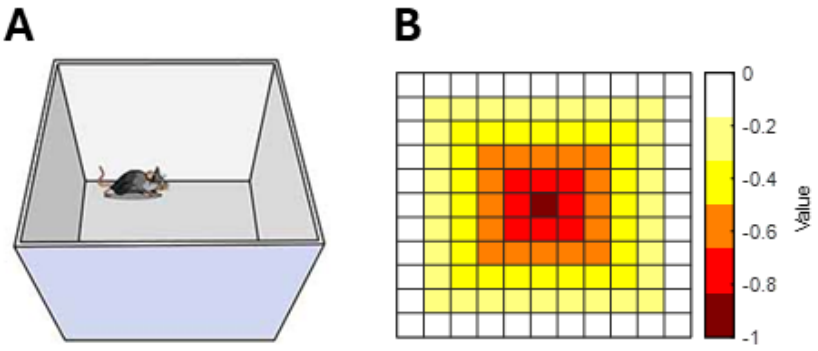


Figure 4.2, The open field test. A, *Schematic representation of the open field test. The test consists of a square enclosure with high walls. Rodents are placed in the arena and their exploratory behaviour is observed.* ***B,*** *To model aversion to open spaces, I label each state according to its distance from the nearest wall. The plot shows the extrinsic value at each state by colour – central regions are associated with more negative extrinsic value. Note: Figure A is reproduced from Figure 1 of La-Vu M, Tobias BC, Schuette PJ and Adhikari A (2020) To Approach or Avoid: An Introductory Overview of the Study of Anxiety Using Rodent Assays. Front. Behav. Neurosci. 14:145. doi: 10.3389/fnbeh.2020.00145.*

Aversion to light: the light-dark box (Figure 4.3)

The light-dark box involves a box consisting of two compartments: a darkened compartment and a light compartment. The behaviour-of-interest here is *scototaxis*, which is when an organism prefers and consequently moves towards regions of darkness over light regions. This behaviour is interpreted as an anxious response to the increased risk of predation in brighter regions of space. To model the potential aversiveness of light, I simply encode the presence of light in a certain state, s, as $E(s) = -1$ for states in the light compartment of the box, whereas $E(s) = 0$ for states in the darkened compartment. The negative sign captures the assumption that the relevant species of organism is averse to light (although this can be adapted for other species with differing attitudes to light).

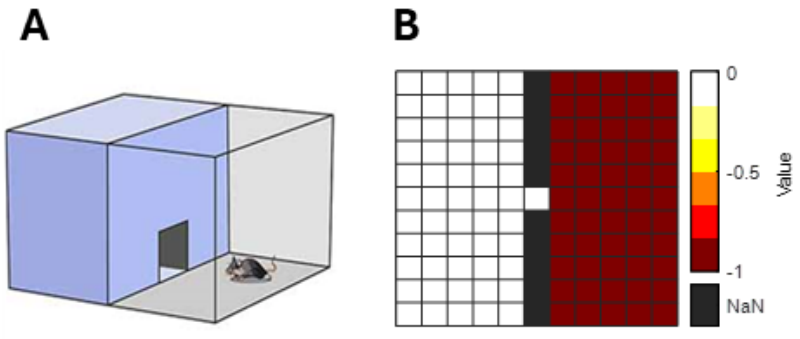


Figure 4.3, The light dark box. A, *Schematic illustration of the light-dark box test. This apparatus consists of two connected compartments: one brightly lit (shown in light blue) and one dark (shown in grey). Rodents are typically initially placed in the dark compartment and their exploratory behaviour is observed.* ***B,*** *To model aversion to brightly lit spaces, I label the states according to which compartment it belongs to. States in the light compartment have negative value, whilst those in the dark compartment are neutral. Note: Figure A is reproduced from Figure 1 of La-Vu M, Tobias BC, Schuette PJ and Adhikari A (2020) To Approach or Avoid: An Introductory Overview of the Study of Anxiety Using Rodent Assays. Front. Behav. Neurosci. 14:145. doi: 10.3389/fnbeh.2020.00145.*

Aversion to open and elevated spaces: the elevated plus maze (Figure 4.4)

The elevated plus maze involves an elevated platform consisting of two sets of arms, arranged in a plus shape. A pair of opposing arms are enclosed with barriers, which may provide perceived safety, whilst the other arms are exposed. This contrast therefore probes an organism's aversion to exposed, elevated spaces. As in the light-dark box, I simply encoded the value of a state as negative if it was on the open arms of the maze ($E(s) = -1$) or 0 otherwise ($E(s) = 0$).

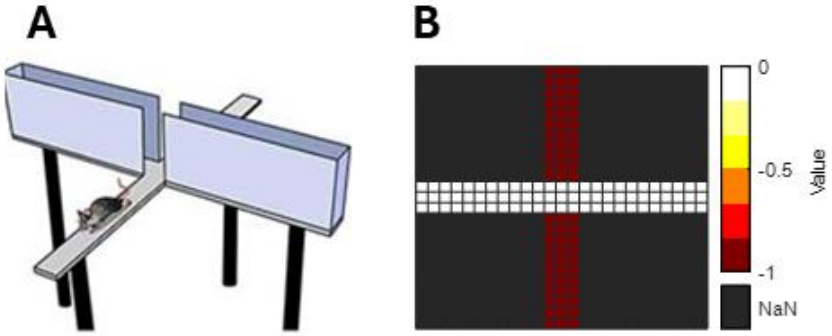


Figure 4.4, The elevated plus maze. A, *Schematic representation of the elevated plus maze. This apparatus consists of two open arms and two closed arms arranged in a plus shape, elevated above the ground. Rodents are typically initially placed on the centre of the platform and their exploratory behaviour is observed.* ***B,*** *To model aversion to open and elevated spaces, I label the states according to which arm they are located at. States on the open arms have negative value, whilst the remaining states are neutral. Note: Figure A is reproduced from Figure 1 of La-Vu M, Tobias BC, Schuette PJ and Adhikari A (2020) To Approach or Avoid: An Introductory Overview of the Study of Anxiety Using Rodent Assays. Front. Behav. Neurosci. 14:145. doi: 10.3389/fnbeh.2020.00145.*

4.4.3 Empirical dataset

To validate the model with empirical data, I re-analysed an existing dataset from a study implementing a mixed-reality version of the elevated plus maze in human participants (Biedermann et al., 2017). Briefly, the dataset consists of two studies. In Experiment 1, 100 participants were tested on the elevated plus maze. In Experiment 2, 51 participants were randomly allocated to receive an acute dose of lorazepam, yohimbine (an indole alkaloid that acts as a pro-anxiety drug) or placebo, after which they were tested on the elevated plus maze. In both experiments, the participants were healthy individuals without psychiatric disorders, and the test involved five minutes of free exploration on the mixed-reality elevated plus maze. The maximum length of the maze was 350cm. For a full overview of the data, I refer the reader to Biedermann et al. (2017).

As I was primarily interested in the raw movement data in this chapter, I excluded two participants' datasets based on the quality of the movement data. From Experiment 1, I excluded one participant's data as it contained one instance of an extremely large change in position across adjacent timepoints, and another participant's data as they appeared to spend a large portion of time seemingly off the platform, both of which would be indicative of errors in location recording. I also excluded two participants from Experiment 2 whose drug labels were missing. This led to final samples sizes of n = 98 and n = 49 for Experiments 1 and 2, respectively.

Data preparation

To allow for the modelling of discrete movement actions, I discretised the raw movement data into a grid-like Cartesian space. I ensured a balanced representation of the maze's central location by opting for an odd-numbered map width, thus incorporating a central state. This promotes fair exploration and minimises directional biases that might otherwise affect initial placements. To determine the unit width of the Cartesian space, I considered the resulting width of the extended arms of the platform (as these widths are directly proportional). To keep the overall number of possible states low whilst allowing for some granularity in position data on the extended arms, I opted for a map width which enabled an arm width of 3 units.

To discretise the position data over time, I set the initial discretised position to the nearest Cartesian co-ordinate of the raw starting position. Then, after each time step, the next discretised position was set as the Cartesian co-ordinate with the lowest Euclidean distance to the raw position whilst fulfilling two criteria: first, subsequent discretised positions could only be at most one step in either the x or y axis from the previous position (but not both, i.e. without diagonal movements). Second, if the raw position was not on the platform in discretised space (which could occur due to rounding in

the discretisation process or by participants peeking their heads and the attached motion sensors over the platforms), the discretised position was adjusted to the nearest on-platform position.

4.4.4 Experiment 1: model validation

I fitted variants of the computational model using an expectation maximisation algorithm (Huys et al., 2011) – the algorithm and supporting code is available online at https://www.quentinhuys.com/pub/emfit/. The models were fitted hierarchically, with the algorithm estimating both group-level and individual-level parameters simultaneously. Briefly, the group-level structure involved modelling individual parameters as drawn from independent Gaussian distributions in untransformed parameter space – this effectively served to regularise extreme parameter values via shrinkage. Parameters with both lower and upper bounds were transformed within the model via the sigmoid function, and those with only lower bounds were exponentiated within the model. I compared across models using the integrated Bayesian information criterion (Huys et al., 2011) which combines both model likelihood and complexity into a single measure of model evidence.

To test the contributions of each parameter, I compared across several models. The full model included all four parameters, namely the intrinsic and extrinsic reward weights, the intrinsic reward decay parameter and the temporal discount parameter. For some models, I tested whether removing the agents' sensitivities to the environmental features improved model fit by setting the extrinsic reward weight to 0. For other models, I also fixed the intrinsic reward decay to 0.5, a value that is considered to be optimal with respect to information gain (Strehl & Littman, 2008). For other models, I removed the mechanism of future planning by fixing the temporal discount parameter to 0, which effectively meant that agents could only consider rewards at immediately adjacent states when deciding to act.

To compare between directed and undirected exploration mechanisms, I also included a model incorporating a random exploration mechanism. This was implemented by simply including a single, positive-bounded parameter, b, that decreased an agent's probability of choosing to remain in the same state, Specifically, the actions values of all actions were set to 0, except for the value of remaining in the same state which was set to $-b$: thus higher values of b lead to greater likelihood of movement in any direction. The full model list is reported in Table 4.1.

Table 4.1, Model specifications

Index	Model name	Intrinsic reward weight	Intrinsic reward decay	Extrinsic reward weight	Temporal discount	Random exploration
1	INT	✓				
2	EXT			✓		
3	INT+DEC	✓	✓			

4	INT+EXT	✓		✓		
5	INT+DEC+EXT	✓	✓	✓		
6	INT+TD	✓			✓	
7	INT+DEC+TD	✓	✓		✓	
8	INT+EXT+TD	✓		✓	✓	
9	FULL	✓	✓	✓	✓	
10	RANDOM					✓

Note. Checkmarks (✓) indicate the inclusion of a parameter in the corresponding model. Model names reflect the parameters included: INT - intrinsic reward weight, DEC - intrinsic reward decay, EXT - extrinsic reward weight, TD - temporal discount. The RANDOM model represents purely random exploration. Models which did not include a specific reward weight parameter fixed its value at 0. Models which did not include the intrinsic reward decay parameter fixed its value at 0.5. Models which did not include the temporal discount parameter fixed its value at 0.

4.4.5 Experiment 2: drug effects

To assess the putative computational mechanisms of lorazepam and yohimbine on task behaviour, relative to placebo, I took forward the most parsimonious of the models fitted to the data from Experiment 1 and used it for inference in Experiment 2. As there were fewer participants in Experiment 2, this protected against a bias for selecting simpler models in this dataset.

To estimate models hierarchically in between-groups designs such as in this experiment, there are two common approaches. One is to estimate all groups jointly under a single prior which leads to shrinkage towards the full group-level mean, and the other is to estimate the groups under independent priors (i.e. one for each group) which leads to within-group shrinkage. Based on a simulation study which found that the former can severely underestimate true group differences (Valton et al., 2020), I opt for the independent priors approach.

4.4.6 Statistical analyses of model parameters

Given the findings from the original study that the model-agnostic behavioural metrics correlated with self-reported anxiety during task performance, I conducted correlation tests across the parameters from the winning model and self-reported anxiety in Experiment 1, to determine whether these putative model-based cognitive mechanisms were also associated with anxiety. As the model parameters typically violate Gaussian distributional assumptions, I used Kendall's tau to test for correlations – I report approximate Pearson's r value given tau, following the convention of $r = \sin(\tau \cdot \pi/2)$. I note that the dataset I received from the authors only contained 63% of the original self-reported measures, therefore the results that follow concerning the self-reported measures will not be totally consistent with their findings in this regard. I flag this caveat below with the relevant analyses where applicable.

I conducted mediation analyses to investigate if the model parameters could explain the effects of anxiety on behaviour. Specifically, I specified the model with anxiety as the predictor of the model-agnostic metric (e.g. time spent on the open arms), with all the parameters from the most

parsimonious model as parallel mediators of this relationship. The mediation model parameters were estimated via bootstrapping with 50,000 iterations per model.

To test for drug effects on the model parameters in Experiment 2, I conducted one-way tests comparing the effects of drug condition (lorazepam, yohimbine, placebo) on parameter values via Kruskal-Wallis tests. The Kruskal-Wallis test statistic is reported as H_2, where the subscript 2 indicates the degrees of freedom (number of groups minus 1). Given the sufficiently small sample size (n = 49), I computed exact probabilities for each test. I conducted post-hoc tests where applicable by comparing the effects of each drug to placebo via Mann-Whitney U tests (i.e. two post-hoc comparisons per test). I did not control for multiple comparisons given that there were only two comparisons.

4.5 Results – model simulations

4.5.1 Exploration behaviour

First, I show that the model can generate exploratory behaviour. Here, I simulated the movement of agents within a maze-like environment, which required agents to navigate around impassable walls to fully explore the available space (Figure 4.5A). In each simulation, agents began in a corner of the environment and executed a fixed number of movement actions (consistent across all agents). For these simulations, I did not specify any environmental features as potentially aversive. Therefore, each agent was produced with a random combination of three parameters: the intrinsic reward weight (β_{int}), the intrinsic reward decay (τ) and the temporal discount (γ) - each parameter was drawn randomly from independent uniform distributions. To quantify exploration, I measured the total distance covered by each agent, defined as the number of unique states visited. As Figure 4.5 shows, the intrinsic reward weight parameter was strongly, positively associated with total distance covered (r = 0.53, p < 0.001; Figure 4.5B, E), indicating that agents motivated more highly by intrinsic rewards explored more extensively into the maze. Agents with near-zero intrinsic reward weight parameters tend to move at random, and so fail to explore the deeper parts of the maze. The intrinsic reward decay was also, albeit less strongly, associated with total distance covered (r = 0.23, p < 0.001; Figure 4.5C, F; Steiger comparison to intrinsic reward weight correlation: Z_H = 4.28, p < .001). Here, low decay values indicate that the agents grow less 'bored' of a state over time, meaning that they remain in familiar locations for longer before travelling to unvisited locations – this explains why near-zero decay parameters lead to lower exploration. The temporal discount parameter was not associated with total distance covered (r = -0.08, p = 0.188; Figure 4.5D, G). These simulations confirm that the parameters intended to model exploratory behaviour effectively influence agent movement as expected.

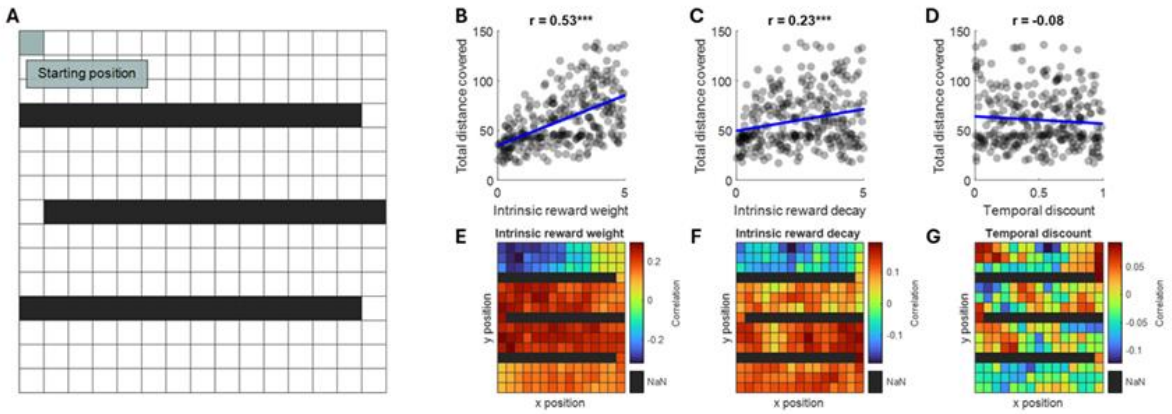


Figure 4.5, Simulating the effects of the model parameters on exploration in a maze-like environment. A, *The spatial environment. To test how well the artificial agents could explore space, I constructed an environment involving barriers that the agents needed to navigate around to access other parts of the environment.* ***B-D,*** *The effect of each model parameter on exploration. I simulated 300 agents with random parameter values, and tested how parameter values affect exploration, which I quantified as the total distance covered. The intrinsic reward weight and decay were positively associated with exploration, whilst the temporal discount parameter was not associated with exploration performance. Significance levels shown according to the following: p < 0.05 – *; p < 0.01 – **; p < 0.001 – ***.* ***E-G,*** *The effect of parameter values on position. For each state, I computed the correlation between parameter values and state visit counts across agents. Warmer colours indicate that agents with higher values of the relevant parameter were more likely to visit the state, and vice versa for cooler colours. The results show that higher values of the intrinsic reward weight drive exploration to further parts of the environment. A similar, but weaker effect can be seen with the intrinsic reward decay. There was not a clear effect of the temporal discount parameter on positions.*

4.5.2 Simulations – demonstration of classical approach-avoidance conflict effects

Next, I assess the model's ability to reproduce the classical findings from the open field test, light-dark box and elevated plus maze. These simulations were similar to those in the previous section, with the exception that the agents were now sensitive (to varying degrees) to each environment's feature-of-interest (e.g. the central position in the open field test). This was modelled by producing agents with four parameters: the intrinsic reward weight (β_{int}), the intrinsic reward decay (τ), the extrinsic reward weight (β_{ext}) and the temporal discount (γ). Again, each parameter was drawn randomly from independent uniform distributions. For each test, I quantified behaviour according to the standard metrics used to measure avoidance behaviour in each test.

The open field test

Traditionally, avoidance in this test is measured by time spent in the central zone. However, since the selection of this zone is somewhat arbitrary, and to use a more continuous measure of position, I instead quantified avoidance behaviour as the average distances from the centre of the field. Agents were initially placed in a corner of the field. Simulations showed that the intrinsic reward weight was negatively associated with distance from the centre of the field (i.e., reduced thigmotaxis r = -0.23, p < 0.001; Figure 4.6C, G), meaning that the agents with higher exploration drives were more willing to explore the centre of the field. As expected, the extrinsic reward weight was positively associated with distance from centre (i.e., increased thigmotaxis: r = 0.61, p < 0.001; Figure 4.6D, H), in other

words the agents who were more averse to the centre avoided it. The intrinsic reward decay and temporal discount parameters did not significantly correlate with distance from centre (intrinsic reward decay: r = -0.11, p = 0.056; Figure 4.6E, I; temporal discount: r = -0.11, p = 0.051; Figure 4.6F, J).

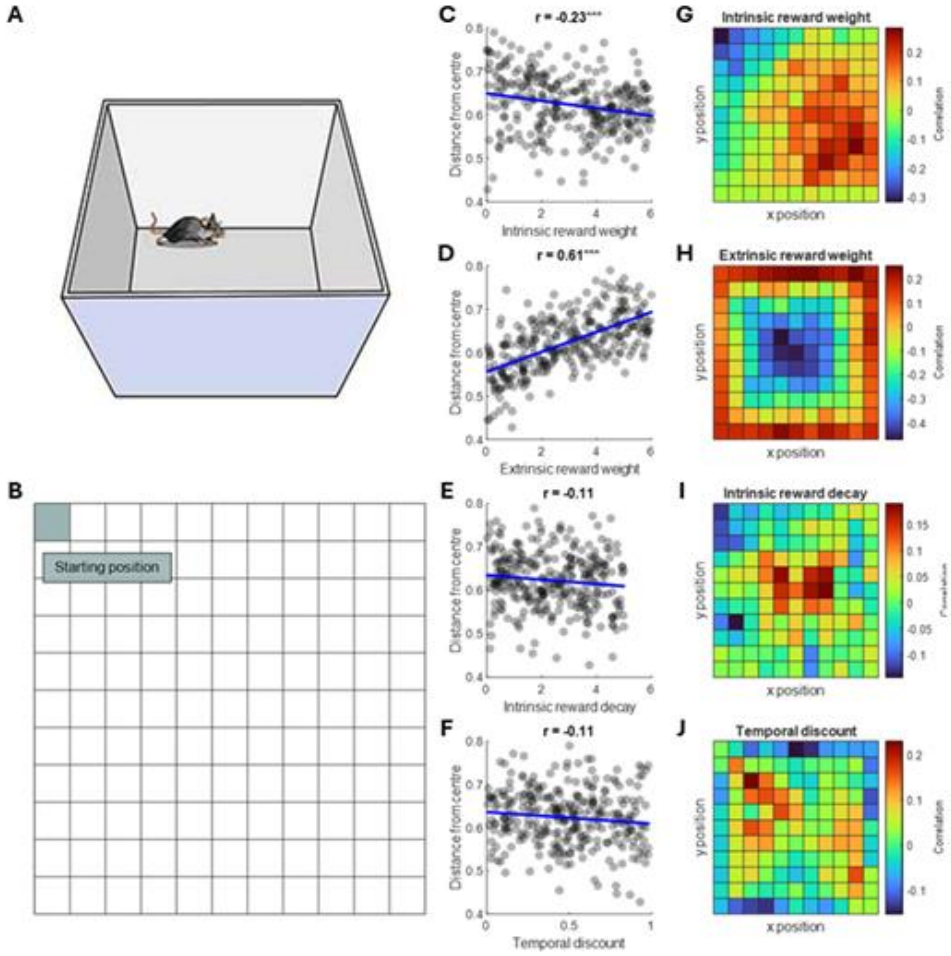


Figure 4.6, Simulating the effects of the model parameters in the open field test. A, *The open field test and the discretised version of the test. In line with convention, agents started from a corner of the field.* ***B,*** *The agents' starting position across all simulations.* ***C-F,*** *The effect of each model parameter on avoidance behaviour, based on simulations of 300 agents with random parameter values. The metric of avoidance was the average distance from the centre of the field (normalised such that centre = 0 and periphery = 1). The intrinsic reward weight reduced avoidance, whist the extrinsic reward weight had the opposite effect. The other two parameters did not significantly affect avoidance. Significance levels shown according to the following: p < 0.05 – *; p < 0.01 – **; p < 0.001 – ***.* ***G-J,*** *The effect of parameter values on position. Higher values of the intrinsic reward weight were associated with more central positions, and the extrinsic reward weight produced the opposite effect. Qualitatively, higher values of the intrinsic reward decay led agents further away from the starting position, as they discounted novelty more rapidly, leading to them seeking new states (irrespective of central/peripheral position).*

The light dark box

Avoidance in this test is measured by two key metrics: 1) the time spent in the light compartment of the box, and 2) the latency to first entry of the light compartment. Agents were initially placed in the corner of the dark compartment, in line with convention. Here, agents with higher intrinsic reward weights spent more time in the light compartment (i.e. reduced scototaxis: $r = 0.28$, $p < 0.001$; Figure 4.7C) and showed quicker initial exploration of the light compartment ($r = -0.30$, $p < 0.001$; Figure 4.7G). Conversely, the extrinsic reward weight showed the opposite effects (time spent in light compartment/increased scototaxis: $r = -0.32$, $p < 0.001$; Figure 4.7D; latency to first entry: $r = 0.21$, $p < 0.001$; Figure 4.7H). The intrinsic reward decay was positively correlated with time in the light compartment ($r = 0.15$, $p = 0.010$; Figure 4.7E) and slower exploration of light compartment exploration ($r = -0.14$, $p = 0.018$; Figure 4.7I), and the temporal discount parameter also correlated with faster exploration of the light compartment ($r = 0.11$, $p = 0.048$; Figure 4.7J).

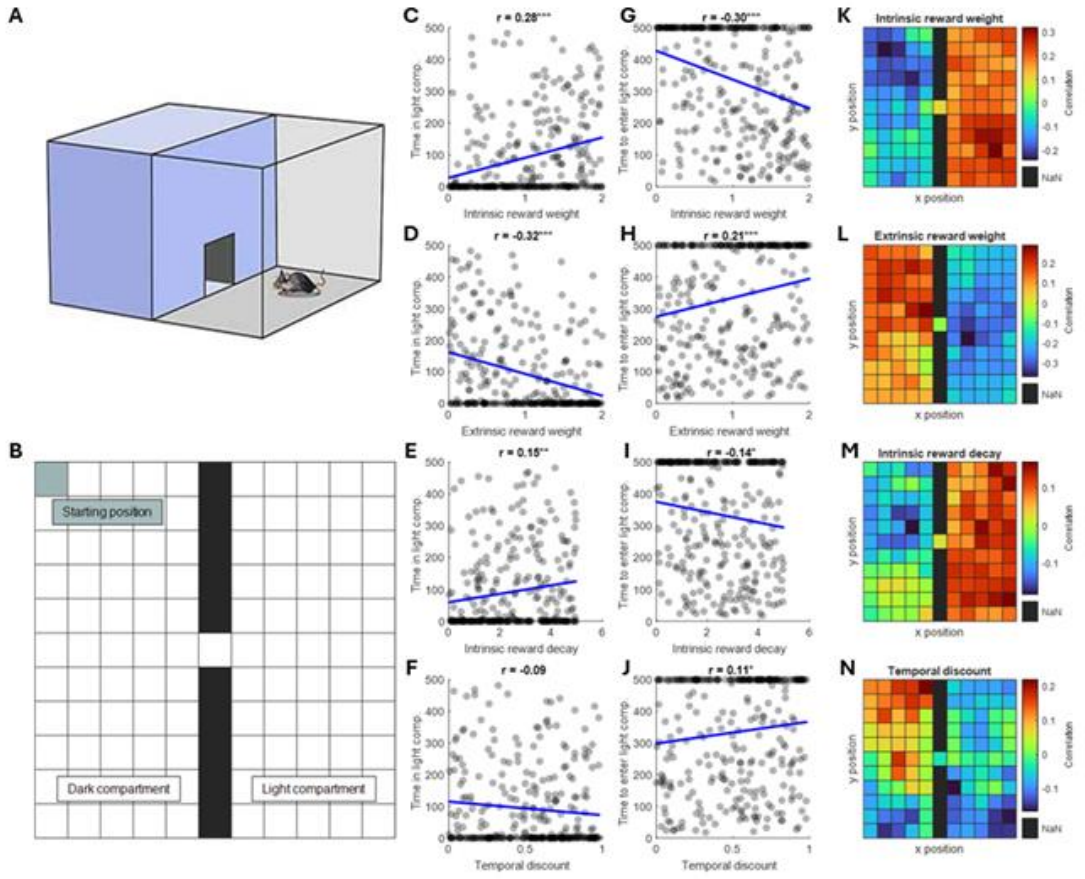


***Figure** 4.**7, Simulating the effects of the model parameters in the light dark box. A,** The light dark box and the discretised version of the test. In line with convention, agents started from a corner in the dark compartment. **B,** The agents' starting position across all simulations. **C-J,** The effects of each model parameter on avoidance behaviour, based on simulations of 300 agents with random parameter values. Significance levels shown according to the following: $p < 0.05$ – *; $p < 0.01$ – **; $p < 0.001$ – ***. One metric of avoidance was the time spent in the light compartment (**C-F**) and the other was the latency of the first visit to the light compartment (**G-J**). The intrinsic reward weight increased time spent in the light compartment and reduced latencies of the first visit, whist the extrinsic reward weight and the temporal discount parameter had the opposite effect. The intrinsic reward decay did not significantly affect avoidance. **K-N,** The effect of parameter values on*

position. Higher values of the intrinsic reward weight were associated occupancy of the light compartment, and the extrinsic reward weight and temporal discount parameter produced the opposite effect.

The elevated plus maze

Similarly to the light-dark box, the key metrics of avoidance in this test are: 1) the time spent in the open arms of the maze, and 2) the latency to first entry of the open arm. Agents were initially placed in the centre of the map, in line with convention. Agents with higher intrinsic reward weights spent more time on the open arms (r = 0.31, p < 0.001; Figure 4.8C) and explored them more quickly (r = -0.20, p < 0.001; Figure 4.8G). Conversely, extrinsic reward weight demonstrated the opposite effect (time spent on open arms: r = -0.44, p < 0.001; Figure 4.8D; latency to first entry: r = 0.51, p < 0.001; Figure 4.8H). The intrinsic reward decay was positively correlated with time on the open arms (r = 0.22, p < 0.001; Figure 4.8E) and faster entry onto the open arms (r = -0.16, p = 0.006; Figure 4.8I).

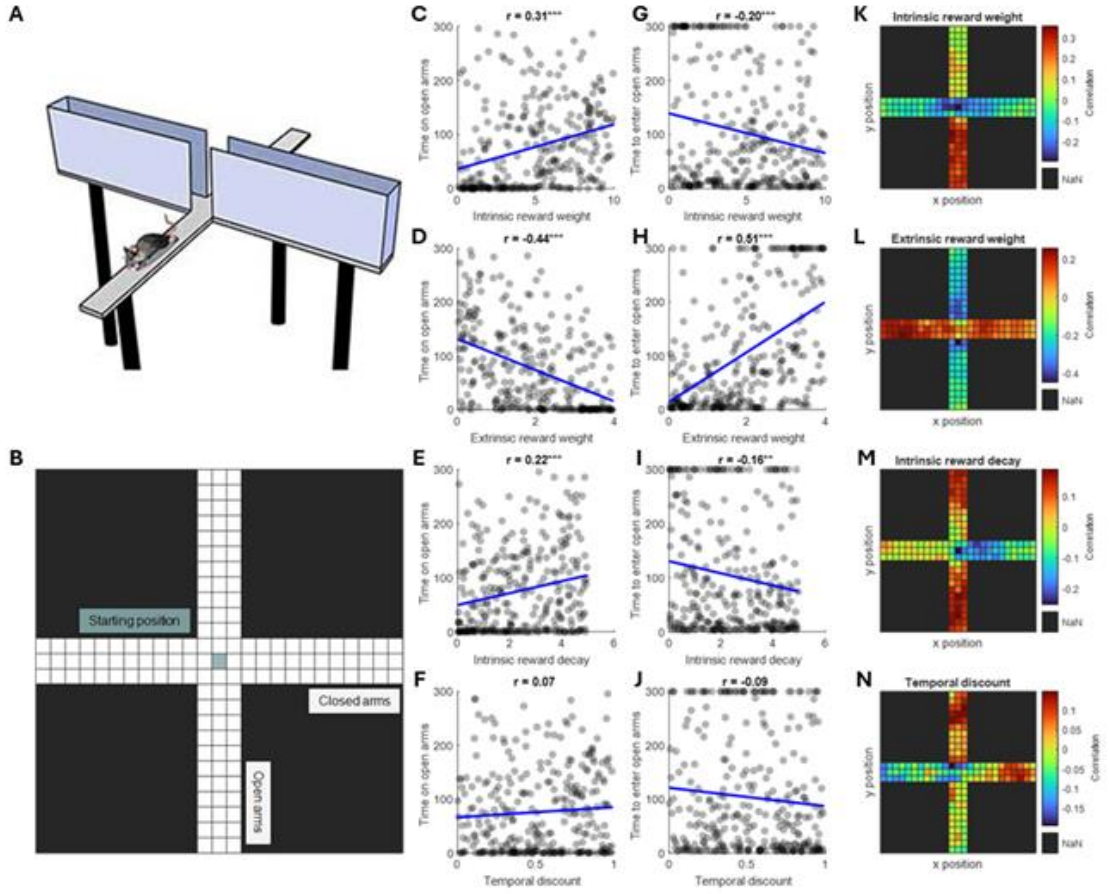


Figure 4.8, Simulating the effects of the model parameters in the elevated plus maze. A, *The elevated plus maze and the discretised version of the test. In line with convention, agents started from a corner in the dark compartment.* ***B,*** *The agents' starting position across all simulations.* ***C-J,*** *The effects of each model parameter on avoidance behaviour, based on simulations of 300 agents with random parameter values. Significance levels shown according to the following: p < 0.05 – *; p < 0.01 – **; p < 0.001 – ***.* ***C-F,*** *One metric of avoidance was the time spent on the open arms. The intrinsic reward weight increased time spent on the open arms whist the extrinsic reward weight had the opposite effect. There was a positive effect of the intrinsic reward decay on time spent on the open arms which approached threshold, but the temporal discount parameter did not affect time spent.* ***G-J,*** *The other metric of avoidance was the latency of the first visit to the open arms. The intrinsic reward weight and the temporal discount parameter decreased latencies whist the extrinsic reward weight had the opposite effect.* ***K-N,*** *The effect of parameter values on position. Qualitatively, higher values of the intrinsic*

reward weight and intrinsic reward decay showed greater occupancy of the open arms, whereas the extrinsic reward weight produced the opposite effect.

The simulations above demonstrate that the proposed model flexibly generates the typical avoidance behaviour from a range of classical exploration paradigms. Greater aversion to threat-associated environmental features induces avoidance, whilst greater motivation to explore reduces avoidance. Notably, in all cases, variance in the traditional metrics of avoidance could be explained by both mechanisms. In the next section, I sought to validate the model against empirical data.

4.6 Results – application of the model to empirical data

4.6.1 Experiment 1: model validation

I fitted ten candidate models to the data from 98 participants in this experiment. The winning model was the 'INT+EXT+TD' model which included three parameters: both intrinsic and extrinsic reward weights and the temporal discount parameter (Figure 4.9A). The distributions of the fitted parameters from this model are shown in Figure 4.9B-D. When I performed parameter recovery analyses of these fitted parameters, I found strong correlations between the fitted and refitted parameters (Pearson's r values for intrinsic reward weight: 0.98, extrinsic reward weight: 0.95, temporal discount parameter: 0.87).

I then checked the winning model's ability to recapitulate the behavioural data by simulating data based on each participants' parameter estimates, as a posterior predictive check. This showed that the model-agnostic metrics based on the empirical data and simulated data were strongly correlated across the classical metrics of the time spent in the open arms ($r = 0.73$, $p < 0.001$; Figure 4.9E), and latency to first entry of the open arm ($r = 0.84$, $p < 0.001$; Figure 4.9F). These findings reassured me that the proposed computational model could indeed simulate human-like data.

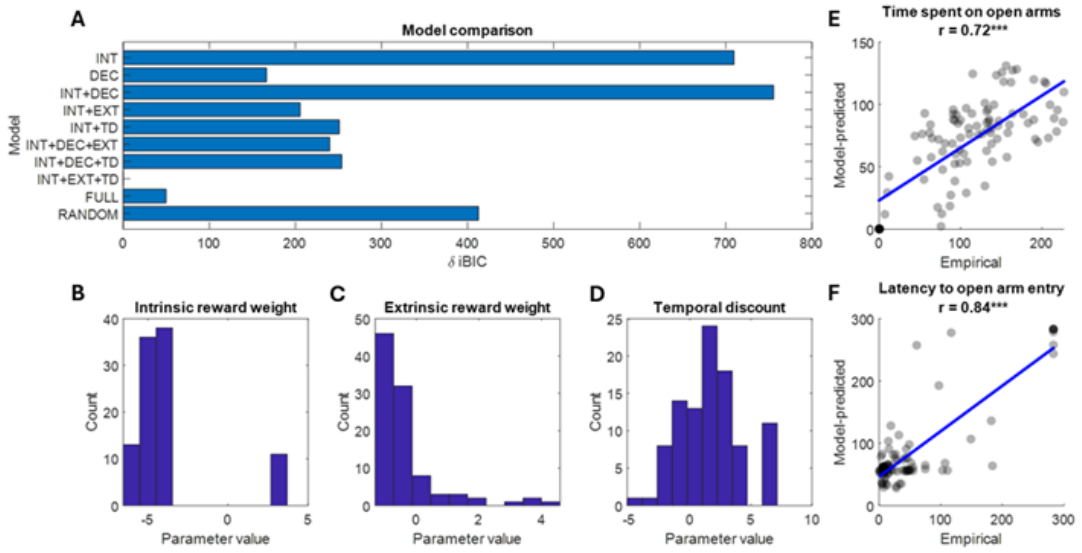


Figure 4.9, Modelling empirical data from the elevated plus maze (Experiment 1 from Biedermann et al., 2017). A, *Model comparison. Bar plots of the difference in the integrated Bayesian information criterion (d IBIC) for each model, compared to the model with the lowest value. The most parsimonious model included three parameters: the temporal discount parameter, and the intrinsic and extrinsic reward weights.* ***B-D,*** *Distributions of the model parameters.* ***E-F,*** *Posterior predictive checks. To assess how well the winning model produced human-like behaviour, I simulated agents who had parameters matching the estimated parameters from the winning model, and compared the similarity of the model-agnostic metrics of avoidance across the empirical and simulated (i.e. model-predicted data). Both classical metrics of avoidance, namely the time spent on the open arms and the latency of the first open arm entry, were significantly and positively correlated, indicating that the model demonstrated good predictive performance. Significance levels shown according to the following: p < 0.05 – *; p < 0.01 – **; p < 0.001 – ***.*

In the original study (Biedermann et al., 2017), the authors found significant associations between self-reported anxiety and several model-agnostic metrics of task behaviour (tested by comparing groups according to median split of self-reported anxiety; time on open arms: $t_{98} = 3.31$, $p = 0.001$, $r = 0.32$; number of open arm entries: $t_{98} = 4.05$, $p < 0.001$, $r = 0.38$; latency of first visit to open arms: $t_{98} = –2.94$, $p = 0.004$, $r = 0.28$; latency of first visit to an extreme end of the open arms: $t_{98} = –4.24$, $p < 0.001$, $r = 0.39$). As noted above, the present dataset did not contain the full anxiety data, but these effects were qualitatively similar in this reduced dataset, albeit with some failing to achieve statistical significance in this restricted sample (tested via Pearson's correlations; time on open arms: $r_{60} = -0.30$, $p = 0.020$; number of open arm entries: $r_{60} = -0.36$, $p = 0.005$; latency of first visit to open arms: $r_{60} = 0.22$, $p = 0.093$; latency of first visit to an extreme end of the open arms: $r_{60} = 0.22$, $p = 0.086$).

When I investigated the associations between the model parameters and self-reported anxiety, I found a moderate negative correlation between anxiety and the intrinsic reward weight (tau = -0.22, $p = 0.019$, $r \approx -0.34$) but the correlation with temporal discount parameter and extrinsic reward weights were non-significant (temporal discount parameter: tau = -0.13, $p = 0.176$, $r \approx -0.20$; extrinsic reward weight: tau = 0.09, $p = 0.344$, $r \approx 0.14$). This suggests that the above associations between anxiety and behaviour are explained by anxiety's effects on the willingness to explore the environment, rather than individual differences in aversion to the open arms. To test this more

explicitly in an exploratory analysis, I investigated whether the significant relationships between anxiety and task performance (specifically, the time spent on the open arms and the number of open arm entries) could be explained individual differences in the intrinsic reward weight. To this end, I conducted mediation analyses, including all three model parameters as parallel mediators of the effect of anxiety on each model-agnostic metric (i.e. in two separate models, one each for time spent on the open arms and the number of open arm entries). This revealed that the intrinsic reward weight was a significant mediator of the relationship between anxiety and time spent on the open arms (indirect effect = -3.79, 95% CI = [-10.14, -1.00], p = 0.003). However, its mediating effect on the number of open arm entries did not reach threshold, although the 95% confidence interval approached the boundary of zero (indirect effect = -0.10, 95% CI = [-0.31, 0.0001], p = 0.084) – however, it should be noted that this analysis is very likely underpowered. These results provide preliminary evidence that individual differences in the willingness to explore may underlie the observed relationships between anxiety and task performance, rather than sensitivities to exposure on the open arms.

4.6.2 Experiment 2: effects of lorazepam and yohimbine on behaviour

I fitted the data from Experiment 2 with the most parsimonious model from Experiment 1 (INT+EXT+TD) to test the putative cognitive mechanisms of lorazepam and yohimbine, relative to placebo. After fitting the data, I first conducted posterior predictive checks to verify the model's ability to recapitulate the model-agnostic effects observed in the original study. These effects were that lorazepam increased time spent on the open arms while yohimbine had the opposite effect, and lorazepam administration also resulted in lower latencies to the first open arm entry (Biedermann et al., 2017). The posterior predictive checks revealed that the model generally replicated the observed patterns in the empirical data, albeit with some consistent misestimation biases in both measures (Figure 4.10A-B). Despite these biases, this analysis indicates that the model faithfully captured the differential effects of drug on behaviour, even with a small set of parameters and relatively low sample sizes.

Next, I examined which, if any, of the model parameters were affected by the drugs (Figure 4.10C-E). I observed a significant effect of drug condition on all three parameters (temporal discount parameter: H_2 = 36.75, p < 0.001; intrinsic reward weight: H_2 = 14.90, p < 0.001; extrinsic reward weight: H_2 = 26.10, p < 0.001). Planned post-hoc test revealed that the difference in the intrinsic reward weights between the lorazepam and placebo groups narrowly missed significance ($U_{19,15}$ = 389, p = 0.052) with median parameter values higher in the lorazepam group (median = 1.52) compared to the placebo group (median = 0.56). Yohimbine, on the other hand, had a significantly negative effect on intrinsic reward weights, compared to placebo ($U_{19,15}$ = 163, p = 0.004). With

respect to the extrinsic reward weights, there was a significant and negative effect of lorazepam on these parameter values, compared to placebo ($U_{19,15}$ = 214, p < 0.001), whereas the effect of yohimbine compared to placebo did not achieve significance ($U_{19,15}$ = 277, p = 0.068). Finally, lorazepam increased the temporal discount parameter (i.e. decreased temporal discounting; $U_{19,15}$ = 475, p < 0.001), whereas yohimbine had the opposite effect (i.e. increased temporal discounting: $U_{19,15}$ = 170, p = 0.010). Overall, these analyses suggest distinct computational effects of lorazepam and yohimbine on the task.

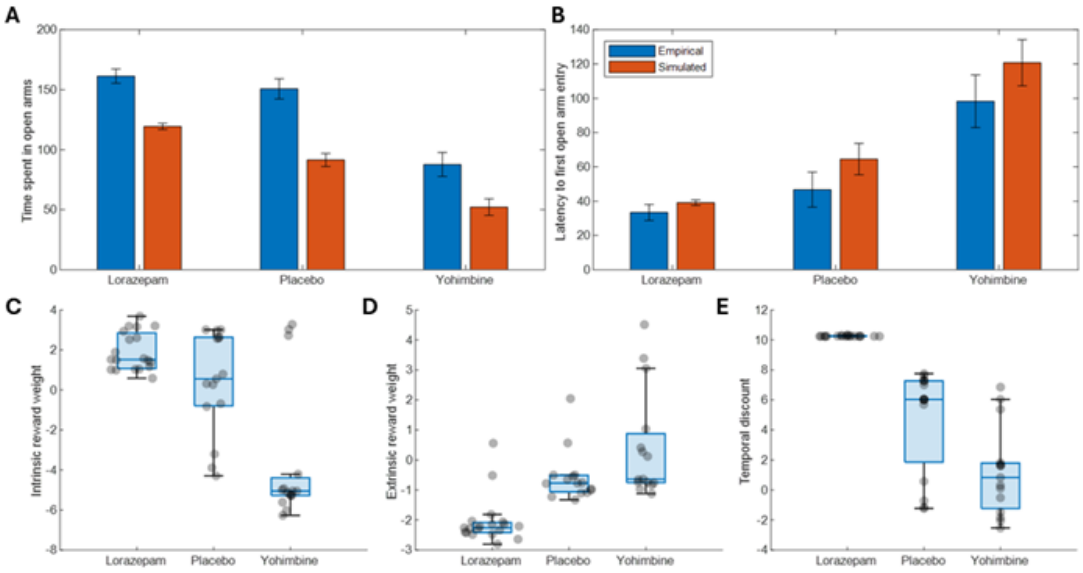


Figure 4.10, Computational analyses of the effects of lorazepam and yohimbine on behaviour in the elevated plus maze (Experiment 2 from Biedermann et al., 2017). A-B, *Posterior predictive checks. To assess how well the winning model from Experiment 1 could predict behaviour in this experiment's data, I fitted the model to the data and simulated agents from the estimated parameters. The simulated data (orange) showed qualitatively similar patterns of data to the empirical data (blue), although there were consistent differences in values.* ***C-E,*** *Parameter distributions across drug conditions.*

4.7 Discussion

In this chapter, I introduced a computational model for exploration in aversive environments. The model offers a simple but effective account of exploratory behaviour and is applicable to a range of existing naturalistic tests of anxiety-like behaviour. When fitted to empirical human data, I found that the model could recapitulate individual differences in behaviour, and when the model was applied to better understand the computational mechanisms of pharmacological interventions in the task, I found evidence that lorazepam and yohimbine exerted their effects on behaviour through partially distinct cognitive mechanisms – an insight that would not be possible without a computational modelling analysis.

4.7.1 Implications

The results from the between-subjects pharmacological study highlight the strength of the computational approach applied to this behavioural paradigm. The model-agnostic analyses from the original human elevated plus maze study showed that lorazepam increased the time participants spent on the anxiogenic open arms compared to placebo, whilst yohimbine had the opposite effect (Biedermann et al., 2017). However, as my simulations demonstrate, there are two conceptually (i.e. computationally) independent mechanisms that can influence the time spent on the open arms. Through the computational analyses, I found that lorazepam and yohimbine exerted their effects on behaviour through partially distinct mechanisms: lorazepam decreased participants' sensitivities to the aversive arms of the elevated plus maze, whilst yohimbine decreased participants' intrinsic motivation to explore the platform. Further these drugs had opposing effects on temporal discounting, with lorazepam showing a decreasing effect whereas yohimbine had an increasing effect on temporal discounting.

This ability to tease apart competing mechanisms has some important implications for the use of these tasks in the drug development pipeline. First, by identifying how a drug acts at the mechanistic level rather than solely focusing on descriptive behavioural outcomes, we can potentially tailor treatments more precisely. For instance, drugs could be recommended based on their effects on disease-relevant mechanisms, allowing for individualised treatment approaches. Second, in the context of non-human animal studies, the computational approach allows for improved interpretability of behaviour as self-reported behavioural motives are inaccessible. Indeed, increased locomotor activity in rodents can be interpreted as either high anxiety behaviour, as a result of slower habituation, or low anxiety behaviour, reflecting a greater propensity for exploration (Neville et al., 2023). Such ambiguity arises due to the problem of equifinality, where different underlying cognitive mechanisms can produce the same observable summary-statistics of behaviour. Better interpretability of behavioural data is particularly relevant in the context of drug development pipelines, where such measures are employed in preclinical trials, as a computational understanding of drug effects could help to better assess if a candidate drug affects disease-relevant mechanisms over irrelevant ones – in other words, this could enhance predictions about whether its effects will translate clinically.

4.7.2 Biological plausibility

Regarding the neural plausibility of the model, there are several aspects that align with theories of spatial representation in the brain. For instance, the hippocampus has been extensively implicated in encoding spatial maps (Morgan et al., 2011; O'Keefe & Nadel, 1978), which can depend on environmental features including aversive stimulus locations (Calvin et al., 2024; Hayman et al., 2003;

Nikbakht et al., 2024; OKeefe & Burgess, 1996), thus providing a neural basis for representing spatial maps of rewards and punishments. Notably, a recent study (Calvin et al., 2024) observed that hippocampal representations of threat locations during approach-avoidance conflict were abolished by diazepam (a benzodiazepine of the same family as lorazepam). This finding is conceptually similar to my finding that lorazepam reduced the extrinsic reward weight - both suggesting a common anxiolytic mechanism of attenuating the encoding of threat.

The hippocampus is also thought to encode past trajectories in space through a process known as replay, where stored representations are reactivated, even in the absence of specific sensory inputs (Karlsson & Frank, 2009). This stored history of spatial positions gives a potential mechanism of how organisms can track state visit counts which can be used to direct exploration. Furthermore, contemporary theories propose that the hippocampus serves a function akin to the successor representation (Ekman et al., 2023; Stachenfeld et al., 2017). This aligns with my model's approach of efficiently managing changes in reward information while maintaining stable transition dynamics in the environment.

The convergence of these neuroscientific principles with my computational framework supports its plausibility as an approximation of the underlying neural processes involved in exploration in aversive environments. While further research is needed to fully bridge the gap between this computational model and its neural mechanisms, the current evidence suggests that the model's components, from spatial representation to threat processing and exploration strategies, find support in the current understanding of hippocampal function and its role in spatial navigation and memory.

4.7.3 Limitations/opportunities for development

For the sake of simplicity, I did not model stochasticity in the spatial environments, as is traditional in MDPs; this could be relevant to anxiety, as previous theoretical work suggests that the level of an agent’s pessimism about their future spatial trajectory can lead some to exhibit anxiety- or depressive-like behaviour (Zorowitz et al., 2020). Pessimism about state transitions may also provide another explanation for avoidance of the open arms of the elevated plus maze, as agents may be pessimistic about their ability not to fall off the platform. Future experiments that induce some stochasticity in movement will be an interesting avenue for future research. Of note, the ‘closed arms’ in the original elevated plus maze study by Biedermann et al. (2017) were not truly closed in the sense that one could still fall off the platform from these positions. This potentially introduces an aversive element to these arms, in addition to the open arms, if one assumes that the risk of falling induces avoidance. Indeed, it is likely that both mechanisms (i.e. aversion to exposure and to falling) play a role in guiding exploration in the task.

I also opted not to implement a fully model-based algorithm to explain exploratory behaviour. Instead, I employed a successor-representation-like algorithm (Dayan, 1993), which efficiently handles changes in reward information but not in the spatial structure of the environment (Russek et al., 2017). This suited my modelling approach as the behavioural paradigms of interest in this chapter maintain stable spatial maps, while the intrinsic reward signal in the model evolves continuously over time. However, compared to a purely model-based algorithm which would plan the optimal sequence of actions from any given state to maximise reward, the algorithm approximates this process by assuming that agents choose actions based on nearby rewards without explicitly computing an action sequence at any given time. Future research will be necessary to compare the explanatory power of these different approaches. Nevertheless, the strong correlations observed between the empirical and model-predicted data in posterior predictive checks suggest that this approximation is a sufficiently accurate representation of human exploration behaviour.

An interesting shortcoming of the model was its tendency to overestimate avoidance behaviour in the drug experiment. As shown in Figure 4.10A and B, the model consistently predicted less time spent on the open arms and longer latencies to first entry compared to the empirical data. This systematic bias suggests that the current model formulation may be overemphasizing the aversiveness of the open arms. Though there are many potential explanations for this this discrepancy, an important one is that the model may not adequately capture the effects of habituation, whereby the perceived threat of the open arms could diminish over time, even if it has not been explored yet. Future work could incorporate time-dependent parameters to model habituation effects, or indeed other additional parameters to capture other important cognitive aspects of behaviour. Addressing these factors will help to improve the model's absolute predictive accuracy.

4.7.4 Naturalistic computational psychiatry

More broadly, the work in this chapter is an example of the potential of ‘naturalistic computational psychiatry’, which consistent with a general trend in the cognitive sciences towards measures with better ecological validity, such as through using more ecological stimuli (e.g. movies; Kirk et al., 2023), or by enriching the environment that a participant/subject can interact with (Bach et al., 2014; Wise et al., 2023). Traditionally, cognitive research has relied on more simple and abstract measures of decision-making such as two arm bandit tasks, to study a narrow selection of cognitive processes at a very precise level. The problem with naturalistic measures is that they recruit a diverse array of cognitive processes, many of which may not be directly relevant to the research question. However, as my findings show, the theory-driven computational approach is well-suited to disentangling multiple candidate mechanisms of behaviour, making it well-placed to tackle the problem of

multidimensionality in naturalistic data - there is already evidence that studying behaviour in more naturalistic environments reveals novel cognitive and neuroscientific insights that cannot be observed in traditional measures (Scholl & Klein-Flügge, 2018). Going forward, more work is needed that embraces this multidimensionality, as patients do not present with simple problems isolated to abstract concepts; rather, they navigate complex issues within a highly multidimensional world.

4.8 Conclusion

With a relatively simple model, I explain some classical approach-avoidance conflict effects during spontaneous exploration in aversive spatial environments. The model was also supported by empirical human data, and was able to provide novel mechanistic interpretations of the effects of lorazepam and yohimbine on exploration behaviour. This work broadens the scope for more naturalistic measures that can be analysed within a computational framework, which may improve the ecological validity of computational psychiatry research going forward.

5 General Discussion

5.1 Overview

This thesis has been an exploration of translational models of anxiety; how behaviour in forward-translated tests relates to anxiety in humans, what they measure in cognitive-computational terms, and how the mechanisms they engage respond to changes in anxiety. I focused on two of the most popular classes of translational anxiety models: conditioned tests, in which approach and avoidance responses need to be learned and balanced simultaneously; and unconditioned tests, in which an organism's natural drive to explore novel environments conflicts with the avoidance of aversive environmental features.

In this final chapter of the thesis, I will summarise the key findings from each experimental chapter, and then discuss their broader implications.

5.2 Summary of findings

5.2.1 Chapter 2: developing and validating a human conditioned test of anxiety

I began by developing a human conditioned test of anxiety, specifically by forward-translating the task structure of the Vogel and Geller-Seifter conflict tests (Geller et al., 1962; Vogel et al., 1971). These tests create a conflict between reward-pursuit and punishment-avoidance by offering rodents a choice between two actions, which are asymmetrically associated with reward and punishment. The 'approach' action (which can be a lever-press or a lick at a waterspout) can lead to a reward (food/water) but also an electric shock. Withholding these actions represents 'avoidance' and results in neither reward nor punishment. I translated this asymmetric structure into a two-armed bandit task that was designed to optimise computational psychometric properties. In the task, human participants chose between a 'conflict' option associated with relative high reward probability but also some chance of producing an aversive auditory stimulus (i.e. the 'approach' action), or a 'non-conflict' option which had lower reward probability but no chance of producing the aversive stimuli (i.e. the 'avoidance' action). In two large and independent online samples with participants drawn from the general (online) population, I found that participants who reported feeling more anxious during the task were more inclined to choose the non-conflict option over the conflict option, demonstrating more avoidance compared to less anxious participants. When I analysed the choice data using model-free reinforcement learning algorithms, I found that this effect was explained by asymmetries in individual sensitivities to reward and punishment – specifically, participants experiencing greater -task-induced anxiety had higher sensitivity to punishment, relative to reward, and further, this association appeared to explain the relationship between anxiety and avoidance behaviour. These findings validate the relevance of tests like the Vogel and Geller-Seifter conflict tests

to human anxiety, and for the first time, provide a mechanistic perspective on how anxiety and avoidance are related in these tests. Specifically, the results indicate that anxiety is associated with a computational bias towards higher sensitivity to punishments relative to rewards, which in turn drives avoidance behaviour in approach-avoidance conflict situations.

5.2.2 Chapter 3: computational mechanisms of avoidance which are sensitive to anxiety manipulations

My findings from Chapter 2 hinted at a potential computational mechanism – relative sensitivity to rewards versus punishments - of how anxiety leads to avoidance behaviour, but this was inferred on the basis of a cross-sectional and correlational study design. In order to test this hypothesis more robustly, in Chapter 3 I used experimental manipulations of anxiety. Specifically, I used an acute dose of lorazepam to test the effect of reducing anxiety during task performance on behaviour and its underlying computational mechanisms, and I used threat of unpredictable electric shocks to test the effect of heightened anxiety on these measures. Based on the findings from Chapter 2, I predicted that decreasing anxiety would decrease avoidance behaviour and shift sensitivities towards rewards over punishments, whereas increasing anxiety would increase avoidance behaviour and shift sensitivities towards punishments over rewards. The results showed that threat-of-shock induced a state of heightened anxiety as expected and it also led to an increase in avoidance behaviour, but there was no evidence for an effect of lorazepam on either anxiety or on descriptive model-agnostic metrics of avoidance. However, the computational analyses showed that each of the anxiety manipulations each the expected effects on the reward-punishment sensitivity index: lorazepam shifted sensitivities towards rewards whereas threat-of-shock shifted sensitivities towards punishments. These findings suggest that this task engages computational mechanisms that respond in expected directions when subjected to anxiety manipulation – in other words, the results provide evidence for the computational validity of this human task with respect to anxiety.

5.2.3 Chapter 4: a computational model of spontaneous exploration in aversive spatial environments

In Chapter 4, I turned my focus to unconditioned tests of anxiety such as the elevated plus maze. As a generative-computational approach had not previously been applied to the behaviour in these tests, I first proposed a model of spontaneous spatial exploration based on a concept of directed exploration. Simulations from this model consistently explained the classical findings from the open field test, the light-dark box and the elevated plus maze through relatively simple computational mechanisms. Having found potential for the model's explanatory power through these simulations, I used data from a previous study implementing a mixed-reality human version of the elevated plus

maze (Biedermann et al., 2017) to validate the model empirically. The results from these analyses showed that the model could recapitulate the behavioural effects in the data, and the model parameters provided novel insights into the mechanisms of the effects of the anxiolytic, lorazepam, and the anxiogenic, yohimbine, on task performance. Specifically, the model-based analyses suggested that lorazepam decreased open arm avoidance by reducing individual sensitivities to open arm exposure, and also reduced temporal discounting. On the other hand, yohimbine increased avoidance by reducing individual drives to explore the environment. These analyses demonstrated the potential of the computational approach in deriving more precise, mechanistic inferences from unconditioned tests of anxiety that involve spatial exploration in aversive environments.

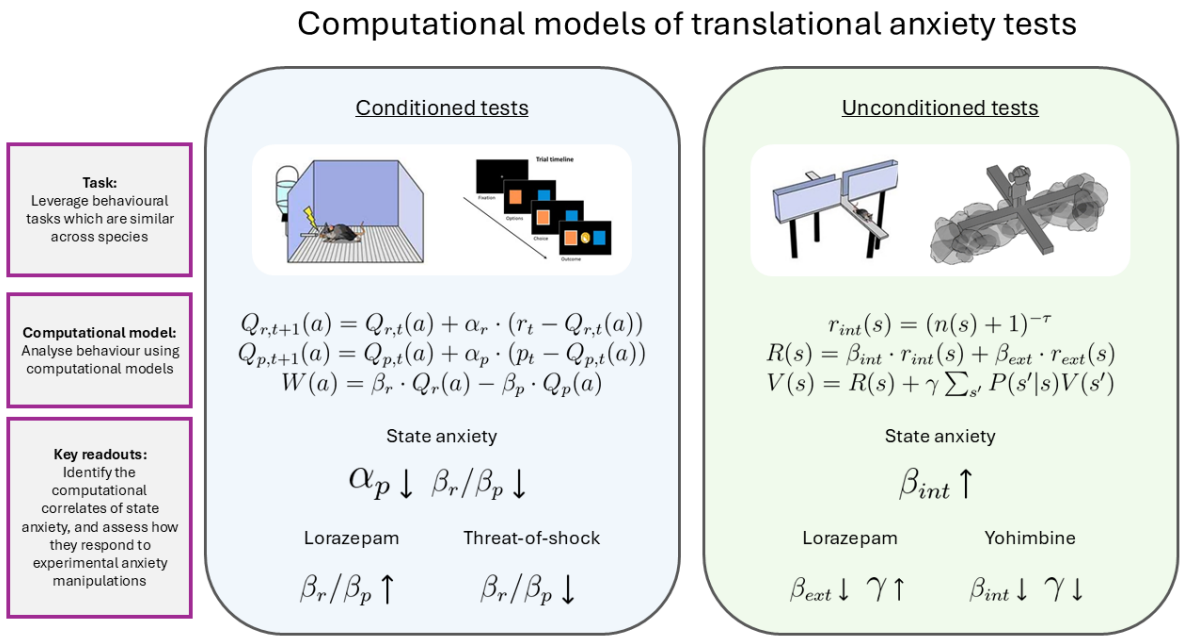


Figure 5.1, Summary of computational models for translational anxiety tests. *This figure presents an overview of the key findings from conditioned and unconditioned tests of anxiety explored in this thesis. For each type of test, it shows the behavioural task, the computational model used to analyse behaviour, and key readouts including how computational parameters relate to state anxiety and respond to anxiolytic (lorazepam) and anxiogenic (threat-of-shock, yohimbine) manipulations. In conditioned tests, state anxiety and threat-of-shock were associated with decreased reward-punishment sensitivity ratios (βr/βp), while lorazepam increased this ratio. In unconditioned tests, state anxiety was associated with decreased intrinsic reward weight (βint). Lorazepam decreased extrinsic reward weight (βext) and increased temporal discounting (γ), while yohimbine decreased βint and γ. These findings demonstrate how computational modelling can provide mechanistic insights into anxiety-related behaviours across different experimental paradigms.*

5.3 Implications: returning to the outstanding questions

In Chapter 1, the General Introduction, I laid out three key questions from the literature preceding the work from this thesis. I will discuss the implications of my results for each of these questions in turn.

5.3.1 Are classical preclinical anxiety tests relevant to human anxiety?

As I noted in the General Introduction, the classical preclinical tests of anxiety were developed as surrogates of broad concepts in anxiety, notably of approach-avoidance conflict. These tests have seen surprisingly little empirical validation with respect to their relevance to human anxiety, despite their massive popularity in the non-human animal anxiety literature (Griebel & Holmes, 2013). It was only relatively recently that a study took an explicitly translational approach in testing a forward-translated version of the conditioned anxiety tests to validate its association with anxiety in humans (Aupperle et al., 2011). In this study, avoidance behaviour was significantly correlated with anxiety experienced during the task, albeit only in male participants and not in females, but behaviour in the task did not correlate with clinical measures of anxiety (Aupperle et al., 2015; Aupperle et al., 2011) such as the State-Trait Anxiety Inventory (Spielberger et al., 1971) or the Behavioural Inhibition Scale (Carver & White, 1994). This suggests that this specific behavioural paradigm may measure processes that are more relevant to state anxiety rather than clinical symptoms of anxiety or anxious traits.

The results from Chapters 2 and 3 are consistent with this notion. I used a different design to that developed by Aupperle et al., by also adding a learning component to the task to better approximate the 'conditioned' aspect of the non-human animal tests. However, the results relating to state and clinical/trait anxiety were broadly similar. I indeed found replicable associations between avoidance behaviour and self-reported state anxiety during approach-avoidance reinforcement learning. However, I found little, or at best, weak, evidence for the task's sensitivity to clinical symptoms of anxiety. In the discovery sample in Chapter 2, I found that avoidance behaviour in the task was significantly correlated with self-reported scores of experiential avoidance, a concept related theoretically to clinical anxiety, but this effect did not replicate in the independent replication sample, and there were no associations based on the Generalised Anxiety Disorder scale (Spitzer et al., 2006). In Chapter 3, exploratory analyses provided preliminary evidence that the effect of lorazepam on behaviour may depend on behavioural inhibition traits, but this requires further replication/validation. Therefore, based on the present results, it seems that the behaviours engaged in human conditioned tests of anxiety are primarily relevant to state anxiety. One important caveat is that the participant samples in both Chapters 2 and 3 did not include many individuals reporting high levels of clinical symptoms, which may preclude the assessment of this task's clinical relevance. I will discuss this further below.

On the other hand, human behaviour in unconditioned anxiety tests does seem to show sensitivity to clinically relevant symptoms. The original study on the human version of the elevated plus maze (Biedermann et al., 2017), whose data I re-analysed using a computational approach, found that avoidance behaviour was associated with both state anxiety and clinical symptoms of acrophobia

(fear of heights), although it was not significantly correlated with a more general measure of clinical anxiety, the State-Trait Anxiety Inventory. Similarly, Walz et al. (2016) demonstrated that behaviour in a human open field test was related to agoraphobia symptoms. Specifically, patients with agoraphobia and participants with high anxiety sensitivity (a risk factor for panic disorder and agoraphobia) exhibited increased thigmotaxis - a tendency to stay close to walls - compared to healthy controls and those with low anxiety sensitivity. This thigmotactic behaviour correlated with clinical measures of agoraphobic fear. Together, these findings suggest that unconditioned tests may be more ecologically valid proxies of real-world avoidance compared to conditioned tests. The naturalistic settings of these paradigms - an elevated maze or an open field - may better approximate specific anxiety-provoking situations encountered in daily life by individuals with clinical anxiety.

It is noteworthy that in both studies, the behavioural measures were particularly sensitive to specific subtypes of anxiety - acrophobia in the elevated plus maze and agoraphobia in the open field test. This specificity suggests that these paradigms are useful for studying avoidance to specific objects of anxiety, or specific phobias, but may be less useful for disorder covering a broader array of anxiety such as generalised anxiety disorder.

5.3.2 What constructs do these tests measure?

In Chapter 1, I explored a brief history of the Diagnostic and Statistical Manual of Mental Disorders (DSM; American Psychiatric Association, 2013b) and its attempts to classify clinical anxiety as a group of categorical disorders. This theoretical framework inspired the development of behavioural tests that were aimed at reproducing aspects of these disorder categories in non-human animals. At least in the case of tests proposing to measure anxiety, it has been commonly observed that benzodiazepines (which are robustly anxiolytic in humans) reduce ‘anxiety-like’ behaviour in non-human animals in these tests, supporting their role as models of human anxiety disorders. However, more recent research has called into question the validity of these categories, which leaves us with the question of what exactly they measure, if not ‘anxiety’ as per the DSM.

A better way to describe these tests could be to state the neurobiological processes they engage, and this could include descriptions at the molecular, circuit- and cognitive levels. This is increasingly the approach taken in contemporary perspectives on translational anxiety models. However, it is not without its limitations. Cognitive descriptions, while useful, can often lack the specificity needed for precise characterisation of these tests. They may use broad terms like "decision-making" or "approach-avoidance conflict" that do not fully capture the nuanced differences between various anxiety tests. This lack of precision in cognitive descriptions can hinder our ability to draw clear

parallels between animal models and human anxiety, potentially limiting the translational value of these tests.

In line with recent trends (Redish et al., 2022), I argue that an even better, more precise way to describe what these tests measure is to frame them as information-processing problems, consistent with the key assumption of computational neuroscience that the brain is an information-processing organ. Laying out in computational terms the processes engaged in these tasks can serve as precise statements that allow us to easily dissociate between and draw similarities across different tests.

For example, the computational models of behaviour in conditioned tests of anxiety propose that these tests involve a mechanism of error-dependent learning to update the expected values of actions:

$Q_t^o(a) = Q_t^o(a) + \alpha \cdot [o^t - Q_t^o(a)]$ [The Rescorla-Wagner learning mechanism]

The learning rates (α) are free parameters in these models, indicating that the rate of updating expected values is one axis that individual differences can manifest in behaviour.

On the other hand, the open field test, light-dark box and elevated plus maze involve elements of value-based planning across space and time and spontaneous exploration, which are captured by the computations that were supported by the modelling analyses:

$V(s) = R(s) + \gamma \sum_{s'} P(s'|s)V(s')$ [The Bellman equation]

$r_{int}(s) = \frac{1}{(n(s)+1)^{\tau}}$ [Intrinsic exploration rewards via state visit counts]

Again, the free parameters of the temporal discount parameter (γ) and the intrinsic reward decay (τ) indicate the scope for individual differences.

Comparing the equations that make up the computational model for each task/test helps to clearly define their differences. One obvious difference between the conditioned and unconditioned tests is that the former involve learning (i.e. conditioning), and this is evident through the learning equations in the conditioned test models, and their absence in the models for the unconditioned tests. However, it is less obvious that unconditioned tests involve a count-based spatial exploration strategy, which the unconditioned tests do not – this is made clear by the differences in the respective models.

At the same time, the equations also indicate the similarities of certain tests beyond the simple face-level. Perhaps it is quite obvious that the Vogel and Geller-Seifter engage similar cognitive-computational mechanisms, but the commonality of the Rescorla-Wagner rule and its theoretical

neural underpinnings (Roesch et al., 2012) raises the possibility that both tasks depend on dopaminergic signalling as a central mechanism that ties the tasks together, even if the behavioural responses (engage water spout /lever-press) and the reward outcome (water/food) differ. Similarly, the simulations from Chapter 4 demonstrated that common mechanisms could account for the behaviours across the open field test, light-dark box and elevated plus maze, which might also imply a common neural mechanism, potentially via the hippocampus (O'Keefe & Nadel, 1978; Stachenfeld et al., 2017).

Finally, there is one common equation to the conditioned and unconditioned models:

$$W = \beta_R R + \beta_P P \qquad \text{[Integrated action values]}$$

This equation makes clear that both tasks involve value-based decision-making, between things we want to approach and things we want to avoid; in other words, this shows that both induce approach-avoidance conflict.

Returning to the question of what exactly these behavioural tests measure, the answer, at least at the cognitive-computational level, may be most precisely stated via the equations of the computational model themselves. We have just seen that these constructs specifically relate to learning to predict outcomes from experience, planning over time and space, exploration and decision-making. But above this benefit of precision, this method of description has another important advantage – it circumvents any problems with anthropomorphising the behaviour of non-human animals through words like 'worried' or 'conflicted'. Such terms can be ambiguous and ascribe unlikely capacities to non-human animals. Can rats or mice feel *worried* about an upcoming shock or potential predator, in the same way that humans do? Given the language barrier across species, it is impossible to tell. On the other hand, information processing mechanisms are arguably more objectively comparable as they can be measured directly from behaviour. At the same time, while these abstract computational descriptions allow us to sidestep the issue of anthropomorphism, they still require validation in terms of their relevance to human anxiety. As such, the next crucial step was to examine how these constructs, identified via computational modelling, relate to and respond to changes in human anxiety.

5.3.3 How do these constructs respond to changes in anxiety?

The computational constructs described above all potentially affect the read-out of behaviour (i.e., approach or avoidance) in their own unique ways, but not all of them may be relevant to anxiety. I used computational modelling as a tool to disentangle these processes, and potentially identify anxiety-relevant mechanisms from other cognitive processes that may influence behaviour. I sought

to do this by examining how the computational mechanisms in each task change in response to anxiety, either through correlations across anxiety levels or via experimental manipulations.

The findings of Chapters 2 and 3 consistently pointed to the reward-punishment sensitivity index as a key computational mechanism in anxiety-related approach-avoidance behaviour. In Chapter 2, task-induced anxiety was negatively correlated with both punishment learning rate and the reward-punishment sensitivity index. However, mediation analyses revealed that only the latter significantly mediated anxiety's effect on avoidance behaviour. This suggests that anxiety primarily influences behaviour by influencing the relative weighting of rewards versus punishments during decision-making, rather than by influencing learning, per se. The experimental manipulations in Chapter 3 provide further support for this interpretation. Both anxiolytic (lorazepam) and anxiogenic (threat-of-shock) interventions significantly affected the reward-punishment sensitivity index in predicted directions: lorazepam increased sensitivity to rewards relative to punishments, while threat-of-shock had the opposite effect. Crucially, mediation analyses again suggested that these shifts in the reward-punishment sensitivity index explained the effects of the interventions on behaviour. Interestingly, although threat-of-shock influenced multiple computational mechanisms, including learning rates, mediation analyses consistently identified the relative reward-punishment sensitivity index as the primary drivers of behavioural change. Taken together, these findings provide compelling evidence that the reward-punishment sensitivity index serves as a key computational mechanism through which anxiety modulates approach-avoidance behaviour.

In Chapter 4, I found that self-reported anxiety was negatively correlated with an intrinsic reward weight parameter, which represents an individual's motivation to explore novel states. Mediation analyses suggested that this parameter explained the relationship between anxiety and reduced time spent in open arms of the elevated plus maze. Interestingly, the analogous effects with respect to the extrinsic reward weight parameter were not significant. This finding goes against the conceptualisation of the elevated plus maze as a test that induces avoidance due to aversion to exposed regions of the environment. However, it should be noted that in the human version of the test, the closed arms were not fully closed as they traditionally are in rodent set-ups of the test, which may have biased the results by reducing the expected difference in aversion to the open vs closed arms. Another interpretation is that the reduced willingness to explore in more anxious humans could also have reflected ‘freezing’– a universal fear response (Roelofs, 2017), which obscured any biases in preferences for the open vs closed arm. But given the lack of baseline movement data, it is difficult to tease apart the potential conflicts between freezing and willingness to explore. However, lorazepam decreased the extrinsic reward parameter (i.e. aversion to the open arms), perhaps providing some evidence of the relationship between anxiety and open arm aversion,

at least in the mixed-reality instantiation of the elevated plus maze. Further work will be necessary to test the robustness of this relationship.

The other drug manipulation, yohimbine, an anxiogenic, had uniquely non-overlapping effects on the computational mechanisms compared to lorazepam. Unlike lorazepam, it had no effect on open arm sensitivity or temporal discounting, but it did significantly reduce the intrinsic reward weight, in other words, the willingness to explore. This indicates that different drugs impact anxiety and behaviour via distinct mechanisms, presumably as a result of distinct pharmacological pathways, and highlights the strength of modelling in teasing apart competing mechanisms explaining behaviour that might be missed by traditional behavioural analyses alone.

Taken together, these results support the often-cited claim that computational modelling is superior to traditional behavioural analyses, by allowing for a more mechanistic understanding of behaviour and how interventions impact behaviour. In the context of translational models of anxiety, the process of identifying specific parameters that respond consistently to changes in anxiety and mediate its effects on behaviour marks an important step towards cross-species translational models that better capture the relevant aspects of clinical anxiety. These findings provide initial indications that the computational measures from these tasks could serve as biomarkers of anxiety, which can be elicited analogously in humans and animals.

5.4 Broader implications

5.4.1 Computational theories of anxiety

My findings contribute to a broader literature on the computational understanding of anxiety. With respect to model-free reinforcement learning, one common finding in the literature has been that learning rates, very often punishment-specific learning rates, are an important mechanism in the context of anxiety symptoms. For example, whilst healthy control participants dynamically adjust their rate of learning in accordance with the temporal predictability of the environment (Behrens et al., 2007), individuals with high trait anxiety show less of this flexibility, especially for punishing outcomes (Browning et al., 2015; Huang et al., 2017). Further, a meta-analysis of studies comparing patients with anxiety and/or depression to controls on model-free reinforcement learning parameters found that the symptomatic group showed elevated punishment learning rates (Pike & Robinson, 2022). These consistent findings point to a specific role of learning mechanisms.

However, my results from Chapters 2 and 3 did not find this to be the case, at least in the context of state anxiety. Punishment learning rates were indeed correlated with state anxiety in Chapter 2, but the mediation analyses from both Chapters 2 and 3 suggested that individual differences in valuation of reward vs punishment (i.e. the reward-punishment sensitivity index), which represents more of a

decision-making mechanism compared to a learning one, were a stronger predictor of the relationships between anxiety and avoidance. This difference is perhaps explained by task design, as computational parameter values can be affected by experimental context (Eckstein et al., 2022), and my task was designed specifically to study approach-avoidance conflict. Prior tasks in the literature typically involve symmetrical reinforcement structures, where learning about rewards and punishments occurs independently. In contrast, my task used asymmetrical options, where rewards and punishments were in tension across options. This design may have shifted the emphasis from learning processes to decision-making processes, as participants had to constantly weigh the relative impact of rewards and punishments when making choices. Consequently, this approach-avoidance conflict setup might be more sensitive to individual differences in reward-punishment sensitivity rather than learning rates.

Further, the computational model of spatial exploration from Chapter 4 is a novel contribution to computational theories of anxiety, and it demonstrates how computational modelling approaches can be extended to more complex behaviours. As such, it contributes to a growing recognition of the need to move towards more naturalistic paradigms in psychiatric research that better reflect real-world complexity. This shift is particularly important when investigating how learning and decision-making processes may be altered in mental health conditions, as findings need to extend beyond the laboratory to make meaningful impacts on people's lives. While some efforts have been made to increase ecological validity through gamified task variants, there are opportunities to explore more complex, naturalistic behaviours. For instance, recent work suggests that examining more complex planning processes in naturalistic settings could provide insights into symptoms like inattention and how it relates to model-based reinforcement learning (Toby Wise et al., 2024). By embracing the challenges of complex and potentially noisy environments, naturalistic computational psychiatry promises greater ecological validity and also a way to bridge the gap between computational models and the lived experiences of individuals with mental health conditions.

5.4.2 Potential clinical implications

The broader context of studying classical preclinical tests of anxiety in this thesis was that these measures are currently limited in their clinical utility, as they struggle to predict the efficacy of anxiolytic drugs in humans. Whilst I developed and validated new tasks and computational models of anxiety-related behaviour through my research, their immediate clinical value is not clear. The tasks demonstrated sensitivity to experimentally induced anxiety and pharmacological interventions in healthy humans, but I did not study their direct application in clinical settings. Regardless, it is still perhaps worth speculating on their potential clinical applications looking forward.

The computational parameters derived from these tasks, particularly the reward-punishment sensitivity index, could potentially serve as biomarkers for anxiety-related processes and apply to precision psychiatry approaches. Such markers could be used to subtype anxiety disorders or predict individual responses to treatments. For example, individuals with a particularly low reward-punishment sensitivity index might benefit from interventions specifically designed to increase reward sensitivity or decrease punishment sensitivity.

Furthermore, the tasks could be valuable in drug development pipelines, helping to bridge the gap between preclinical animal studies and human clinical trials. By providing a common computational framework across species, these tasks and their models promise to improve the translation of findings from animal models to human patients, potentially leading to more efficient and targeted development of anxiety treatments. The results from Chapter 4 are particularly interesting in this regard. The effects of anxiety-selective drugs on the elevated plus maze, and other tests of anxiety, are often conceptualised along an approach-avoidance axis (i.e. a drug either promotes approach or avoidance). This indeed appeared to be the case in the context of the opposing effects of lorazepam and yohimbine on human behaviour – lorazepam promoted exploration (approach) whilst yohimbine induced avoidance of the open arms. However, as noted earlier, the computational analyses suggested that lorazepam and yohimbine exerted their opposing effects via distinct and orthogonal mechanisms, specifically by reducing aversion to open arm exposure and decreasing the willingness to explore, respectively.

In the drug development process, being able to understand drug effects at this more granular level could be invaluable in the drug development pipeline, offering more sensitive and mechanistically informative measures of drug efficacy. It could help identify promising compounds earlier in the development process, disentangle signal from noise, or provide guidance on which specific anxiety-related processes a given compound is most likely to affect. With all these promises in mind however, extensive validation in clinical populations, such as individuals with diagnosed anxiety disorders, will be necessary before these tasks and their readouts (i.e. parameters/indices generated via computational analyses) could be considered for any clinical application.

5.5 Limitations and future directions

While the research of this thesis has demonstrated the value of taking a computational perspective on translational models of anxiety, there are several important missing parts of the story. I will frame these as questions for future research and attempt to provide suggestions for how they can be addressed.

5.5.1 Are the computational models of this thesis good accounts of non-human animal behaviour?

I have argued that the computational models implemented in my experimental chapters are species-agnostic descriptions of how humans and non-human animals alike behave in each relevant context. However, there has been an elephant in the room, and that is that my studies and experiments have not included animal subjects (elephant or rodent). Future work is necessary to empirically validate these models as satisfactory accounts of non-human animal behaviour, especially the novel spatial exploration models proposed in Chapter 4. As noted in Chapter 2, many studies by now have found evidence that model-free reinforcement learning mechanisms are good models of how rodents learn to pursue reward and avoid punishment (Mobbs et al., 2020; Schultz, 2013), but no study has modelled learning when subjects were required to do both simultaneously. Yet, as both mice and rats can perform learning tasks that involve two stimulus dimensions (Brown & Tait, 2016), it is arguably unlikely that a wholly distinct mechanism will be required to explain simultaneous reward and punishment learning.

It is worth thinking about the extent to which computational strategies in a given test/task across species should be similar to meet the criterion of computational validity. The extreme case of the same model and parameter distribution is highly unlikely and also unnecessary: for example, ethical constraints limit the maximum possible magnitudes of electric shocks for humans, but they are not as constrained in rodents (who also cannot be informed that they are 'harmless'), which can lead to differences in parameter values of punishment sensitivity. It may also be forgivable that the model specification differs across species to account for ethological idiosyncrasies. For example, humans tend to extrapolate inference across actions in two-choice tasks (e.g. if this action is bad, the other must be good; Palminteri et al., 2017), but to my knowledge, this effect has not been replicated in rodent decision-making tests. This implies that an additional counterfactual updating mechanism might provide a more parsimonious fit in the human data, although the general learning mechanism might be similar across species.

If it becomes clear that non-human subjects behave in the tests through totally different computational strategies compared to humans, refinements should be made to the current forms of the tests, or new ones should be developed with the explicit intention to match the strategies seen in humans, especially those that are relevant to how patients with anxiety disorders behave.

5.5.2 Are these computational mechanisms relevant to clinical anxiety?

As noted above, most of the research participants across my experimental chapters did not report high levels of anxiety symptoms, indeed this was specifically an exclusion criterion for the samples in

Chapters 3 and 4. The work from this thesis was therefore broadly aligned with dimensional approaches to psychiatry (Cuthbert & Insel, 2013; Insel et al., 2010) and the *experimental psychopathology* approach (Grillon et al., 2019), which assume that clinical anxiety and normal, adaptive anxiety lie on a continuum, and studying anxiety-related processes in healthy controls can inform something about cognition in anxiety patients.

However, the ultimate goal of this research is to improve outcomes for those struggling with debilitating anxiety, and so conducting studies including patients with clinical anxiety is necessary going forward. Specifically with regards to the context of this thesis, it will be important to validate the clinical relevance of the computational mechanisms engaged in the approach-avoidance reinforcement learning task developed in Chapter 2 and the spatial exploration tasks of Chapter 4. From the results from Chapters 2 and 3 relating to the reward-punishment sensitivity index, I would hypothesise that this computational measure may be exaggerated towards punishments in patients with anxiety disorders. For those who do show this bias, knowing that lorazepam can shift it away from punishments and towards rewards would be useful to allow for targeted treatment. That the correlation between the reward-punishment sensitivity index and anxiety symptoms (as opposed to state anxiety) was not replicable in the large online samples of Chapter 2 may indicate evidence against this hypothesis, but it is possible that the relationship between symptoms and this computational mechanism is not robust at the lower end of the symptom distribution, where most of the study participants were found.

Another important consideration for future research is to examine of how these computational mechanisms might differ across various anxiety subtypes. The current work has primarily focused on anxiety as a broad construct, without specifically differentiating between subtypes such as generalised anxiety disorder and social anxiety, or distinctions between somatic and cognitive anxiety. These subtypes may engage different computational processes or show varying degrees of alterations in the mechanisms identified in this thesis. For instance, individuals with social anxiety might show more pronounced shifts in reward-punishment sensitivities specifically in social contexts, while those with GAD might exhibit these shifts more consistently across multiple situations. Similarly, those experiencing predominantly somatic anxiety might show different patterns of spatial exploration compared to those with primarily cognitive anxiety symptoms.

Future work will be necessary to empirically validate the computational validity of these computational mechanisms as relevant to clinical anxiety, across a range of anxiety disorders. Although I cannot speak to any results yet as the work is ongoing, the approach-avoidance reinforcement learning task from Chapter 2 is currently being tested in a sample of generalised

anxiety disorder patients in the context of a clinical trial. There are also ongoing efforts to develop other mixed-reality translations of exploration-based anxiety tests like that of Chapter 4 in human patient samples, which will allow for the scaling of naturalistic measures of anxiety-related spatial planning.

5.5.3 Small effect sizes

Beyond the need for clinical validation, another important limitation that emerged across the experimental chapters was the observation of relatively small effect sizes for key findings. In Chapter 2, the correlation between task-induced anxiety and avoidance behaviour (as measured by the proportion of safe option choices) was small, with Kendall's tau values on the order of -0.075 across samples. Similarly, in Chapter 3, the effects of lorazepam and threat-of-shock on the proportion of approach/avoidance choices were subtle, with differences of 0.2% and -1.6% respectively, which did not reach statistical significance in simple t-tests. Even when examining computational parameters such as relative reward/punishment sensitivity, which did show statistically significant effects of the interventions, the effect sizes remained modest ($d = 0.22$ for lorazepam, $d = 0.40$ for threat-of-shock).

These small effect sizes have implications for interpreting and applying the findings. They raise questions about the clinical significance and real-world applicability of the observed relationships between anxiety, avoidance behaviour, and the underlying computational mechanisms. The subtle nature of these effects may limit the predictive power and utility of the approach-avoidance reinforcement learning task as a diagnostic tool or measure of treatment response in clinical settings.

Several factors may contribute to these small effect sizes. The use of non-clinical samples likely restricted the range of anxiety levels observed, potentially limiting the strength of correlations with avoidance behaviour. This limitation aligns with the discussion in section 5.2 regarding the need for validation in clinical populations. Additionally, the anxiety manipulations, particularly lorazepam administration, may have been too subtle to elicit large behavioural changes, especially given that lorazepam did not significantly reduce self-reported anxiety in healthy participants. Therefore, some directions for future research could be, as noted in the previous section, to study clinical samples with more severe anxiety – which may reveal larger effects and would enhance clinical relevance - but also to optimise tasks or study parameters to improve sensitivity. For example, the approach-avoidance reinforcement learning task design may have room for optimisation, such as in the dynamics of the reward/punishment probabilities, to improve sensitivity to anxiety. It may also be possible to optimise the experimental anxiety manipulations themselves to increase their impact on anxiety, for example by increasing the dose of lorazepam or changing the route of administration.

Small effect sizes are an important limitation, but it is worth thinking about them in the context of computational psychiatry as a relatively young and rapidly evolving field. Much of the current research, including the work presented in this thesis, is still in a 'discovery' phase. In this early stage, even subtle effects could provide valuable insights and direction for future research. The consistency of effects across my experiments, despite their small size, supports their validity and suggests they reflect real, albeit subtle, relationships. As it appears increasingly the case in the field, a key benefit of computational modelling in my analyses was its ability to reveal these subtle effects which were not detectable via traditional model-agnostic analyses. That said, computational modelling is no silver bullet for inefficient experimental design; while my experiments have yielded interesting insights, there is certainly room for further optimisation in both task and experimental design.

5.5.4 Do conditioned and unconditioned tests of anxiety overlap?

The classical conditioned and unconditioned tests of anxiety differ with respect to the computational mechanisms they engage, but there is some overlap at the point of choice – both forms of test involve trading off rewarding actions against punishing actions. An interesting question is then to what extent behaviour correlates across these tests within-subjects.

This is fundamentally an empirical question that has yet to be thoroughly investigated. To address this, future studies should aim to administer both types of tests to the same individuals, allowing for direct comparison of behaviour and computational parameters across tasks. Such research could involve participants completing both the approach-avoidance reinforcement learning task (as described in Chapters 2/3) and a spatial exploration task (like the one modelled in Chapter 4). This would allow the examination of correlations between key behavioural measures (e.g., avoidance behaviour in the reinforcement learning task and open arm exploration in a virtual elevated plus maze) as well as between computational parameters (e.g., the reward-punishment sensitivity index from the reinforcement learning model and the intrinsic reward weight from the spatial exploration model).

If strong associations are identified, it would suggest that these different tests are tapping into similar underlying anxiety-related processes, despite their surface-level differences. This would provide evidence for a common computational substrate of anxiety that manifests across diverse behavioural contexts. On the other hand, weak or absent correlations would suggest that these tests are capturing distinct aspects of anxiety-related cognition and behaviour. This outcome would highlight the multifaceted nature of anxiety and underscore the importance of using a diverse battery of tests to fully characterise anxiety-related processes. Additionally, factor analysis or principal component analysis could be employed to identify latent variables that might explain

patterns of covariation across different behavioural and computational measures. This could potentially reveal higher-order constructs that bridge across different types of anxiety tests.

5.6 Summary and concluding remarks

This thesis set out to address some fundamental questions in translational anxiety research: what exactly are the classical preclinical tests of anxiety measuring, and how do they relate to human anxiety? Through the empirical work presented here, I sought to answer these questions by applying computational approaches to both human and animal paradigms.

Chapter 2 addressed these issues by translating classical conditioned tests of anxiety for use in humans. This work revealed that these tests engage specific computational mechanisms - notably, the balance between reward and punishment sensitivities. Importantly, I found this computational measure correlated with self-reported state anxiety in humans, providing a direct link between the test's measures and subjective experiences of anxiety. Chapter 3 further validated this computational mechanism by demonstrating its sensitivity to both anxiolytic (lorazepam) and anxiogenic (threat-of-shock) manipulations in humans, mirroring how these tests in animals are typically validated. Chapter 4 extended this computational lens to unconditioned tests of anxiety, such as the elevated plus maze. By developing a novel computational model of spatial exploration under threat, this work provided new insights into what these tests measure, specifically processes related to exploration drive and threat sensitivity. Importantly, the model parameters were found to correlate with human anxiety levels and respond to anxiolytic and anxiogenic drugs, demonstrating their relevance to human anxiety.

Collectively, these findings demonstrate that by reframing classical tests of anxiety in computational terms, we can move beyond loose behavioural descriptions to understand the precise information-processing mechanisms these tests engage and how they relate to human anxiety. This approach offers a more rigorous way to conceptualise and compare anxiety-related processes across species, potentially bridging the long-standing translational gap in anxiety research. Indeed, the implications of this work extend beyond simply improving our understanding of existing tests. They suggest a fundamental shift in how we should approach translational models of anxiety. Rather than conceptualising these behavioural tests as models of 'anxiety' or specific DSM categories of anxiety disorders, we should view them as probes of specific computational processes. This perspective acknowledges the limitations of broad, ill-defined constructs like 'anxiety-like' behaviours, which many researchers in the field already recognise. Instead, it advocates for a more precise, mathematically grounded description of the information-processing problems these tests engage.

The key take-home message of the thesis is that, extending the linguistic connotation of *translational* approaches, computational models can act as the *lingua franca* that bridges cross-species anxiety research, in a way that is species-agnostic and provides a unifying framework for studying the neural computations that underlie symptoms of anxiety, and how certain treatments or drugs affect these computations. Looking forward, I believe this computational perspective is a promising path for driving much-needed innovation in translational anxiety research, and in the longer term I very much hope that this approach contributes to improving outcomes for individuals suffering from anxiety.

6 References

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review*, *30*(2), 217-237. https://doi.org/10.1016/j.cpr.2009.11.004

American Psychiatric Association. (1952). *Diagnostic and statistical manual of mental disorders*.

American Psychiatric Association. (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed., text rev. ed.).

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed., text rev. ed.).

American Psychiatric Association. (2013a). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

American Psychiatric Association. (2013b). *Diagnostic and statistical manual of mental disorders: DSM-5* (Vol. 5). American psychiatric association Washington, DC.

Arantes, R., Tejada, J., Bosco, G. G., Morato, S., & Roque, A. C. (2013). Mathematical methods to model rodent behavior in the elevated plus-maze. *J Neurosci Methods*, *220*(2), 141-148. https://doi.org/10.1016/j.jneumeth.2013.04.022

Auer, P., Cesa-Bianchi, N., & Fischer, P. (2002). Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, *47*(2-3), 235-256. https://doi.org/Doi 10.1023/A:1013689704352

Aupperle, R. L., Melrose, A. J., Francisco, A., Paulus, M. P., & Stein, M. B. (2015). Neural Substrates of Approach-Avoidance Conflict Decision-Making. *Human Brain Mapping*, *36*(2), 449-462. https://doi.org/10.1002/hbm.22639

Aupperle, R. L., & Paulus, M. P. (2010). Neural systems underlying approach and avoidance in anxiety disorders. *Dialogues Clin Neurosci*, *12*(4), 517-531. https://www.ncbi.nlm.nih.gov/pubmed/21319496

Aupperle, R. L., Sullivan, S., Melrose, A. J., Paulus, M. P., & Stein, M. B. (2011). A reverse translational approach to quantify approach-avoidance conflict in humans. *Behavioural Brain Research*, *225*(2), 455-463. https://doi.org/10.1016/j.bbr.2011.08.003

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. https://doi.org/Pii S0033291718004117

10.1017/S0033291718004117

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. https://doi.org/10.1038/s41562-019-0628-0

Bach, D. R. (2021). Cross-species anxiety tests in psychiatry: pitfalls and promises. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-021-01299-4

Bach, D. R., Guitart-Masip, M., Packard, P. A., Miro, J., Falip, M., Fuentemilla, L., & Dolan, R. J. (2014). Human Hippocampus Arbitrates Approach-Avoidance Conflict (vol 24, pg 541, 2014). *Current Biology*, *24*(12), 1435-1435. https://doi.org/10.1016/j.cub.2014.05.051

Bandelow, B., Reitt, M., Röver, C., Michaelis, S., Görlich, Y., & Wedekind, D. (2015). Efficacy of treatments for anxiety disorders: a meta-analysis. *International Clinical Psychopharmacology*, *30*(4), 183-192. https://doi.org/10.1097/Yic.0000000000000078

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. https://doi.org/10.1037/0033-2909.133.1.1

Barlow, D. H. (2002). *Anxiety and its disorders: The nature and treatment of anxiety and panic, 2nd ed*. The Guilford Press.

Baron, R. M., & Kenny, D. A. (1986). The Moderator Mediator Variable Distinction in Social Psychological-Research - Conceptual, Strategic, and Statistical Considerations. *Journal of*

Personality and Social Psychology, *51*(6), 1173-1182. https://doi.org/Doi 10.1037/0022-3514.51.6.1173

Beck, A. T., & Clark, D. A. (1988). Anxiety and depression: An information processing perspective. *Anxiety research*, *1*(1), 23-36.

Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. *Neuron*, *100*(2), 490-509. https://doi.org/10.1016/j.neuron.2018.10.002

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. https://doi.org/10.1038/nn1954

Bellemare, M. G., Srinivasan, S., Ostrovski, G., Schaul, T., Saxton, D., & Munos, R. (2016). Unifying Count-Based Exploration and Intrinsic Motivation. *Advances in Neural Information Processing Systems 29 (Nips 2016)*, *29*. <Go to ISI>://WOS:000458973703019

Bhattacharya, S. K., Satyan, K. S., & Chakrabarti, A. (1997). Anxiogenic action of caffeine: an experimental study in rats. *J Psychopharmacol*, *11*(3), 219-224. https://doi.org/10.1177/026988119701100304

Biedermann, S. V., Biedermann, D. G., Wenzlaff, F., Kurjak, T., Nouri, S., Auer, M. K., Wiedemann, K., Briken, P., Haaker, J., Lonsdorf, T. B., & Fuss, J. (2017). An elevated plus-maze in mixed reality for studying human anxiety-related behavior. *Bmc Biology*, *15*. https://doi.org/ARTN 125

10.1186/s12915-017-0463-6

Bishop, S. J. (2007). Neurocognitive mechanisms of anxiety: an integrative account. *Trends in Cognitive Sciences*, *11*(7), 307-316. https://doi.org/10.1016/j.tics.2007.05.008

Bishop, S. J., & Gagne, C. (2018). Anxiety, Depression, and Decision Making: A Computational Perspective. *Annual Review of Neuroscience, Vol 41*, *41*, 371-388. https://doi.org/10.1146/annurev-neuro-080317-062007

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433-436. https://doi.org/Doi 10.1163/156856897x00357

Brown, V. J., & Tait, D. S. (2016). Attentional Set-Shifting Across Species. *Translational Neuropsychopharmacology*, *28*, 363-395. https://doi.org/10.1007/7854_2015_5002

Browning, M., Behrens, T. E., Jocham, G., O'Reillly, J. X., & Bishop, S. J. (2015). Anxious Individuals Have Difficulty Learning the Causal Statistics of Aversive Environments. *Biological Psychiatry*, *77*(9), 47s-48s. <Go to ISI>://WOS:000352207500123

Bystritsky, A. (2006). Treatment-resistant anxiety disorders. *Molecular Psychiatry*, *11*(9), 805-814. https://doi.org/10.1038/sj.mp.4001852

Calvin, O. L., Erickson, M. T., Walters, C. J., & Redish, A. D. (2024). Dorsal hippocampus represents locations to avoid as well as locations to approach during approach-avoidance conflict. *bioRxiv*. https://doi.org/10.1101/2024.03.10.584295

Campos, A. C., Fogaca, M. V., Aguiar, D. C., & Guimaraes, F. S. (2013). Animal models of anxiety disorders and stress. *Revista Brasileira De Psiquiatria*, *35*, S101-S111. https://doi.org/10.1590/1516-4446-2013-1139

Carlisi, C. O., & Robinson, O. J. (2018). The role of prefrontal-subcortical circuitry in negative bias in anxiety: Translational, developmental and treatment perspectives. *Brain Neurosci Adv*, *2*, 2398212818774223. https://doi.org/10.1177/2398212818774223

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. https://doi.org/Doi 10.1037/0022-3514.67.2.319

Charney, D. S., Heninger, G. R., & Redmond, D. E., Jr. (1983). Yohimbine induced anxiety and increased noradrenergic function in humans: effects of diazepam and clonidine. *Life Sci*, *33*(1), 19-29. https://doi.org/10.1016/0024-3205(83)90707-5

Chebib, M., & Johnston, G. A. R. (1999). The 'ABC' of GABA receptors: A brief review. *Clinical and Experimental Pharmacology and Physiology*, *26*(11), 937-940. https://doi.org/DOI 10.1046/j.1440-1681.1999.03151.x

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284.

CIPD. (2023). *Health and wellbeing at work*.

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. https://doi.org/10.1016/j.cpr.2009.11.003

Clark, L., Li, R. R., Wright, C. M., Rome, F., Fairchild, G., Dunn, B. D., & Aitken, M. R. F. (2012). Risk-avoidant decision making increased by threat of electric shock. *Psychophysiology*, *49*(10), 1436-1443. https://doi.org/10.1111/j.1469-8986.2012.01454.x

Cook, D., Brown, D., Alexander, R., March, R., Morgan, P., Satterthwaite, G., & Pangalos, M. N. (2014). Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. *Nature Reviews Drug Discovery*, *13*(6), 419-431. https://doi.org/10.1038/nrd4309

Corr, P. J. (2004). Reinforcement sensitivity theory and personality. *Neuroscience and Biobehavioral Reviews*, *28*(3), 317-332. https://doi.org/10.1016/j.neubiorev.2004.01.005

Craske, M. G., Hermans, D. E., & Vansteenwegen, D. E. (2006). *Fear and learning: From basic processes to clinical implications*. American Psychological Association.

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: an inhibitory learning approach. *Behaviour Research and Therapy*, *58*, 10-23. https://doi.org/10.1016/j.brat.2014.04.006

Crawley, J., & Goodwin, F. K. (1980). Preliminary-Report of a Simple Animal Behavior Model for the Anxiolytic Effects of Benzodiazepines. *Pharmacology Biochemistry and Behavior*, *13*(2), 167-170. https://doi.org/Doi 10.1016/0091-3057(80)90067-2

Cryan, J. F., & Holmes, A. (2005). The ascent of mouse: Advances in modelling human depression and anxiety. *Nature Reviews Drug Discovery*, *4*(9), 775-790. https://doi.org/10.1038/nrd1825

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *Bmc Medicine*, *11*. https://doi.org/Artn 126

10.1186/1741-7015-11-126

Daniel-Watanabe, L., & Fletcher, P. C. (2021). Are Fear and Anxiety Truly Distinct? *Biological Psychiatry Global Open Science*.

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. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876-879. https://doi.org/10.1038/nature04766

Dayan, P. (1993). Improving Generalization for Temporal Difference Learning - the Successor Representation. *Neural Computation*, *5*(4), 613-624. https://doi.org/DOI 10.1162/neco.1993.5.4.613

Deady, M., Collins, D. A. J., Johnston, D. A., Glozier, N., Calvo, R. A., Christensen, H., & Harvey, S. B. (2022). The impact of depression, anxiety and comorbidity on occupational outcomes. *Occupational Medicine-Oxford*, *72*(1), 17-24. https://doi.org/10.1093/occmed/kqab142

Dillon, D. G., Rosso, I. M., Pechtel, P., Killgore, W. D. S., Rauch, S. L., & Pizzagalli, D. A. (2014). Peril and Pleasure: An Rdoc-Inspired Examination of Threat Responses and Reward Processing in Anxiety and Depression. *Depression and Anxiety*, *31*(3), 233-249. https://doi.org/10.1002/da.22202

Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., van den Hout, M. A., & Baas, J. M. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and Anxiety*, *32*(4), 239-253. https://doi.org/10.1002/da.22353

Eckstein, M. K., Master, S. L., Xia, L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. E. (2022). The interpretation of computational model parameters depends on the context. *Elife*, *11*. https://doi.org/10.7554/eLife.75474

Ekman, M., Kusch, S., & de Lange, F. P. (2023). Successor-like representation guides the prediction of future events in human visual cortex and hippocampus. *Elife*, *12*. https://doi.org/ARTN e78904

10.7554/eLife.78904

Eldar, E., Hauser, T. U., Dayan, P., & Dolan, R. J. (2016). Striatal structure and function predict individual biases in learning to avoid pain. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(17), 4812-4817. https://doi.org/10.1073/pnas.1519829113

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.

Ferrari, A. J., Santomauro, D. F., Herrera, A. M. M., Shadid, J., Ashbaugh, C., Erskine, H. E., Charlson, F. J., Degenhardt, L., Scott, J. G., McGrath, J. J., Allebeck, P., Benjet, C., Breitborde, N. J. K., Brugha, T., Dai, X. C., Dandona, L., Dandona, R., Fischer, F., Haagsma, J. A., . . . Collabor, G. M. D. (2022). Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Psychiatry*, *9*(2), 137-150. https://doi.org/10.1016/S2215-0366(21)00395-3

Fleiss, J. L. (1986). *The design and analysis of clinical experiments*. Wiley. https://doi.org/10.1002/9781118032923

Freedman, R., Lewis, D. A., Michels, R., Pine, D. S., Schultz, S. K., Tamminga, C. A., Gabbard, G. O., Gau, S. S. F., Javitt, D. C., Oquendo, M. A., Shrout, P. E., Vieta, E., & Yager, J. (2013). The Initial Field Trials of DSM-5: New Blooms and Old Thorns. *American Journal of Psychiatry*, *170*(1), 1-5. https://doi.org/10.1176/appi.ajp.2012.12091189

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active Inference: A Process Theory. *Neural Computation*, *29*(1), 1-49. https://doi.org/10.1162/NECO_a_00912

Galatzer-Levy, I. R., & Bryant, R. A. (2013). 636,120 Ways to Have Posttraumatic Stress Disorder. *Perspectives on Psychological Science*, *8*(6), 651-662. https://doi.org/10.1177/1745691613504115

Gamez, W., Chmielewski, M., Kotov, R., Ruggero, C., Suzuki, N., & Watson, D. (2014). The Brief Experiential Avoidance Questionnaire: Development and Initial Validation. *Psychological Assessment*, *26*(1), 35-45. https://doi.org/10.1037/a0034473

Garcez, H., Fernandes, C., Barbosa, F., Pereira, M. R., Silveira, C., Marques-Teixeira, J., & Goncalves, A. R. (2020). Effects of benzodiazepines administration on identification of facial expressions of emotion: a meta-analysis. *Psychopharmacology (Berl)*, *237*(1), 1-9. https://doi.org/10.1007/s00213-019-05393-y

Geller, I., Kulak, J. T., & Seifter, J. (1962). The Effects of Chlordiazepoxide and Chlorpromazine on a Punishment Discrimination. *Psychopharmacologia*, *3*(5), 374-385. https://doi.org/Doi 10.1007/Bf00408322

Gershman, S. J. (2018). Deconstructing the human algorithms for exploration. *Cognition*, *173*, 34-42. https://doi.org/10.1016/j.cognition.2017.12.014

Gillan, C. M., Vaghi, M. M., Hezemans, F. H., Grothe, S. V., Dafflon, J., Bruhl, A. B., Savulich, G., & Robbins, T. W. (2021). Experimentally induced and real-world anxiety have no demonstrable effect on goal-directed behaviour. *Psychological Medicine*, *51*(9), 1467-1478. https://doi.org/Pii S0033291720000203

10.1017/S0033291720000203

Glover, L. R., Postle, A. F., & Holmes, A. (2020). Touchscreen-based assessment of risky-choice in mice. *Behavioural Brain Research*, *393*. https://doi.org/ARTN 112748

10.1016/j.bbr.2020.112748
Gray, J. A. (1982). The Neuropsychology of Anxiety - an Inquiry into the Functions of the Septo-Hippocampal System. *Behavioral and Brain Sciences*, *5*(3), 469-484. https://doi.org/Doi 10.1017/S0140525x00013066
Green, S., & Hodges, H. (1991). Animal models of anxiety. In *Behavioural models in psychopharmacology: Theoretical, industrial and clinical perspectives.* (pp. 21-49). Cambridge University Press.
Griebel, G. (1995). 5-Hydroxytryptamine-Interacting Drugs in Animal-Models of Anxiety Disorders - More Than 30 Years of Research. *Pharmacology & Therapeutics*, *65*(3), 319-395. https://doi.org/Doi 10.1016/0163-7258(95)98597-J
Griebel, G., & Holmes, A. (2013). 50 years of hurdles and hope in anxiolytic drug discovery. *Nature Reviews Drug Discovery*, *12*(9), 667-687. https://doi.org/10.1038/nrd4075
Grillon, C., Baas, J. M. P., Pine, D. S., Lissek, S., Lawley, M., Ellis, V., & Levine, J. (2006). The Benzodiazepine Alprazolam Dissociates Contextual Fear from Cued Fear in Humans as Assessed by Fear-potentiated Startle. *Biological Psychiatry*, *60*(7), 760-766. https://doi.org/https://doi.org/10.1016/j.biopsych.2005.11.027
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*, *44*(12), 1999-2010. https://doi.org/10.1038/s41386-019-0445-1
Guitart-Masip, M., Nuys, Q. J. M., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. *Neuroimage*, *62*(1), 154-166. https://doi.org/10.1016/j.neuroimage.2012.04.024
Hales, C. A., Robinson, E. S. J., & Houghton, C. J. (2016). Diffusion Modelling Reveals the Decision Making Processes Underlying Negative Judgement Bias in Rats. *Plos One*, *11*(3). https://doi.org/ARTN e0152592

10.1371/journal.pone.0152592
Hall, C. S. (1934). Emotional behavior in the rat I Defecation and urination as measures of individual differences in emotionality. *Journal of Comparative Psychology*, *18*(3), 385-403. https://doi.org/DOI 10.1037/h0071444
Handley, S. L., & Mithani, S. (1984). Effects of Alpha-Adrenoceptor Agonists and Antagonists in a Maze-Exploration Model of Fear-Motivated Behavior. *Naunyn-Schmiedebergs Archives of Pharmacology*, *327*(1), 1-5. https://doi.org/Doi 10.1007/Bf00504983
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. *British Journal of Psychiatry*, *195*(2), 102-108. https://doi.org/10.1192/bjp.bp.108.051193
Hartley, C. A., & Phelps, E. A. (2012). Anxiety and Decision-Making. *Biological Psychiatry*, *72*(2), 113-118. https://doi.org/10.1016/j.biopsych.2011.12.027
Hayes, S. C. (1976). The role of approach contingencies in phobic behavior. *Behavior Therapy*.
Hayes, S. C., Wilson, K. G., Gifford, E. V., Follette, V. M., & Strosahl, K. (1996). Experiential avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. *Journal of Consulting and Clinical Psychology*, *64*(6), 1152-1168. https://doi.org/Doi 10.1037/0022-006x.64.6.1152
Hayman, R. M. A., Chakraborty, S., Anderson, M. I., & Jeffery, K. J. (2003). Context-specific acquisition of location discrimination by hippocampal place cells. *European Journal of Neuroscience*, *18*(10), 2825-2834. https://doi.org/10.1111/j.1460-9568.2003.03035.x
Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517-523. https://doi.org/10.1016/j.tics.2009.09.004
Hofmann, S. G., & Smits, J. A. J. (2008). Cognitive-behavioral therapy for adult anxiety disorders: A meta-analysis of randomized placebo-controlled trials. *Journal of Clinical Psychiatry*, *69*(4), 621-632. https://doi.org/DOI 10.4088/jcp.v69n0415

Huang, H., Thompson, W., & Paulus, M. P. (2017). Computational Dysfunctions in Anxiety: Failure to Differentiate Signal From Noise. *Biological Psychiatry*, *82*(6), 440-446. https://doi.org/10.1016/j.biopsych.2017.07.007

Huys, Q. J. M., Cools, R., Golzer, M., Friedel, E., Heinz, A., Dolan, R. J., & Dayan, P. (2011). Disentangling the Roles of Approach, Activation and Valence in Instrumental and Pavlovian Responding. *Plos Computational Biology*, *7*(4). https://doi.org/ARTN e1002028

10.1371/journal.pcbi.1002028

Huys, Q. J. M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai Trees in Your Head: How the Pavlovian System Sculpts Goal-Directed Choices by Pruning Decision Trees. *Plos Computational Biology*, *8*(3). https://doi.org/ARTN e1002410

10.1371/journal.pcbi.1002410

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. https://doi.org/10.1038/nn.4238

Hyman, S. E. (2010). The Diagnosis of Mental Disorders: The Problem of Reification. *Annual Review of Clinical Psychology, Vol 6*, *6*, 155-179. https://doi.org/10.1146/annurev.clinpsy.3.022806.091532

Hyman, S. E. (2012). Revolution Stalled. *Science Translational Medicine*, *4*(155). https://doi.org/ARTN 155cm11

10.1126/scitranslmed.3003142

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research Domain Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders. *American Journal of Psychiatry*, *167*(7), 748-751. https://doi.org/10.1176/appi.ajp.2010.09091379

Ironside, M., Amemori, K., McGrath, C. L., Pedersen, M. L., Kang, M. S., Amemori, S., Frank, M. J., Graybiel, A. M., & Pizzagalli, D. A. (2020). Approach-Avoidance Conflict in Major Depressive Disorder: Congruent Neural Findings in Humans and Nonhuman Primates. *Biological Psychiatry*, *87*(5), 399-408. https://doi.org/10.1016/j.biopsych.2019.08.022

Jean-Richard-Dit-Bressel, P., Lee, J. C., Liew, S. X., Weidemann, G., Lovibond, P. F., & McNally, G. P. (2021). Punishment insensitivity in humans is due to failures in instrumental contingency learning. *Elife*, *10*. https://doi.org/ARTN e69594

10.7554/eLife.69594

Kakoschke, N., Albertella, L., Lee, R. S. C., & Wiers, R. W. (2019). Assessment of Automatically Activated Approach-Avoidance Biases Across Appetitive Substances. *Current Addiction Reports*, *6*(3), 200-209. https://doi.org/10.1007/s40429-019-00254-2

Kaplan, E. M., & DuPont, R. L. (2005). Benzodiazepines and anxiety disorders: a review for the practicing physician. *Current Medical Research and Opinion*, *21*(6), 941-950. https://doi.org/10.1185/030079905x48401

Karlsson, M. P., & Frank, L. M. (2009). Awake replay of remote experiences in the hippocampus. *Nature Neuroscience*, *12*(7), 913-U132. https://doi.org/10.1038/nn.2344

Kesselheim, A. S., Hwang, T. J., & Franklin, J. M. (2015). Two decades of new drug development for central nervous system disorders. *Nature Reviews Drug Discovery*, *14*(12), 815-816. https://doi.org/10.1038/nrd4793

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H. U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research*, *21*(3), 169-184. https://doi.org/10.1002/mpr.1359

Kirk, P. A., Holmes, A. J., & Robinson, O. J. (2023). Anxiety Shapes Amygdala-Prefrontal Dynamics During Movie Watching. *Biological Psychiatry: Global Open Science*, *3*(3). https://doi.org/10.1016/j.bpsgos.2022.03.009

Kirlic, N., Young, J., & Aupperle, R. L. (2017). Animal to human translational paradigms relevant for approach avoidance conflict decision making. *Behaviour Research and Therapy*, *96*, 14-29. https://doi.org/10.1016/j.brat.2017.04.010

Koen, N., & Stein, D. J. (2011). Pharmacotherapy of anxiety disorders: a critical review. *Dialogues Clin Neurosci*, *13*(4), 423-437. https://doi.org/10.31887/DCNS.2011.13.4/nkoen

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., . . . Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, *126*(4), 454-477. https://doi.org/10.1037/abn0000258

Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B. W., Berry, J. T., & Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, *114*(1-3), 163-173. https://doi.org/10.1016/j.jad.2008.06.026

Letkiewicz, A. M., Kottler, H. C., Shankman, S. A., & Cochran, A. L. (2023). Quantifying aberrant approach-avoidance conflict in psychopathology: A review of computational approaches. *Neuroscience and Biobehavioral Reviews*, *147*. https://doi.org/ARTN 105103

10.1016/j.neubiorev.2023.105103

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. https://doi.org/10.1038/nn.2904

Lilienfeld, S. O. (2014). The Research Domain Criteria (RDoC): an analysis of methodological and conceptual challenges. *Behaviour Research and Therapy*, *62*, 129-139. https://doi.org/10.1016/j.brat.2014.07.019

Maia, T. V. (2010). Two-factor theory, the actor-critic model, and conditioned avoidance. *Learning & Behavior*, *38*(1), 50-67. https://doi.org/10.3758/Lb.38.1.50

Mathews, A., & Mackintosh, B. (1998). A cognitive model of selective processing in anxiety. *Cognitive Therapy and Research*, *22*(6), 539-560. https://doi.org/Doi 10.1023/A:1018738019346

Menard, J., & Treit, D. (1999). Effects of centrally administered anxiolytic compounds in animal models of anxiety. *Neuroscience and Biobehavioral Reviews*, *23*(4), 591-613. https://doi.org/Doi 10.1016/S0149-7634(98)00056-6

Metha, J. A., Brian, M. L., Oberrauch, S., Barnes, S. A., Featherby, T. J., Bossaerts, P., Murawski, C., Hoyer, D., & Jacobson, L. H. (2020). Separating Probability and Reversal Learning in a Novel Probabilistic Reversal Learning Task for Mice. *Frontiers in Behavioral Neuroscience*, *13*. https://doi.org/ARTN 270

10.3389/fnbeh.2019.00270

Milad, M. R., & Quirk, G. J. (2012). Fear Extinction as a Model for Translational Neuroscience: Ten Years of Progress. *Annual Review of Psychology, Vol 63*, *63*, 129-151. https://doi.org/10.1146/annurev.psych.121208.131631

Miller, G. (2010). Is Pharma Running Out of Brainy Ideas? *Science*, *329*(5991), 502-504. https://doi.org/doi:10.1126/science.329.5991.502

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. https://doi.org/10.1016/j.biopsych.2017.01.017

Mobbs, D., Headley, D. B., Ding, W. L., & Dayan, P. (2020). Space, Time, and Fear: Survival Computations along Defensive Circuits. *Trends in Cognitive Sciences*, *24*(3), 228-241. https://doi.org/10.1016/j.tics.2019.12.016

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. https://doi.org/10.1016/j.neuron.2021.05.021

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, *16*(1), 72-80. https://doi.org/10.1016/j.tics.2011.11.018

Morgan, L. K., MacEvoy, S. P., Aguirre, G. K., & Epstein, R. A. (2011). Distances between Real-World Locations Are Represented in the Human Hippocampus. *Journal of Neuroscience*, *31*(4), 1238-1245. https://doi.org/10.1523/Jneurosci.4667-10.2011

Murphy, S. E., Downham, C., Cowen, P. J., & Harmer, C. J. (2008). Direct effects of diazepam on emotional processing in healthy volunteers. *Psychopharmacology (Berl)*, *199*(4), 503-513. https://doi.org/10.1007/s00213-008-1082-2

Nair, A., Rutledge, R. B., & Mason, L. (2020). Under the Hood: Using Computational Psychiatry to Make Psychological Therapies More Mechanism-Focused. *Frontiers in Psychiatry*, *11*. https://doi.org/ARTN 140

10.3389/fpsyt.2020.00140

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R-2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, *14*(134). https://doi.org/ARTN 20170213

10.1098/rsif.2017.0213

Nikbakht, N., Pofahl, M., Miguel-López, A., Kamali, F., Tchumatchenko, T., & Beck, H. (2024). Efficient encoding of aversive location by CA3 long-range projections. *Cell Reports*, *43*(3). https://doi.org/ARTN 113957

10.1016/j.celrep.2024.113957

Nutt, D. J., & Malizia, A. L. (2001). New insights into the role of the GABA

-benzodiazepine receptor in psychiatric disorder. *British Journal of Psychiatry*, *179*, 390-396. https://doi.org/DOI 10.1192/bjp.179.5.390

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Clarendon Press ;

Oxford University Press.

Oberrauch, S., Sigrist, H., Sautter, E., Gerster, S., Bach, D. R., & Pryce, C. R. (2019). Establishing operant conflict tests for the translational study of anxiety in mice. *Psychopharmacology*, *236*(8), 2527-2541. https://doi.org/10.1007/s00213-019-05315-y

OKeefe, J., & Burgess, N. (1996). Geometric determinants of the place fields of hippocampal neurons. *Nature*, *381*(6581), 425-428. https://doi.org/DOI 10.1038/381425a0

Palminteri, S., Lefebvre, G., Kilford, E. J., & Blakemore, S. J. (2017). Confirmation bias in human reinforcement learning: Evidence from counterfactual feedback processing. *Plos Computational Biology*, *13*(8). https://doi.org/ARTN e1005684

10.1371/journal.pcbi.1005684

Park, H., Lee, D., & Chey, J. (2017). Stress enhances model-free reinforcement learning only after negative outcome. *Plos One*, *12*(7). https://doi.org/ARTN e0180588

10.1371/journal.pone.0180588

Patel, V., Chisholm, D., & Parikh, R. (2016). Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities (vol 387, pg 1672, 2016). *Lancet*, *387*(10028), 1618-1618. <Go to ISI>://WOS:000374145100026

Paulus, M. P., Feinstein, J. S., Castillo, G., Simmons, A. N., & Stein, M. B. (2005). Dose-dependent decrease of activation in bilateral amygdala and insula by lorazepam during emotion processing. *Archives of General Psychiatry*, *62*(3), 282-288. https://doi.org/DOI 10.1001/archpsyc.62.3.282

Pedersen, M. L., Ironside, M., Amemori, K., McGrath, C. L., Kang, M. S., Graybiel, A. M., Pizzagalli, D. A., & Frank, M. J. (2021). Computational phenotyping of brain-behavior dynamics underlying approach-avoidance conflict in major depressive disorder. *Plos Computational Biology*, *17*(5). https://doi.org/ARTN e1008955

10.1371/journal.pcbi.1008955

Pellow, S., Chopin, P., File, S. E., & Briley, M. (1985). Validation of Open - Closed Arm Entries in an Elevated Plus-Maze as a Measure of Anxiety in the Rat. *Journal of Neuroscience Methods*, *14*(3), 149-167. https://doi.org/Doi 10.1016/0165-0270(85)90031-7

Perusini, J. N., & Fanselow, M. S. (2015). Neurobehavioral perspectives on the distinction between fear and anxiety. *Learn Mem*, *22*(9), 417-425. https://doi.org/10.1101/lm.039180.115

Phaf, R. H., Mohr, S. E., Rotteveel, M., & Wicherts, J. M. (2014). Approach, avoidance, and affect: a meta-analysis of approach-avoidance tendencies in manual reaction time tasks. *Frontiers in Psychology*, *5*. https://doi.org/ARTN 378

10.3389/fpsyg.2014.00378

Pike, A. C., Lowther, M., & Robinson, O. J. (2021). The Importance of Common Currency Tasks in Translational Psychiatry. *Current Behavioral Neuroscience Reports*, *8*(1), 1-10. https://doi.org/10.1007/s40473-021-00225-w

Pike, A. C., & Robinson, O. J. (2022). Reinforcement Learning in Patients With Mood and Anxiety Disorders vs Control Individuals A Systematic Review and Meta-analysis. *Jama Psychiatry*. https://doi.org/10.1001/jamapsychiatry.2022.0051

Pittig, A., Alpers, G. W., Niles, A. N., & Craske, M. G. (2015). Avoidant decision-making in social anxiety disorder: A laboratory task linked to in vivo anxiety and treatment outcome. *Behaviour Research and Therapy*, *73*, 96-103. https://doi.org/10.1016/j.brat.2015.08.003

Pizzagalli, D. A., Jahn, A. L., & O'Shea, J. P. (2005). Toward an objective characterization of an anhedonic phenotype: A signal detection approach. *Biological Psychiatry*, *57*(4), 319-327. https://doi.org/10.1016/j.biopsych.2004.11.026

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, *40*(3), 879-891. https://doi.org/Doi 10.3758/Brm.40.3.879

Prut, L., & Belzung, C. (2003). The open field as a paradigm to measure the effects of drugs on anxiety-like behaviors: a review. *European Journal of Pharmacology*, *463*(1), 3-33. https://doi.org/https://doi.org/10.1016/S0014-2999(03)01272-X

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

Redish, A. D., Kepecs, A., Anderson, L. M., Calvin, O. L., Grissom, N. M., Haynos, A. F., Heilbronner, S. R., Herman, A. B., Jacob, S., Ma, S. S., Vilares, I., Vinogradov, S., Walters, C. J., Widge, A. S., Zick, J. L., & Zilverstand, A. (2022). Computational validity: using computation to translate behaviours across species. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *377*(1844). https://doi.org/ARTN 20200525

10.1098/rstb.2020.0525

Robbins, T. W., & Cardinal, R. N. (2019). Computational psychopharmacology: a translational and pragmatic approach. *Psychopharmacology (Berl)*, *236*(8), 2295-2305. https://doi.org/10.1007/s00213-019-05302-3

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*. https://doi.org/ARTN 203

10.3389/fnhum.2013.00203

Roelofs, K. (2017). Freeze for action: neurobiological mechanisms in animal and human freezing. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *372*(1718). https://doi.org/ARTN 20160206

10.1098/rstb.2016.0206
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. *European Journal of Neuroscience*, *35*(7), 1190-1200. https://doi.org/10.1111/j.1460-9568.2011.07986.x
Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of statistical software*, *48*, 1-36.
Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J., & Daw, N. D. (2017). Predictive representations can link model-based reinforcement learning to model-free mechanisms. *Plos Computational Biology*, *13*(9). https://doi.org/ARTN e1005768

10.1371/journal.pcbi.1005768
Rutledge, R. B., Chekroud, A. M., & Huys, Q. J. M. (2019). Machine learning and big data in psychiatry: toward clinical applications. *Current Opinion in Neurobiology*, *55*, 152-159. https://doi.org/10.1016/j.conb.2019.02.006
Salters-Pedneault, K., Tull, M. T., & Roemer, L. (2004). The role of avoidance of emotional material in the anxiety disorders. *Applied & Preventive Psychology*, *11*(2), 95-114. https://doi.org/10.1016/j.appsy.2004.09.001
Santomauro, D. F., Herrera, A. M. M., Shadid, J., Zheng, P., Ashbaugh, C., Pigott, D. M., Abbafati, C., Adolph, C., Amlag, J. O., Aravkin, A. Y., Bang-Jensen, B. L., Bertolacci, G. J., Bloom, S. S., Castellano, R., Castro, E., Chakrabarti, S., Chattopadhyay, J., Cogen, R. M., Collins, J. K., . . . Collabo, C.-M. D. (2021). Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *Lancet*, *398*(10312), 1700-1712. https://doi.org/10.1016/S0140-6736(21)02143-7
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. https://doi.org/10.1038/nprot.2012.001
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. https://doi.org/10.1016/j.bbr.2017.09.050
Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J Neurophysiol*, *80*(1), 1-27. https://doi.org/10.1152/jn.1998.80.1.1
Schultz, W. (2013). Updating dopamine reward signals. *Current Opinion in Neurobiology*, *23*(2), 229-238. https://doi.org/10.1016/j.conb.2012.11.012
Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593-1599. https://doi.org/10.1126/science.275.5306.1593
Sebold, M., Garbusow, M., Jetzschmann, P., Schad, D. J., Nebe, S., Schlagenhauf, F., Heinz, A., Rapp, M., & Romanczuk-Seiferth, N. (2019). Reward and avoidance learning in the context of aversive environments and possible implications for depressive symptoms. *Psychopharmacology*, *236*(8), 2437-2449. https://doi.org/10.1007/s00213-019-05299-9
Seow, T. X. F., & Hauser, T. U. (2022). Reliability of web-based affective auditory stimulus presentation. *Behavior research methods*, *54*(1), 378-392. https://doi.org/10.3758/s13428-021-01643-0
Seymour, B., Daw, N., Dayan, P., Singer, T., & Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *Journal of Neuroscience*, *27*(18), 4826-4831. https://doi.org/10.1523/Jneurosci.0400-07.2007
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. https://doi.org/10.1038/nature02581
Sharp, P. B., Russek, E. M., Huys, Q. J. M., Dolan, R. J., & Eldar, E. (2022). Humans perseverate on punishment avoidance goals in multigoal reinforcement learning. *Elife*, *11*. https://doi.org/10.7554/eLife.74402

Shear, M. K., Bjelland, I., Beesdo, K., Gloster, A. T., & Wittchen, H. U. (2007). Supplementary dimensional assessment in anxiety disorders. *International Journal of Methods in Psychiatric Research*, *16*, S52-S64. https://doi.org/10.1002/mpr.215

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass Correlations - Uses in Assessing Rater Reliability. *Psychological Bulletin*, *86*(2), 420-428. https://doi.org/Doi 10.1037/0033-2909.86.2.420

Sierra-Mercado, D., Deckersbach, T., Arulpragasam, A. R., Chou, T. N., Rodman, A. M., Duffy, A., McDonald, E. J., Eckhardt, C. A., Corse, A. K., Kaur, N., Eskandar, E. N., & Dougherty, D. D. (2015). Decision making in avoidance-reward conflict: a paradigm for non-human primates and humans. *Brain Structure & Function*, *220*(5), 2509-2517. https://doi.org/10.1007/s00429-014-0796-7

Simon, N. W., Gilbert, R. J., Mayse, J. D., Bizon, J. L., & Setlow, B. (2009). Balancing Risk and Reward: A Rat Model of Risky Decision Making. *Neuropsychopharmacology*, *34*(10), 2208-2217. https://doi.org/10.1038/npp.2009.48

Smith, R., Kirlic, N., Stewart, J. L., Touthang, J., Kuplicki, R., Khalsa, S. S., Feinstein, J., Paulus, M. P., & Aupperle, R. L. (2021). Greater decision uncertainty characterizes a transdiagnostic patient sample during approach-avoidance conflict: a computational modelling approach. *Journal of Psychiatry & Neuroscience*, *46*(1), E74-E87. https://doi.org/10.1503/jpn.200032

Spielberger, C. D., Gonzalez-Reigosa, F., Martinez-Urrutia, A., Natalicio, L. F., & Natalicio, D. S. (1971). The state-trait anxiety inventory. *Revista Interamericana de Psicologia/Interamerican journal of psychology*, *5*(3 & 4).

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder - The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092-1097. https://doi.org/DOI 10.1001/archinte.166.10.1092

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, *20*(11), 1643-+. https://doi.org/10.1038/nn.4650

Stan Development Team. (2023). *RStan: the R interface to Stan.* In https://mc-stan.org/

Stein, M. B., & Paulus, M. P. (2009). Imbalance of Approach and Avoidance: The Yin and Yang of Anxiety Disorders. *Biological Psychiatry*, *66*(12), 1072-1074. https://doi.org/10.1016/j.biopsych.2009.09.023

Stewart, A. M., Nguyen, M., Poudel, M. K., Warnick, J. E., Echevarria, D. J., Beaton, E. A., Song, C., & Kalueff, A. V. (2015). The failure of anxiolytic therapies in early clinical trials: what needs to be done. *Expert Opinion on Investigational Drugs*, *24*(4), 543-556. https://doi.org/10.1517/13543784.2015.1019063

Strehl, A. L., & Littman, M. L. (2008). An analysis of model-based Interval Estimation for Markov Decision Processes. *Journal of Computer and System Sciences*, *74*(8), 1309-1331. https://doi.org/10.1016/j.jcss.2007.08.009

Strimbu, K., & Tavel, J. A. (2010). What are biomarkers? *Current Opinion in Hiv and Aids*, *5*(6), 463-466. https://doi.org/10.1097/COH.0b013e32833ed177

Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction, 2nd Edition. *Reinforcement Learning: An Introduction, 2nd Edition*, 1-526. <Go to ISI>://WOS:000481873900019

Talmi, D., Dayan, P., Kiebel, S. J., Frith, C. D., & Dolan, R. J. (2009). How Humans Integrate the Prospects of Pain and Reward during Choice. *Journal of Neuroscience*, *29*(46), 14617-14626. https://doi.org/10.1523/Jneurosci.2026-09.2009

Tang, H. R., Houthooft, R., Foote, D., Stooke, A., Chen, X., Duan, Y., Schulman, J., De Turck, F., & Abbeel, P. (2017). #Exploration: A Study of Count-Based Exploration for Deep Reinforcement Learning. *Advances in Neural Information Processing Systems 30 (Nips 2017)*, *30*. <Go to ISI>://WOS:000452649402078

The Lancet Global, H. (2020). Mental health matters. *Lancet Glob Health*, *8*(11), e1352. https://doi.org/10.1016/s2214-109x(20)30432-0

The MathWorks Inc. (2015). *MATLAB (R2015b)*. In (Version 8.6.0.267246) The MathWorks Inc. https://www.mathworks.com

Tolman, E. C. (1948). Cognitive Maps in Rats and Men. *Psychological Review*, *55*(4), 189-208. https://doi.org/DOI 10.1037/h0061626

Treit, D., Engin, E., & McEown, K. (2010). Animal Models of Anxiety and Anxiolytic Drug Action. *Behavioral Neurobiology of Anxiety and Its Treatment*, *2*, 121-160. https://doi.org/10.1007/7854_2009_17

Valton, V., Wise, T., & Robinson, O. J. (2020). Recommendations for Bayesian hierarchical model specifications for case-control studies in mental health. *arXiv preprint arXiv:2011.01725*.

van den Bos, R., Koot, S., & de Visser, L. (2014). A rodent version of the Iowa Gambling Task: 7 years of progress. *Frontiers in Psychology*, *5*. https://doi.org/ARTN 203

10.3389/fpsyg.2014.00203

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413-1432. https://doi.org/10.1007/s11222-016-9696-4

Vogel, J. R., Beer, B., & Clody, D. E. (1971). A simple and reliable conflict procedure for testing anti-anxiety agents. *Psychopharmacologia*, *21*(1), 1-7. https://doi.org/10.1007/BF00403989

Vogel, S., & Schwabe, L. (2019). Stress, aggression, and the balance of approach and avoidance. *Psychoneuroendocrinology*, *103*, 137-146. https://doi.org/10.1016/j.psyneuen.2019.01.020

Vos, T. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019 (vol 396, pg 1204, 2020). *Lancet*, *396*(10262), 1562-1562. <Go to ISI>://WOS:000616720800003

Waltmann, M., Schlagenhauf, F., & Deserno, L. (2022). Sufficient reliability of the behavioral and computational readouts of a probabilistic reversal learning task. *Behavior research methods*. https://doi.org/10.3758/s13428-021-01739-7

Walz, N., Muhlberger, A., & Pauli, P. (2016). A Human Open Field Test Reveals Thigmotaxis Related to Agoraphobic Fear. *Biological Psychiatry*, *80*(5), 390-397. https://doi.org/10.1016/j.biopsych.2015.12.016

Wang, O., Lee, S. W., O'Doherty, J., Seymour, B., & Yoshida, W. (2018). Model-based and model-free pain avoidance learning. *Brain Neurosci Adv*, *2*, 2398212818772964. https://doi.org/10.1177/2398212818772964

Watkins, C. J. C. H., & Dayan, P. (1992). Q-Learning. *Machine Learning*, *8*(3-4), 279-292. https://doi.org/DOI 10.1007/bf00992698

Willner, P. (1984). The Validity of Animal-Models of Depression. *Psychopharmacology*, *83*(1), 1-16. https://doi.org/Doi 10.1007/Bf00427414

Willner, P. (1986). Validation criteria for animal models of human mental disorders: learned helplessness as a paradigm case. *Prog Neuropsychopharmacol Biol Psychiatry*, *10*(6), 677-690. https://doi.org/10.1016/0278-5846(86)90051-5

Wilson, M. (1993). DSM-III and the Transformation of American Psychiatry - A History. *American Journal of Psychiatry*, *150*(3), 399-410. <Go to ISI>://WOS:A1993KN84500004

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans Use Directed and Random Exploration to Solve the Explore-Exploit Dilemma. *Journal of Experimental Psychology-General*, *143*(6), 2074-2081. https://doi.org/10.1037/a0038199

Wise, T., Charpentier, C. J., Dayan, P., & Mobbs, D. (2023). Interactive cognitive maps support flexible behavior under threat. *Cell Reports*, *42*(8). https://doi.org/ARTN 113008

10.1016/j.celrep.2023.113008

Wise, T., & Dolan, R. J. (2020). Associations between aversive learning processes and transdiagnostic psychiatric symptoms in a general population sample. *Nature Communications*, *11*(1). https://doi.org/ARTN 4179

10.1038/s41467-020-17977-w

Wise, T., Emery, K., & Radulescu, A. (2024). Naturalistic reinforcement learning. *Trends in Cognitive Sciences*, *28*(2), 144-158. https://doi.org/10.1016/j.tics.2023.08.016

Wise, T., Liu, Y. Z., Chowdhury, F., & Dolan, R. J. (2021). Model-based aversive learning in humans is supported by preferential task state reactivation. *Science Advances*, *7*(31). https://doi.org/ARTN eabf9616

10.1126/sciadv.abf9616

Wise, T., Robinson, O. J., & Gillan, C. M. (2022). Identifying Transdiagnostic Mechanisms in Mental Health Using Computational Factor Modeling. *Biol Psychiatry*. https://doi.org/10.1016/j.biopsych.2022.09.034

Wise, T., Sookud, S., Michelini, G., & Mobbs, D. (2024). Transdiagnostic mental health symptom dimensions predict use of flexible model-based inference in complex environments.

Woods, K. J. P., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention Perception & Psychophysics*, *79*(7), 2064-2072. https://doi.org/10.3758/s13414-017-1361-2

Yamamori, Y., & Robinson, O. J. (2023). Computational perspectives on human fear and anxiety. *Neuroscience and Biobehavioral Reviews*, *144*. https://doi.org/ARTN 104959

10.1016/j.neubiorev.2022.104959

Yamamori, Y., & Robinson, O. J. (2024). Thinking computationally in translational psychiatry. A commentary on Neville et al. (2024). *Cogn Affect Behav Neurosci*. https://doi.org/10.3758/s13415-024-01172-1

Yamamori, Y., Robinson, O. J., & Roiser, J. (2023). Approach-avoidance reinforcement learning as a computational and translational model of anxiety-related avoidance. *bioRxiv*, 2023.2004.2004.535526.

Zorowitz, S., Momennejad, I., & Daw, N. D. (2020). Anxiety, avoidance, and sequential evaluation. *Comput Psychiatr*, *4*. https://doi.org/10.1162/cpsy_a_00026